University of Stuttgart
Institute of Visualization and Interactive Systems (VIS)
Department of Human Computer Interaction

Master Thesis Number:

Mental Task Classification using Single Electrode Brain Computer Interfaces

Master Thesis

AUTHOR: Mariam Hassib

SUPERVISORS: Alireza Sahami
Niels Henze

EXAMINER: Prof. Dr. Albrecht Schmidt

1 NOVEMBER 2012

I hereby declare that I have created this work completely on my own and used no other sources or tools than the ones listed, and that I have marked any citations accordingly.

$STUTTGART, MONTH\ YEAR$
Mariam Hassib

Contents

List of Figures

List of Tables

List of Abbreviations

ADHD .... Attention deficit hyperactivity disorder
ANN .... Artificial Neural Network
API .... Application Programming Interface
BCI .... Brain Computer Interface
CNS .... Central Nervous System
CSV .... Comma-Separated Values
ECG .... Electrocardiogram
ECoG .... Electrocorticography
EEG .... Electroencephalography
ERP .... Event Related Potential
FFT .... Fast Finite Fourier Transform
fMRI .... Functional Magnetic Resonance Imaging
FNIRS .... Near-infrared spectroscopy
GUI .... Graphical User Interface
HCI .... Human Computer Interaction
ICA .... Independent Component Analysis
MLPNN .... Multi-Layer Perceptron Neural Network
NFT .... Neurofeedback Training
PCA .... Principal Component Analysis
SNR .... Signal to Noise Ratio
SSVEP .... Steady State Visual Evoked Potential
SVM .... Suuport Vector Machines
VEP .... Visual Evoked Potential

Abstract

In the recent years, the field of Human-Computer Interaction (HCI) has greatly evolved to involve new and exciting interaction paradigms that allow users to interact with their environment and with technology in a more intuitive and ergonomic way. These interaction paradigms include voice, touch, virtual reality, and more recently, the brain.

A brain-computer interface (BCI) is a an interface system allowing users to control devices without using the normal output pathways of peripherals, instead, by using neural activity generated in the brain. BCIs have a huge potential in a multitude of fields, all the way from providing users with severe motor disabilities with means for interaction with the external world, to entertainment, gaming, user state monitoring, and self tracking systems. The potentials of BCI have sparked the interest of researchers, gaming markets and healthcare providers more and more in the recent years. The is due to the emergence of new commercial lightweight, low cost Electroencephalograph (EEG) equipment that made it possible to create more portable and usable BCI systems and expanded their fields of application.

This Master thesis aims to explore the state of the art commercial BCI as well as the uses and challenges related to them. Commercially available EEG equipment, namely the Neurosky Brainband and Neurosky Mindset, will be investigated. User tests will be carried out to investigate whether such equipment with low accuracy and low cost can be used to recognize various mental activities. This would be performed by first collecting a dataset of brain signals during performing a set of mental tasks, applying a set of signal processing algorithms, then exploring various classification techniques to classify the collected signals.

Part I

Introduction and Motivation

Chapter 1

Introduction

Human Computer Interaction (HCI) has been a constantly evolving field in the recent years especially with the inclusion of new interaction paradigms including touch, voice and even virtual reality interfaces. A relatively new addition to these paradigms is the brain. The ability to record electric potentials from the brain and translate them into real life commands to which a system would respond has recently been a growing interest for researchers. In addition, this newfound paradigm adds new dimensions to HCI. BCI, short for *Brain Computer Interfaces*, carries a lot of hope and potential for a very wide spectrum of scientists and researchers. It is going all the way from providing a better quality of life for people with severe motor disabilities Kaur et al. [2012] to being the new fashion in the gaming and entertainment industry. In both cases, undoubtedly accompanied by many challenges, new findings and even more evolving fields of research and potentials will be appearing.

Using brain signals in the field of HCI has a wide paradigm of application possibilities for both disabled and healthy users

This Master Thesis aims to explore new emergent BCI interfaces using off-the-shelf EEG devices, and their uses especially in the field of mental task recognition.

1.1 What is a Brain Computer Interface (BCI)?

EEG is the only non-invasive, portable and affordable way to collect brain signals

A brain-computer interface is a an interface system allowing users to control devices without using the normal output pathways of peripherals, instead, by using neural activity generated by the brain. The neural activity used in BCI can be recorded using invasive or non-invasive techniques. These techniques include ElectroEncephaloGraphy (EEG), MagnetoEncephaloGraphy (MEG) functional Magnetic Resonance Imaging (fMRI), functional Near InfraRed Spectroscopy (fNIRS), ElectroCorticoGraphy (ECoG) and Subcortical Electrode Arrays (SEA) [van Erp et al., 2012]. The cost, effort, invasiveness and quality of these varying techniques can be summed up in table 1.1

Technology	Disadvantage
Electrocorticogram (ECoG)	Highly invasive, surgery
Magneto-encephalography (MEG)	Extremely expensive
Computed Tomography (CT)	Only anatomical data
Single Photon Emission Computed Tomography(SPECT)	Radiation exposure
Positron Emission Tomography (PET)	Radiation exposure
Magnetic Resonance Imaging (MRI)	Only anatomical data
Functional Magnetic Resonance Imaging (fMRI)	Extremely expensive
Functional Near-Infrared (fNIRS)	Still in infancy,expensive

Table 1.1: Brain signal recording technologies and their disadvantages Lee and Tan [2006]

The components of a typical BCI system: Electrode Cap, Amplifier, Computer

Figure 1.1 shows the different components of a basic EEG-based BCI system, which is of the non-invasive brain sensing technologies. The system is composed of: an electrode cap fitted with electrodes connected to an amplifier that amplifies the captured signal and converts it from analog to digital. The cap is connected via wired or wireless connections to a computing device to process the captured signal. Other BCI systems mentioned in table 1.1 have different compoenents. However, the work presented here is based on such EEG BCI system, the components of which are discussed in chapter 4.

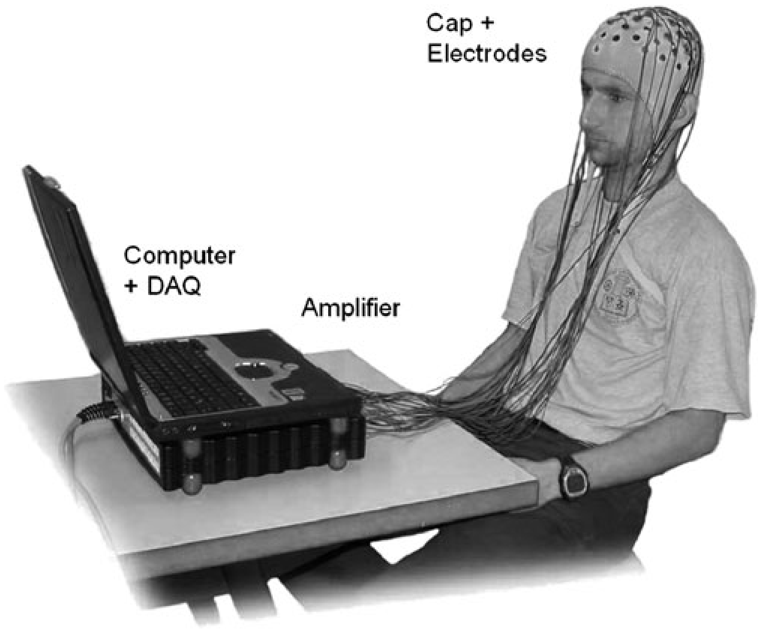


Figure 1.1: Components of a typical BCI system [Graimann et al., 2011]

Until recently, BCI systems have been mostly a research interest in the medical and clinical world. This is due to their obvious potential in enhancing the lives of people with severe motor disabilities and in the hope of using them in the diagnosis and treatment of brain disorders like Alzheimer's , Attention Deficit Hyperactivity Disorder (ADHD), epilepsy, and other complex disorders. The size and complexity of existing BCI systems in the past decades have made them only suitable for lab experiments and research. However, this fact is starting to change due to the availibility and development of low-cost, portable, comsumer-oriented and non-invasive BCIs.

Clinical uses of BCI include diagnosis of ADHD, Alzheimer's and epilepsy

1.2 Commercial BCIs and Their Challenges

In the recent years, BCI systems have taken a leap into the commercial world. The size, portability, signal-aquisition techniques, and cost of equipment evolved greatly, making

NeuroSky and Emotiv are the first commercially available and affordable BCI devices it possible to finally take BCIs out of the lab and into the daily life of normal healthy individuals, and companies like Neurosky[1] and Emotiv[2] have started gaining popularity in both the entertainment market and among researchers. Both companies provide fairly inexpensive hardware for obtaining EEG signals in normal environments outside the restricted lab conditions, hence allowing for new fields of application to emerge.

BCI open source software facilitate designing BCI systems and support complex signal processing of the various types of brain signals

The availibility of open source BCI processing software has also made it possible for the emergence of those types of interfaces. Some of the most widely spread tools are BCI2000[3], OpenVibe[4], and EEGLab[5] which is an open source Matlab plugin. Such tools encouraged developers and research communities to enter the field of BCI with what they offer from signal processing techniques to even designing and simulating complete BCI systems.

While the use of commercial BCI is still limited to research and gaming right now, according to van Erp et al. [2012], non-medical BCIs will evolve in the following areas with varying degrees of societal impact and success in the upcoming years:

Device Control

- Controlling a wheelchair or an artificial limb by translating brain signals into control commands.

User State Monitoring

- Gathering information about a person's state including alertness, fatigue and workload. This can be employed in applications such as monitoring driver alertness.

Evaluation

- Offline and online evaluation of user brain response while doing certain tasks. An example area of usage would be neuromarketing, which is concerned with studying brain response to advertisements, hence identifying which advertisement has the highest impact.

Training and Education

- Many educational possibilities of BCI already exist, in

[1] http://www.neurosky.com
[2] http://www.emotiv.com
[3] www.bci2000.org
[4] http://openvibe.inria.fr
[5] http://sccn.ucsd.edu/eeglab/

which the children's capabilities to be attentive or to do mathematical tasks are enhanced.

- Games tailored to adapt according to the user's mental state increase the feeling of immersion in the virtual gaming envoronment. This field is proving to be successful already.

Entertainment and Gaming

- Providing neurofeedback training for healthy, recovering, or rehabilitating users, in order to improve their attention, working memory or executive functions.

Cognitive Improvement

- A niche field still in the concept development phase. Areas of application include lie detection from EEG signals, for which EEGs are yet not reliable enough to practically consider.

Safety and Security

Although the application areas of off-the-shelf BCI systems seem promising and encouraging, a lot of challenges deter the fast prevalence of this technology in the market. These challenges are not crucial when looking at BCIs from the clinical perspective in medical applications developed in strict lab conditions. However, for market applications targetted towards use in the everyday life, many challenges prevail. A list of those challenges is presented and discussed below:

Properly addressing the usability challenges of BCI will help in its advancement

- The ease of use of system without prior training is a definite user expectation when it comes to applications in the market. A BCI device requiring too long prior preparation and application of gel on the head for fixation is a tedious task for users in [van Erp et al., 2012].

Usability

- The size, weight and comfort of the BCI device is a crucial challenge. Usually BCI devices have to be fitted on the head using an electrode cap (in case of EEG and non-invasive devices). The comfort of the device when worn for long periods of time is also an important aspect.

Portability of hardware

- BCI systems offering high-quality signal aquisition techniques with noise and artifact removal modules were, for the past decade, very expensive. However, with the emergence of companies like NeuroSky and

Cost

Emotiv, the possibility for reduced cost off-the-shelf BCI devices increased tremendously.

Signal Processing

- Brain signals are relatively very weak, hence extensive and complex signal processing and artifact removal needs to be done prior to using those signals. Some commercial devices offer simple processing on chip, however, much more processing needs to be done on a remote processor.

1.3 Thesis Structure

Part 1: Introduction and Motivation

This thesis is organized into 4 main parts and is presented in book format. Part 1, 'Introduction and Motivation' includes two chapters: Introduction1 which introduced BCI as a new paradigm and the challenges related to it, and the Motivation2, which will explain the uses of activity recognition and introduced the research questions as well as the intended contributions of this thesis.

Part 2: Background and Related Work

Part 2, 'Background and Related Work' , contains 3 chapters. Chapter 3 disusses the biomedical background behind BCIs and how the brain works, chapter 4 introduces the blocks of a BCI system, the theoretical background behind each block and introduces the hardware and software technical background of BCIs. Finally chapter 5 discusses previous related work in the fields of mental task recognition and commercial BCI devices.

Part 3: Mental Task Classification

Part 3, 'Mental Task Classification' is a thorough discussion of the technical work done during this thesis. It is divided into 3 main chapters. Chapter 6, discusses the system design and implementaion aspects. Chapter 7 discusses data classification and evaluation. Finally, chapter 8 discusses the limitations, challenges and shortcomings of the system implemented.

Part 4: Conclusion and Future Work

The final part 4, 'Conclusion and Future Work', includes 2 chapters. Chapter 9 concludes the results mentioned in the previous chapters, and chapter **??** presents ideas about how the work can be extended and built on to be enhanced.

Chapter 2

Motivation for Mental Task Classification

While BCIs move towards diminishing the need for motor movements to control computers, the biggest gain of brain interfaces would be extracting the information such interfaces provide about the state of the user. By using BCI systems to classify the different mental and cognitive activities of the brain, a wide range of applications can be realized for the enhancement of the lives of both healthy and physically impaired people. The possibility of translating thoughts into commands, through which one can control a device, has been a science fiction myth for decades, maybe even eras. Yet with BCI systems, it is now a reality.

The full potential of the brain as an input channel lies in the rich information it provides about the state of the user

The application for mental task classification via BCI systems is discussed in this chapter covering different aspects and target groups who can benefit from this techonology.

2.1 Communication and Control

People suffering from motor disabilities due to accidents, strokes, neuro-degenerative diseases such as Amyotrophic Lateral Sclerosis (ALS: a disease which eventually strips an individual of all voluntary muscular activities, while

Beneficiaries of this technology include patients of ALS and paralysis

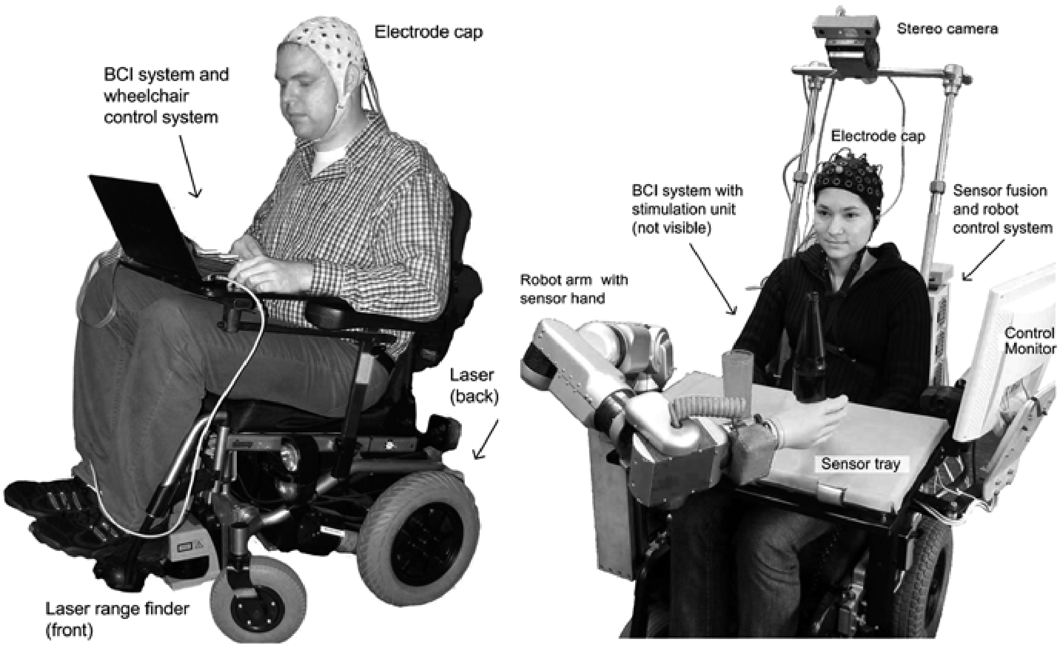


Figure 2.1: Semi-autonomous wheelchairs developed at the University of Bremen operating via EEG signals aquired from an electrode cap [Graimann et al., 2011]

leaving cognitive function intac.), or any other reason, are those benefiting the most from this type of technology.

BCI systems can be used to generate control signals via EEG analysis and classification

Using BCI systems to translate certain mental tasks into control signals to move a wheelchair, call a person, or signal for a cursor to move on the computer screen, would enhance the quality of life for patients suffering motor disabilities and increases their independence that has been lost due to their disability. Successful advances in this area during the past decade have been very useful to the medical community particularly. Figure 2.1 shows semi-autonomous assistive devices developed at the University of Bremen that include high level control for patients with motor disabilities: Intelligent wheelchair Rolland III, and rehabilitation robot FRIEND II.

Examples of easily distinguished mental tasks include motor imagery and mental math

The success of such systems depends on the high quality signal aquisition and processing equipment, as well as on the tasks chosen to train the system. Mainly, tasks chosen to be classified are the ones that a user can be easily trained to modulate, and the system can easily distinguish.

Among those mental tasks are: imagined hand movement, mathematical tasks, or mind spelling through special signal detection which will be discussed in detail in Chapter 3, and more examples of similar work will be illustrated in Chapter 5.

2.2 Neurofeedback and Rehabilitation

Another use of mental task classification is in Neurofeedback Training (NFT). NFT is a type of biofeedback that uses realtime displays of electroencephalography (EEG) or functional magnetic resonance imaging (fMRI) to illustrate brain activity, often with a goal of controlling central nervous system (CNS) activity and training certain brain functions. Sensors are placed on the scalp to measure brain activity, while users are engaged in mental tasks, with measurements displayed using video displays or sound.

Neurofeedback training is a way to train the brain to function at its maximum potential

This type of training can be used in a number of applications including training of ADHD patients to control their concentration levels, discovering reading problems in children [Mostow and Beck, 2007], or training school children to concentrate in mathemetical tasks. Neurofeedback training also has potential in treating autism, headaches and sleep disorders by providing users with feedback about their sleeping patterns.

NFT has a lot of potential in better understanding brain disorder and children learning difficulties

In addition to the aforementioned neurofeedback applications, the use of BCI for brain neurorehabilitation has been proposed. Studies targetting patients who suffered from strokes and have destroyed motor functions show that using a BCI system can be helpful in restoring the lost functionalities. The BCI system detects motor imagery and reinforces the use of damaged neural pathways. Based on these studies, positive rehabilitation training directed by motor intention has a better efficiency than conventional passive training [Graimann et al., 2011].

Neurorehabilitation of stroke patients to aid in motor function restoration is possible using BCIs

2.3 Mental State Tracking

Smart gadgets are now used for the rising trend of self tracking

While the previous two sections were concerned with medical aspects which can benefit from BCI mental task classifcation, this section shows how healthy users can utilize this technology as well. With the prevalence of highly advanced gadgets (smart phones, tablets, notebooks, readers and personal compters), personal tracking of physical activity, emotional state, calorie intake, and a wide variety of information is a current trend of interest. This is made possible via the many sensors provided in smart phones (gyroscopes, light sensors, etc..) and data internet packages.

Self tracking tools promote better understanding of oneself

The rise in the personal information tracking trend led to the development of many applications and tools facilitating self tracking and communities sharing the same interest. An example of this is the Quantified Self [1] platform, which represents a collaboration of users and tools which promote personal tracking. This platform is now supporting some BCI applications offered by companies offering off-the-shelf BCI gadgets like NeuroSky [2]. BCI applications detecting mental states can give feedback to the user about his/her stress levels, concentration, relaxation, fatigue, and emotional state. Tracking the mental state can give one more insight into their trends and can allow people to better understand themselves and organize their lives in more efficient ways.

Mental state tracking can be used in designing user interfaces and improving work flows

In addition to personal self tracking, mental task classification can also be used in improving industrial production environments. For example, recognition of the levels of arousal, fatigue, emotion, workload or other variables relating to the user's mental state and can be recorded, can be used in designing better car interfaces adapting to the driver's mental state [Graimann et al., 2011].

[1] `http://quantifiedself.com/`
[2] `http://neurosky.com/`

There are many others possibilities for uses and applications of mental task classification using BCI; the previous sections only aimed to give a brief overview. The project presented in this thesis is concerned with classifying mental tasks using portable, consumer-oriented, off-the-shelf BCI systems. This research is an attempt to discover the feasibility of using such equipment in recognizing different cognitive tasks providing feedback for the user about their personal trends. This project, being one of the first attempts to classify mental tasks using commercial BCI, also aims to collect a large data corpus of brain signal data to pave the road for future investigation and research in similar areas. The results of this research can be used to later provide applications of many types including personal mental state tracking for healthy users as well as affordable NFT applications for developing cognitive skills in children or patients of mental disorders.

This work presents an attempt of mental task recognition using commercially available BCI

2.4 Research Questions and Contributions

This section introduces the research questions that this Master Thesis aims to answer, and the contributions it aims to offer. To begin with, the opportunities and challenges of commercially available BCI systems in the field of mental activity recognition will be discovered. This work intends to explore the full potential of NeuroSky's single-electrode BCI devices. The research questions tackled throughout this thesis are presented in the list below:

Mental task classification using single-electrode BCI is a yet undiscovered field that this work targets to explore

PROPOSED RESEARCH QUESTIONS:

1. Can low-cost commercial BCI be used to classify mental tasks?
2. What information can be extracted from single dry electrode BCI systems?
3. What are the potentials and challenges of commercial BCI?

The main contributions this thesis will offer include building a data corpus of unfiltered EEG data from 15 subjects using NeuroSky's MindSet and Brainband devices, as well as a filtered, artifact-free corpus. No similar dataset has been previously collected. In addition, this work will contribute to mental task recognition research by investigating the possibility of tracking mental and cognitive activities using off-the-shelf BCI, which can be the basis of building many activity tracking applications on top of this work's findings.

This works also provides a thorough analysis of the advantages and limitations of off-the-shelf EEG devices and proposes ideas for future work and enhancements to the system. In short, this thesis largly contributes in laying the basic building blocks for using new commercially available BCI systems beyond the entertainment industry.

Part II

Background and Related Work

Chapter 3

Biomedical Background

This chapter discusses the biomedical background behind BCI systems. It gives an overview about the structure of the brain, as well as the different types of brain signals that are collected and utilized by BCIs.

3.1 The Brain

The brain is the center of the nervous system and is considered, by far, the most complex organ in the human body. It is also the organ responsible for centralized control over all other organs of the body.

Functions of the frontal, occipital, temporal and parietal lobes of the cerebral cortex

The brain is composed of three main parts: The cerebrum, cerebellum and the medulla (the brain stem). The cerebral cortex is the largest part of the brain and it is associated with all the vital functions. Figure 3.1 shows the 4 different lobes of the cerebral cortex and their functions. The frontal lobe controls mood, concentration, decision-making and planning activities. The occipital lobe is responsible for processing visual information. The temporal lobe is associated with hearing, memory and language functions, and finally, the parietal lobe is associated with taste, touch, and movement coming from the rest of the body [HealthPage.org, Jun 2010]. Although each of those cerebral lobes is associated with a group of functionalities and senses, all of them work in conjunction with one another. Taking 'reading a book' as an example, we would see that the occipital lobe is actively working to process visual information such as words, leters and illustrations in the book. The frontal lobe is also actively working in understanding the meaning of the words being read. Finally, the temporal lobe is also working on processing all the surround sound [Wren, 2001].

The BCI system that is implemented and researched in this thesis uses electric potentials generated mainly from the frontal lobe.

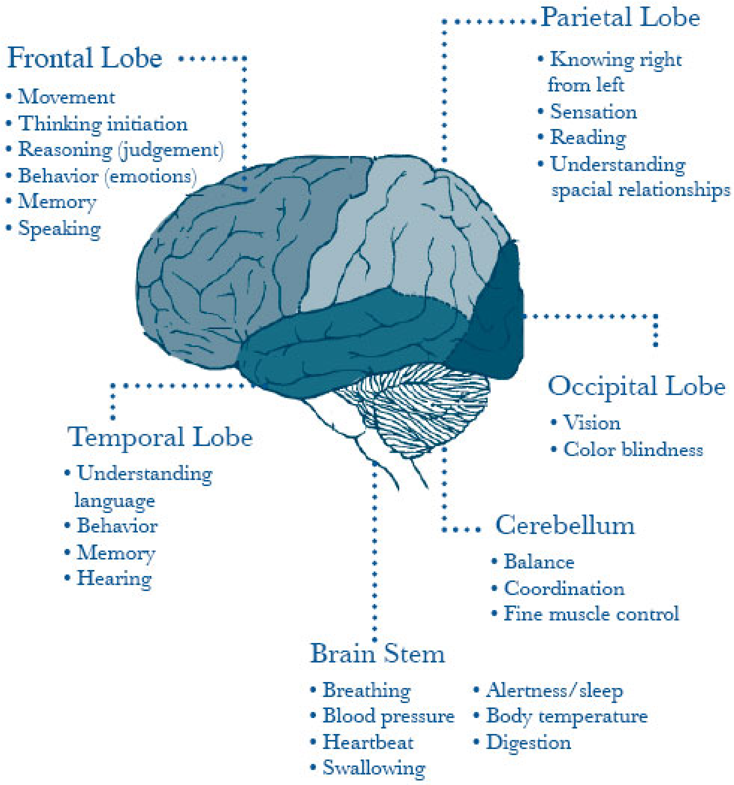


Figure 3.1: The 4 Lobes of the Cerebral Cortex [HealthPage.org, Jun 2010]

The brain is composed of billions of neurons

Taking a closer look at the cellular structure of the brain we find that it is composed of 2 main broad types of cells: neurons and glial cells. Glial cells are used in many critical

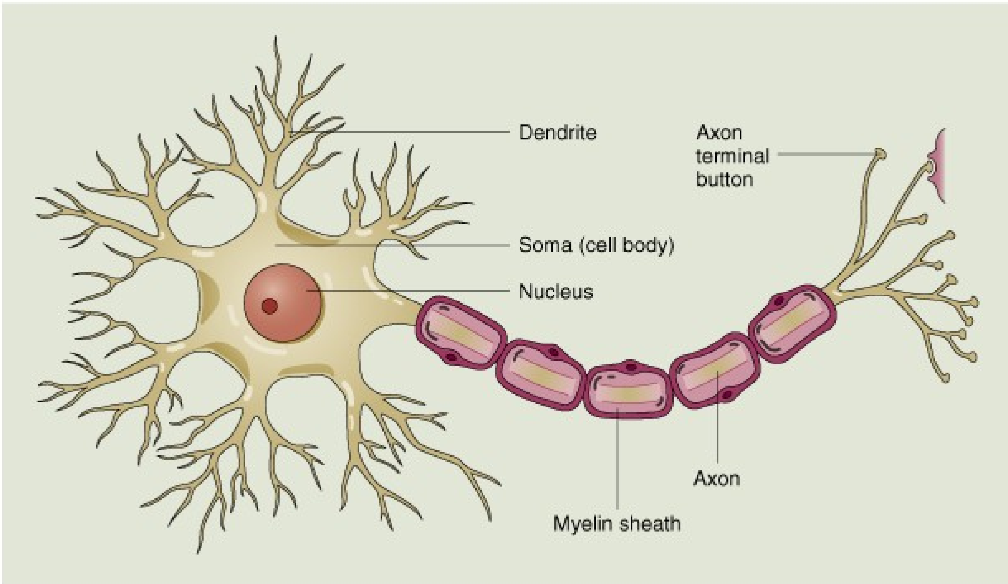


Figure 3.2: The Neuron [of Queensland Inc., 2004 (accessed October 19, 2012]

functions including the structural and metabolic support of the brain. On the other hand, neurons, which are the main propeller behind the idea of BCIs, are electrically excitable cells that transmit information by electrical and chemical signaling. Each neuron is composed of an axon, a cell body and dentrites as shown in figure 3.2.The brain is composed of billions of neurons connected to together to form a complex Neural Network which is the inspiration to the machine learning technique having the same name.

Neurons send information to other neurons through *synapses*

The main functionality of neurons is to receive input information from other neurons in the network, process and send that information to other neurons through the 'Synapses', which are connection between neurons. The term 'information' refers to motor, sensory or cognitive information that needs to be sent between the brain and other organs of the body [Stufflebeam, Jan 2002]. An action potential, also called a 'spike' or 'impulse' occurs when information moves from the cell body of a neuron, through its axon until it reaches the axon terminals and is then transmitted to the billions of neighboring neurons forming the neural network. The definition of the action potential is presented below.

Definition:
Action Potential

ACTION POTENTIAL:
An electrical signal generated near the cell body of a neuron that propagates along the axon to the axon terminals

Measuring electric signals generated from the brain

As information is received and transmitted through the neurons, an electric field is generated due to changes in the membrane potentials of the synapses, neurons and axons. This electric field can be recorded in a number of ways including, if a large number of neurons are synchronized, a small voltage change on the scalp. Voltage changes on the scalp due to neuronal activity is called Electroencephalography (EEG) and it is the technology employed in the BCI discussed in this thesis. It is further discussed in detail in section 3.2. Other methods of measuring the neuronal electric field include: Electrocorticographic activity (ECog) measure on the cortical surface and Intracortical recordings measured within the brain cortical tissue [Graimann et al., 2011] which are illustrated together with EEG in figure 3.3.

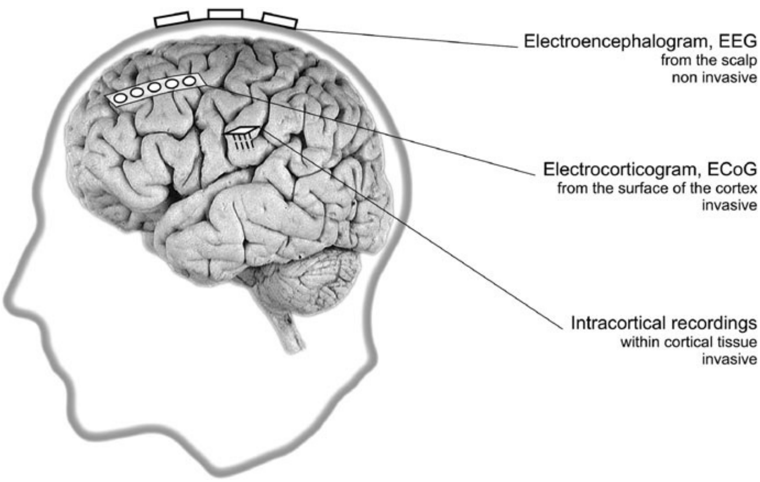


Figure 3.3: 3 examples of measuring electric potentials: EEG, ECog, and Intracrotical Recording [Graimann et al., 2011]

3.2 Electroencephalograph (EEG)

Temporal and spatial resolution of EEGs gives it an advantage of other brain sensing techniques

Although there exist already many non-invasive techniques for recording brain activity, Electroencephalographs (EEGs) are by far the most popular and widespread method of signal aquisition for BCIs. The high temporal resolution of EEGs, that are able to measure the brain signal at a rate less than 1 ms, gives EEGs an obvious advantage over other techniques of measuring brain activity. Modern EEGs also have a relatively high spatial resolution where up to 256 electrode sites can be measured at the same time [Vallabhaneni et al., 2005]. The EEG signal is recorded from the surface of the scalp via electrodes after the application of conductive gel or in some other cases using dry electrodes [NeuroSky, 2009]. EEG signal aquisition methods and electrode placement are further discussed in section 4.1. Figure 3.4 illustrates 14 EEG signals from 14 electrode locations.

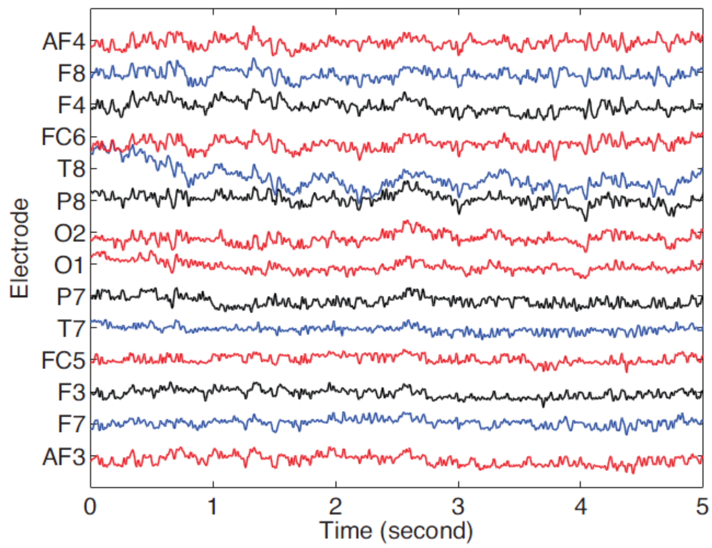


Figure 3.4: Raw EEG signal from 14 electrode locations on the y-axis [Campbell et al., 2010]

Properties and features of EEGs can be extracted from both the time and frequency domains

EEG signals are measured in the time domain as voltage levels in microvolts, and can also be analyzed in the frequency domain producing special waveforms and features discussed in section 3.2.1. Changes in the voltage levels can also be time-locked to particular stimuli called then 'Event-Related Potentials' (ERPs) which are discussed in section 3.2.2. Features extracted from measuring EEGs at both the time and frequency domains have proved useful for BCIs [Graimann et al., 2011] which will be seen later in the discussion of related work.

Thousands of neurons need to be synchronously fired to record an EEG

The human brain is actively generating electric signals at all times, even during rest or sleeping. However, most of this brain activity is invisible to EEGs. For the neuronal activity to be visible to EEGs, a large number of neurons need to be fired in near asynchrony and the electrical signals must be generated along a specific axis oriented perpendicular to the scalp [Smith, 2004].

3.2.1 EEG Frequency Bands

Delta, theta, alpha, beta and gamma frequency bands can be extracted from EEGs

Frequency domin analysis of EEGs provides a set of the most significant features distinguishing differences in brain electrical activity during different cognitive tasks. Signals recorded from EEG are split into several freqnuency bands by applying signal processing techniques such as the Fast Fourier Transform and band-pass filtering. Table 3.1 illustrates the different EEG output signals and the mental states they represent. As Table 3.1 illustrates, there are 5 different observed bands: 0.1-3.5 Hz (delta - δ), 4-7.5 Hz (theta - θ), 8-13 Hz (alpha - α), 14-30 Hz (beta - β) and $\geq$ 30 Hz (gamma - γ). Research about EEG frequency bands and their functionalities is ongoing.

Brainwave Type	Frequency Range	Mental State
Delta	0.1-3Hz	Deep dreamless sleep, unconcious
Theta	4-7Hz	Intuitive,creative,recall,fantasy,imaginary,dream
Alpha	8-12Hz	Relax but not drowsy,tranquil,concious
Low Beta	12-15Hz	Relaxed yet focussed, integrated
Midrange Beta	16-20Hz	Thinking, aware of self and surroundings
High Beta	21-30Hz	Alertness, agitation
Gamma	30-100Hz	Motor functions,higher mental activity

Table 3.1: EEG Brainwave Types and Mental States NeuroSky [2009]

3.2.2 Event Related Potentials (ERP)

Event-related potentials (ERPs), as the name suggests, refer to an electrical potential response to an external or internal event or stimulus that occurs after a fixed time interval. ERPs can occur due to auditory, visual or somatosensory (touch) stimuli.

Detecting ERP compoenents in EEGs is done offline

Since the amplitude of an ERP component in the EEG is much smaller than the spontaneous EEG (typically by a factor of 10) [Smith, 2004], detecting ERP signals from raw EEG in real-time is not possible. Hence, the EEG signal is averaged over fixed time epochs with a large number of trials in which the stimulus was introduced [Grierson, 2008]. Heavy analysis of the EEG signal is then required to extract the ERP component.

P300 signal is a common ERP component

A commonly studied ERP component is the P300 signal. P300 refers to a positive peak in the EEG signal after 300ms of an expected stimulus. A stimulation protocol of this kind is known as an *oddball* paradigm.

Definition: *Oddball Paradigm*

Oddball Paradigm:
A technique used in evoked potential research. Subjects are asked to identify an infrequently occuring target within a series of frequently occuring standard stimuli. Subjects are often asked to respond by pressing a button or counting the number of appearances of the target stimulus. This generates an ERP component in the EEG signal.

The P300 speller enables brain spelling

A common research application utilizing this signal is the *P300 Speller*. It is a virtual keyboard system consisting of characters displayed in a row-column matrix on a screen which are successively highlighted. When the line or the column containing the chosen letter flashes, a P300 ERP occurs. A classifier is then used to determine if this signal corresponds to a positive response or not [Kanoh et al., 2011]. Figure 3.5 shows the P300 speller.

VEP and SSVEP components occur upon visual stimuli

Another type of commonly studied ERPs are the Visual Evoked Potentials (VEP) and the Steady State Visual Evoked

Figure 3.5: The P300 Speller [Krusienski et al., 2006]

Potentials (SSVEP). Both of which are EEG compoenents that occur due to the subjection of a visual stimulus. VEPs depend on the users' ability to control their gaze. SSVEPs are a class of VEPs that require that the subject concentrates on two stimuli flickering at the alpha and beta frequencies on screen. An SSVEP compoenent is generated when the user's focus shifts from one object to the other [Vallabhaneni et al., 2005].

Chapter summary

This overview of the biomedical background behind BCIs briefly presents various concepts that are essential to understanding how EEGs and BCIs work. However, the study of biomedical aspects which make BCIs possible extends far beyond this chapter which only serves as an introduction to the work presented in this Master Thesis.

Chapter 4

BCI Systems

This chapter gives an overview of the basic building blocks of a BCI system, EEG aquisition, as well as the algorithms used in the analysis and classification of EEG signals. It provides an overview to algorithms related to the work presented in this thesis.

4.1 EEG Signal Aquisition

Figure 4.1 is an illustration of the basic operation of a BCI. The signal generated in the brain is aquired via one of the various invasive or non-invasive techniques discussed in the previous sections, then the signal is passed through a signal processing block where it is amplified and digitized for analysis and inspection. In the discussion throughout the rest of the this thesis, EEG BCIs would be the main focus.

Reocrding generated neuronal electrical activity at the scalp is done non-invasively using small round metal plate electrodes. Electrodes are typically fitted in an electrode cap which is worn by the subject to fix them at the correct positions on the scalp.

Two main types of electrodes are currently used in BCIs, wet and dry electrodes. Wet electrodes require the appli-

Wet electrodes require more setup time

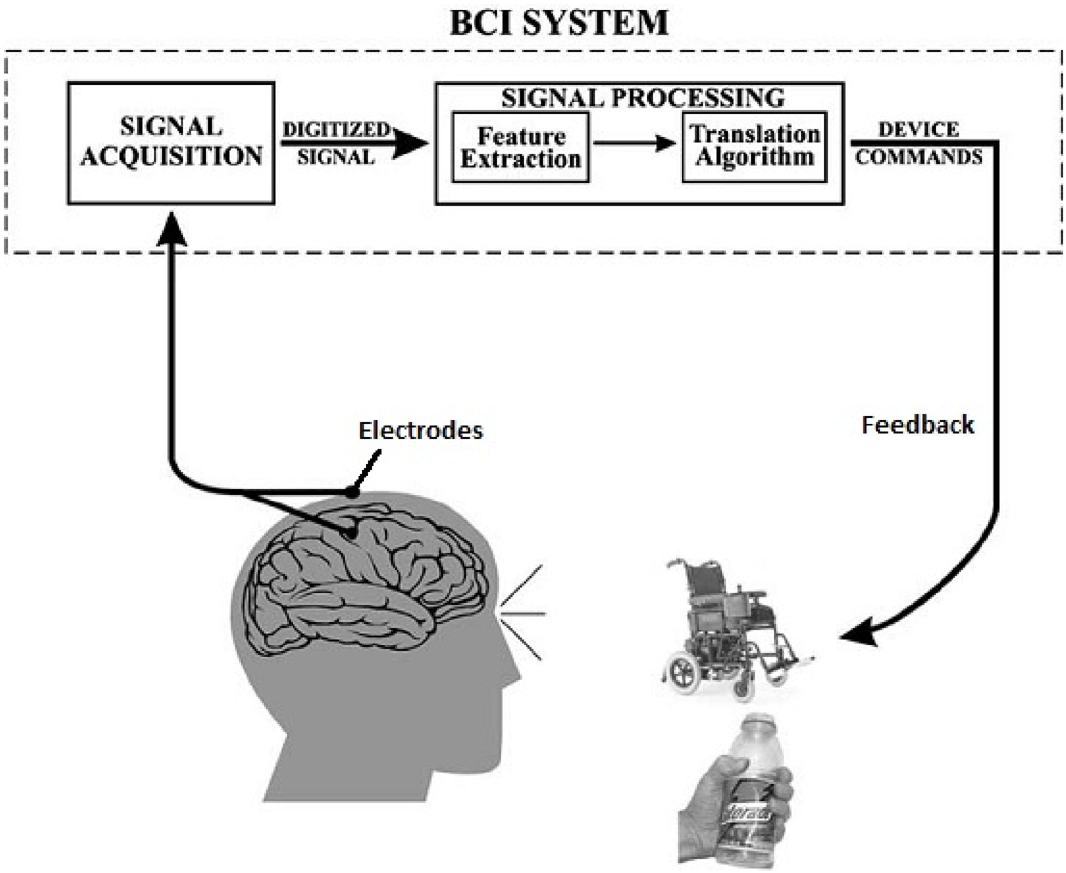


Figure 4.1: An illustration of the components of a BCI system

cation of conductive gels to the scalp prior to placing the electrodes in order to enhance conductivity and capture a better quality signal from the brain. The application of conductive media on the scalp has proven to be an inconvenient and time consuming experience to the subjects. Additionally, frequent application of gels increases the sensitivity of the scalp skin and in long hours of use of such a system regular maintenance is required to ensure that the gel does not dry out and reduce the quality of the captured signl [Saab et al., 2012]. The wet electrode technology has long been the more prevelent option for EEG signal recording despite their impracticality.

Dry electrodes are more practical to use in day-to-day applications

Setting up a wet electrode BCI system is time consuming and inconvenient for use in daily activities and are more suited for research done in laboratory setting. With the recent enhancements in BCIs, the need for more usable, practicle and portable hardware for EEG recording prevailed, hence making way for the dry electrode technology to develop. The BCI hardware used in the work discussed in this thesis is a single dry electrode device from Neurosky that has been simultaneously tested and compared to the wet electrode system Biopac [Biopac] widely used in research and medical applications. Both electrodes were placed very close to each other on the location of the Neurosky device electrode indicated in figure 4.2. The resulting signals were highly close to each other [NeuroSky, 2009].

4.1.1 Electrode Placement

Explaination of the 10-20 system conventions

The placement of electrodes , dry or wet, on the scalp is done based on an international method called the 10-20 system. The naming comes from the percentage distance between the location of electrodes, 10% or 20% of inter-electrode distance. Each electrode location on the scalp in the system has a name composed of a letter representing the underleying brain lobe, and a number ditinguishing the right and left brain hemispheres. The possible letters are F for Frontal, T for Temporal, P for parietal, O for Occipital and C for Central which is not a lobe, but only used for identification purposes. Even numbers refer to locations on the right hemisphere, whereas odd numbers refer to locations on the left brain hemisphere. The area identified as the scalp starts from the 'Nasion', which is the point between the forehead and the nose, and ends at the 'Inion', which is identified as a bump at the end of the skull [Inc., 2001].

After aquiring the generated electrical potential via wet or dry electrodes, the signal has to undergo several stages of processing before being able to extract useful information out of it. There are 3 main stages which are: signal preprocessing, feature extraction and selection, and finally classification. An overview of those 3 stages is given in the following sections.

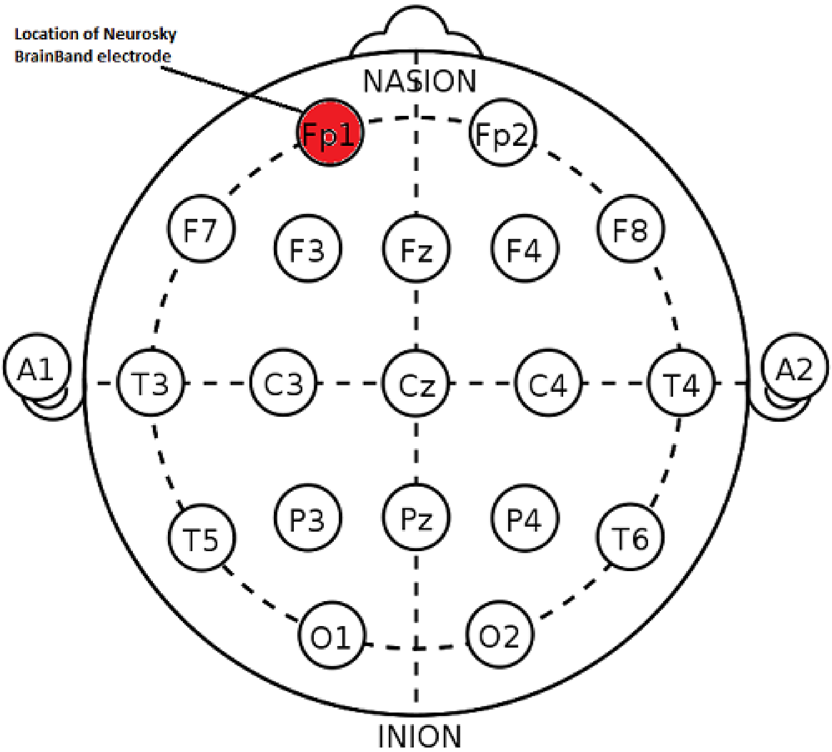


Figure 4.2: 10-20 Electrode Placement System : The location of the Neurosky electrode used in the work described in this thesis, Fp1 (Frontal polar - left hemisphere) is highlighted in red

4.2 EEG Signal Processing

EEG signals are very weak hence have to be amplified via a biosignal amplifier

Since the signal captured from the electrodes, dry or wet, is typically in the order of microvolts or even nanovolts at times, it first has to pass through an amplifier to bring it to a stuiable range for processing. EEG devices used in this project had an embedded amplification circuit, hence no manual amplification needed to be done.

4.2.1 Noise and Artifact removal

After the signal amplification step of the preprocessing stage comes the noise and artifact removal step. EEG recordings are often noisy due to being very weak signals and due to noise coming from the environment in which the recording was done. The main goal of this phase is to improve the quality of the captured signal by enhancing the Signal to Noise ratio. Smaller SNR means a noisier signal and a higher effort on the classification and feature extraction stage of the signal filtering [Graimann et al., 2011].

Physiological artifacts include sweating and muscle movements

Noise and artifacts in the signal can be generated from a number of sources which can be broadly divided into physiological and non-physiological sources. Physiological sources of artifacts include artifacts from muscle movements (Electromyographic (EMG) artifacts) , as well as artifacts from sweating which tamper with the impedence of the electrodes [Graimann et al., 2011]. Physiological artifacts from eye movements or blinking are called Electrocularographic artifacts (EOG). Figure 4.3 shows an EEG signal containing peaks from an EOG blinking artifact, removed using an adaptive filtering technique in the bottom figures. Potentials related to cardiac activity (Electrocardiogram (ECG) pulses) can also be seen in electrodes placed on the left side of the brain on pulsating vessles. This can be especially noted in overweight subjects.

Non-physiological artifacts include electric equipment interference and wrong positioning

Non-physiologic artifacts usually come from either faulty equipment or from the 50 Hz interference from electric equipment and cables in the surrounding environment [NeuroSky, 2009]. Another source of noise can occur from any slight detachment or misplacement of the electrodes from the user's head. As a result, the potential difference rises to ininity, or saturates at a maximum, and then may fall back to zero. Such an artifact, however, can be easily detected in the EEG signal and rectified. This type of artifact has been encountered in the work presented in this thesis and will be illustrated and discussed in further chapters.

Due to those two types of articfacts most BCI user experiments were carried out in quiet and strictly monitored lab environments. Moving BCIs outside the lab and to the real world is considered a paradigm shift that would open many possibilities for using BCI systems in the day to day life of both healthy and handicaped users.

Artifact removal in offline and online studies without corrupting the signal is a complex procedure

The removal of artifacts from the EEG signal is the first step in the long process of EEG signal processing. The sim-

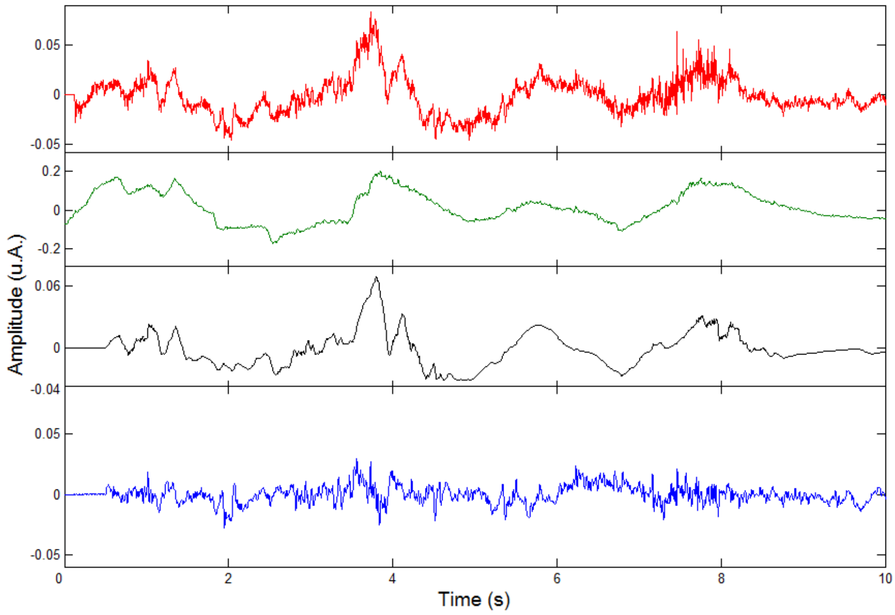


Figure 4.3: Top: EEG signal contaminated with EOG blinking artifacts - Middle & Bottom: Adaptive Filtering to remove EOG artifact [Correa and Leber, 2011]

plest way to remove artifacts would be to discard the whole corrupt trial. This trivial way can only be used in the case of data collection for training during offline studies. However, in case of online single-trial studies, the isolation and removal of artifacts has to be done without corrupting the underleying the EEG signal.

Principal Component Analysis (PCA) and Independent Compoenent Analysis (ICA)

Higher-order statistical separation methods are used for artifact removal

In case of EOG, EMG, and ECG artifacts, several time and frequency domain filtering methods can be used to detect and remove the artifact from the EEG signal. In addition, notch filters can be used to detect and remove the 50 Hz power artifact. The problem of those basic filtering methods is that they often smear of add even additional noise to the signal. Other higher-order statistical separation methods like Principal Component Analysis (PCA) and Independent

Component Analysis (ICA) are more efficient and more popular whith more complex BCI systems involving a large number of electrodes (typically ¿ 32) [Smith, 2004].

PCA and ICA are two popular methods for artifact removal in multi-channel EEG signals

Principal component analysis (PCA) is a Blind Source Separation (BSS) method that decomposes an EEG signal recorded from multiple electrodes into a set of linearly uncorrelated compoenents. As discussed in Smith [2004], the compoenents comprising artifacts such as EOG or ECG can be either manually discarded in the recomposition or correlated with artifacts recorded from simultaneous trials and hence detected and removed.

Independent Component Analysis (ICA) is an extension of PCA. It is developed to deal with what is known as the 'Cocktail Party Problem' defined in the definition box below. As defined in Fatourechi et al. [2007], ICA is a method of BSS which separates a mix independent source signals by forcing the components to be independent. It has proved highly successful in removing occular artifacts (EOG). An advantage of of ICA is that it does not rely on the availability of reference artifacts for separating the noise signal from the EEG. On the other hand, it usually needs prior visual inspection to identify artifact components. Research on automating this process has been also proposed. ICA is also considered to be a computationally expensive task that is not well suited for online studies.

COCKTAIL PARTY PROBLEM:
This problem occurs when a number of people are talking simultaneously in a room and one is trying to follow one of the discussions. This sort of auditory source separation problem can be smoothly handled by the human brain, however, it presents a complex problem in the world of digital signal processing.

Definition: *Cocktail Party Problem*

PCA, ICA and frequency domain filtering are of course not the only artifact isolation and removal techniques available. A lot of research is concerned solely with this complex task, however, this report only aims to give a quick overview about this aspect.

4.3 EEG Feature Extraction

After isolating and removing noise and artifacts and rectifying the EEG signal, the extraction of important and useful features from the EEG signal follows. Feature extraction intends to identify and extract specific characteristics of the EEG signal which encode the commands, messages, or brain activity during performing certain tasks. A variety of different feature extraction methods exist. Some of which extract time domain information, like the evoked potential amplitudes, whereas other information can be extracted by transforming the signal into the frequency domain [Cabrera, 2009]. The goal of this stage is to maximize the discriminative information and therefore to optimally prepare the data for the subsequent classification step.

Frequency domain feature extraction via FFT is popular in EEG research

Section 3.2.1 explained the existence of several frequency domain bands in the EEG signal and illustrated their meanings. The existence of those frequency bands naturally motivates investigating the spectral EEG characteristics through applying frequency domain transformation of the time domain digitized EEG signal. The power of the various frequency bands (delta, alpha, beta, theta and gamma) can be extracted and distinguished by using band pass filtering and Fast Fourier Transform (FFT).

Cepstral coefficients have recently been used in research as EEG features

In some recent studies including Othman et al. [2009] and Temko et al. [2010] , cepstral coefficients, which were originally used for Automatic Speech Recognition (ASR) to extract features from speech, were suggested for EEGs. A cepstrum is calculated by taking the result of a Fourier Transform (FT) of the logarithm of the spectrum of a signal. The power cepstrum particularily is popular is ASR and is defined by equation 4.1.

$$Power\, Cepstrum(x) = \left| \mathcal{F}\left\{ log(|\mathcal{F}\{f(t)\}|^2) \right\} \right|^2 \qquad (4.1)$$

The cepstrum is then converted to the Mel-scale, derived by Stevens et al. [1937], which is a perceptual scale of pitches judged by listeners to be equal in distance from one another.

Mel-frequency cepstral coefficients (MFCCs) are coefficients that collectively make up a Mel frequency cepstrum (MFC). Cepstral coefficients are also used as features in the work done in this thesis. The details of cepstral features extracted will be further discussed in chapter 6.

Among the features extracted from the time domain signal and used in the work in this thesis are the HJorth parameters which were derived by Hjorth [1970]. They have been used in many works including Ansari-Asl et al. [2007] and proved to be successful in EEG classification especially in the field of emotion recognition. There are three main parameters, signal activity, mobility and complexity which are explained by equations 4.2, 4.3 and 4.4. The HJorth activity refers to the variance of the signal whereas the signal mobility calculates the signal's mean frequency where x' stands for the derivative of signal x. Finally, Complexity measures the deviation of the signal from the sine shape.

HJorth parameters proved successful in emotion recognition via EEGs

$$Activity(x) = \frac{\sum_{n-1}^{N} (x(n) - \bar{x})}{N} \tag{4.2}$$

$$Mobility(x) = \sqrt{\frac{var(x')}{var(x)}} \tag{4.3}$$

$$Complexity(x) = \frac{Mobility(x')}{Mobility(x)} \tag{4.4}$$

Although frequency domain features have proven more descriptive of EEG signals, some time domain features are also used in classification. Those include the maximum and minimum signal values, root-mean-square (RMS) of the signal, phase coherance, and the differences or ratios between different electrode channels. The features used in this project will be explained in chapter 6 in details.

More time domain features can be extracted from EEGs such as RMS

4.3.1 Dimensionality Reduction (Feature Selection)

Whereas feature extraction itself is considered a method of dimensionality reduction, in the sense that meaningful features better simplify and describe a dataset, extracting a very large number of features also requires a selection step to produce a subset of the best features describing the dataset.

Wrappers, Filters, and Embedded Feature Selection Algorithms

Three main feature selection methods exist, namely : Filters, Wrappers and Embdedded. The filter feature selection method aims to *filter* the attribute set to produce the most promising subset before learning commences. On the other hand, the wrapper method *wraps* the learning algorithm into the selection procedure, which makes it computationally more expensive Witten et al. [2011]. Finally, the embedded method, where the feature selection and induction processes , i.e. the process of learning the appropriate classifier, are indivisable from each other [Millán et al., 2002]. The two feature selection filter methods that are used in this project are the correlation-based and infomation gain feature selection methods.

Correlation Based Feature Selection (CFS)

CFS is a relatively simplistic and fast feature subset selection method. It was first developed by Hall [1999] who defined the *best* subset of features selected from a feature space to be:

> Good feature subsets contain features highly correlated with the classification, yet uncorrelated to each other

There are two criteria for choosing feature subsets

The algorithm takes into consideration two criteria:

1. How good the individual features are at predicting the class

2. How much the individual features correlate with the other features

Good subsets of features contain features that are highly correlated with the class and uncorrelated with each other. CFS is suitable for handling EEG data because it directly handles correlated and irrelevent features according to Koprinska [2010].

Highly correlated subsets are the chosen ones

Information Gain Feature Selection

Information Gain (IG) feature selection is a very popular and successful feature selection algorithm for highly dimensional data. IG is widely used in the area of Natural Language Processing (NLP) in text classification tasks [Yang and Pedersen, 1997].

Given a set of classes $C = \{c_1, \ldots, c_k\}$, the information gain of a feature f , defined as $IG(f)$, is the expected reduction in entropy H caused by observing f:

$$\begin{aligned} IG(f) &= H(C) - H(C|F), \text{ where} \\ H(C) &= -\sum_{i=1}^{k} P(c_i) \log\, P(c_i), \\ H(C|F) &= -P(f) \sum_{i=1}^{k} P(c_i|f) \log\, P(c_i|f) \end{aligned} \tag{4.5}$$

As explained in Koprinska [2010], the computation of the information gain is done for each feature across all classes and then the features are ranked based on their IG value. Features with a higher IG value are more informative. A subset containing a user-defined number of the highest ranked features is then chosen before moving to the classification step.

Selection of features is based on highest IG values

4.4 EEG Signal Classification

This section explains the theoretical and mathematical background behind some of the most popular classification techniques used in EEG signal classification. In the next chapter, Related Work 5, research done on EEGs using the discussed classification techniques will be overviewed as well.

4.4.1 Naïve Bayes Classifier

Naiive Bayes is based on the rule of conditional probability

Naïve Bayes is a simple supervised learning probabilistic algorithm which assumes that the effect of an attribute value on a given class is independent of the values of the other attributes [Witten et al., 2011]. This naïve assumption produces as simple model which often works surprisingly well.

Naïve Bayes method is based on $P(c|F)$ according to Bayes Rule of conditioinal probability, $P(c|F)$ of an instance f and a class c which is given by:

$$P(c|F) = P(c) \cdot \frac{P(F|c)}{P(F)} \tag{4.6}$$

Theoretical explanation of Naiive Bayes

Where $P(c)$ is the probability of clsas c and $P(F|c)$ is the probability of instance F given class c. Having a set of training instances each consisting of a set of features $F = f_1, f_2, \ldots, f_k$, then the probability $P(F|c)$ can be given by the multiplaction of the probabilities of each individual feature f_k given the class c. Note that the denominator of equation 4.6 can be neglected by normalizing the probabilities of all attributes adding up to 1. Equation 4.6 can now be given as:

$$\begin{aligned} P(F|c) =& P(f_1, f_2, \ldots, f_k|c) \\ =& P(f_1|c) \cdot P(f_2|c) \cdot P(f_3|c) \cdot \ldots \cdot P(f_k|c) \\ =& \Pi_{f \epsilon F} P(f|c) \end{aligned} \tag{4.7}$$

Finally, the probability $P(c)$ of a class c can be computed as the frequency of occurance of c in the training set.

Learning with the naive Bayes classifier is straightforward and involves simply estimating the probabilities in the right side of Equation 4.6 from the training instances. The result is a probabilistic summary for each of the possible classes where the instance is assigned the class with the highest probability [Hall, 1999].

Naïve Bayes assigns the instance to the class with the highest probability

4.4.2 Bayesian Network

To overcome the disadvantages of the Naïve Bayes classifier and produce a classification technique more suited to real life problems in which attributes are actually dependent, Bayesian Networks were derived. Bayesian Networks intend to produce predictions of classes instead of only probabilities.

Bayesian Networks are a theoretically well-founded way of representing probability distributions concisely and comprehensibly in a graphical manner. They are represented as a network of nodes, one for each attribute, connected by directed edges in such a way that there are no cycles, which in graph theory is called a *directed acyclic graph* [Witten et al., 2011]. Figure 4.4 represents a simple Bayesian Network.

Graphical Representation of Bayesian Networks

4.4.3 Decision Trees

As defined in Witten et al. [2011], decision Trees represent a *divide-and-conquer* approach to the problem of learning. Each node of the tree corresponds to a feature and each outgoing edge to a node at the next level is associated with a possible value or range of values of that feature whereas the tree leaf specifies the expected value of the class.

Basic description of the nodes and leaves of a decision tree

In a good decision tree, each node should be represent the most informative feature among the features not yet considered in the path from the root. This can be achieved by calculating the information gain (i.e. Entropy) due to that attribute [Millán et al., 2002]. On main advantage of using decision trees in classification is their simplicity. They do not require pruning of many parameters unlike other

Characteristics of decision trees and their advantages

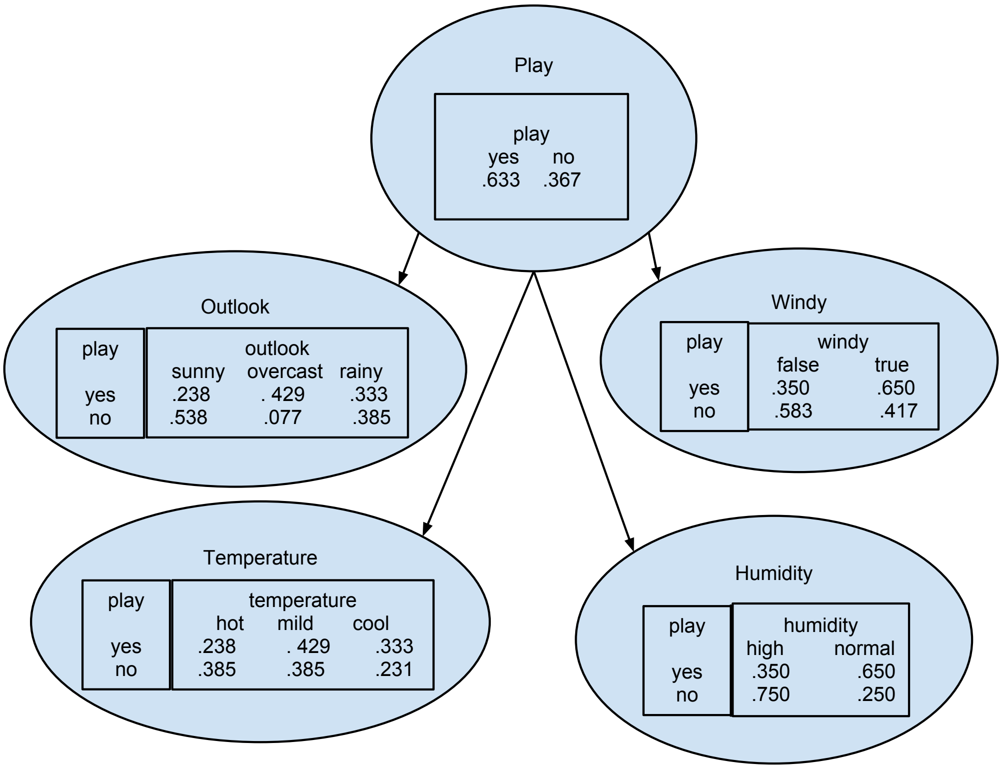


Figure 4.4: A simple Bayesian Network representing weather data [Witten et al., 2011]

classification methods such as Artificial Neural Networks (ANNs). However, the decision tree used in this project, has two parameters which can be set: *confidence value* and a *minimum number of instances* for pruning. Figure 4.5 represents a simple example of decision tree with the leaves representing the two available classes *Play* or *Don't Play* depending on weather conditions represented in the different nodes with values on the edges.

4.4.4 Further Classification Methods

Although the work in this thesis only uses Naïve Bayes, Bayesian Networks and Decision Tress, it is notable to mention that there are other popular classification methods often used in BCI systems. Those include Artificial Neural Net-

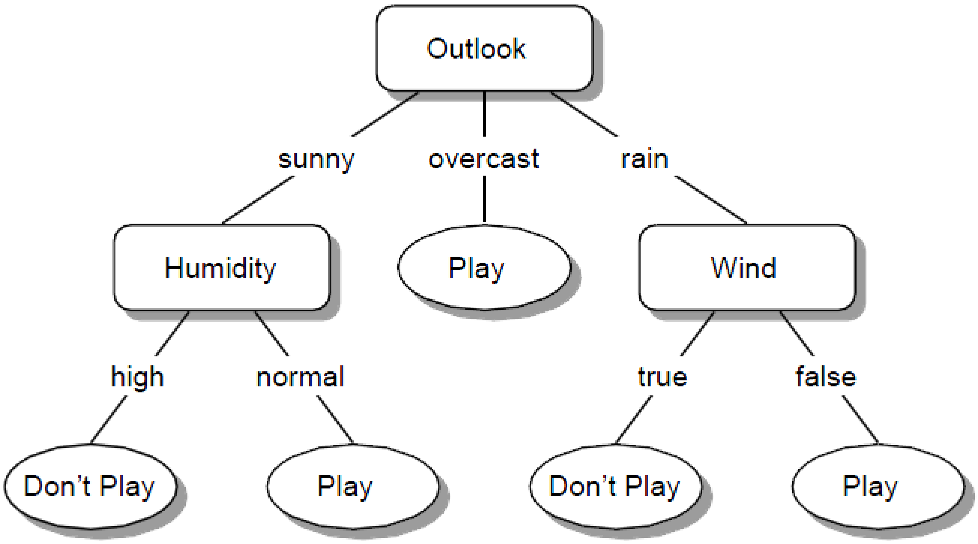


Figure 4.5: A simple decision tree: Leaves represent the 2 classes *Play* or *Don't Play* referring to playing golf if the weather attributes, written on the edges, is good [Witten et al., 2011]

works (ANN) and Support Vector Machines (SVM).

An Artificial Neural Network (ANN) is a computational model that is inspired by biology. The most popular ANN architecture is the feed-forward Neural Network consisting of mainly of three types of layers:input layer, hidden layers and an output layer. This type of ANN is suitable for solving non-linear classification problems by abstractly emulating the structure and operation of the biological nervous system.

ANNs perform well in solving non-linear classifications

The ANN architecture most widely used in EEG classification is the Multilayer Perceptron Neural Network (MLPNN) consisting of input and output layers and any arbitrary number of hidden layers which can be determined by pruning the network according to the complexity of the problem. Good determination of the number of hidden layers is crucial to the performance of the network. Too few hidden layers would lead to poor classification, whereas to many hidden layers would lead to over-fitting and of course, an increased time complexity. The most popular approach of finding the suitable number of hidden layers is simply by

Finding the number of hidden layers in an ANN can be done by trial and error

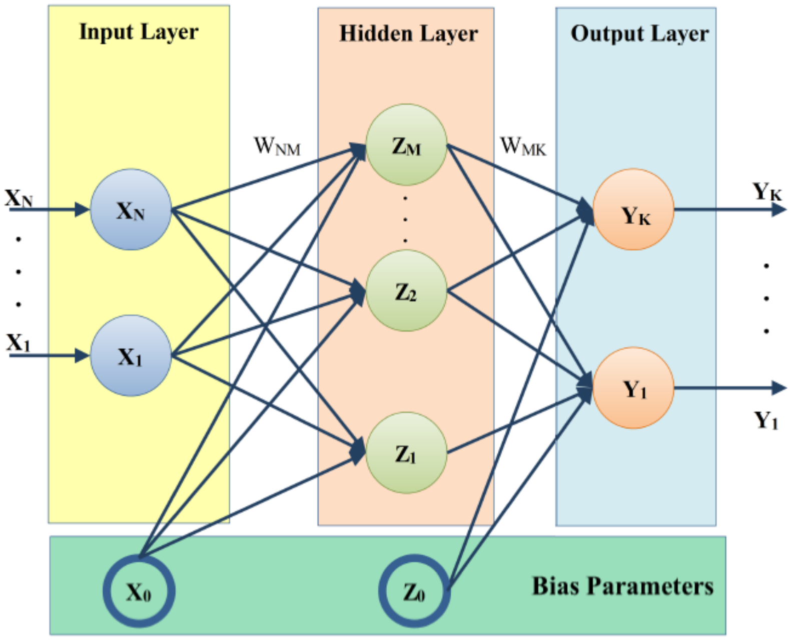


Figure 4.6: An illustration of an Artificial Neural Network (ANN) [Hamzah S., 2011]

trial and error [Subasi and Erçelebi, 2005].

The use of ANNs in EEG classification is very popular and some examples of work using ANNs will be discussed in the following chapter 5. Figure 4.6, shows the structure of feed forward ANN.

SVMs try to find the maximal margin separating different classes

A Support Vector Machine (SVM) is linear modelling algorithm which intends to find a linear separation boundary between different class instances. This boundary must have a maximal margin and this leads realizing better generalizations using SVMs making them a popular classification method. SVMs can also be used for non linear modelling by transforming the space into a hyperplane to deal with more complex classification problems [Kuzovkin].

A sample SVM example is illustrated in 4.7

Figure 4.7 illustrates an SVM classifying two classes. As can be seen in the figure, boundary B1 does not separate the two classes, whereas boundaries B2 and B3 do. However, the euclidean distance between B2 and the nearest instances of both classes is smaller than that of B3. Hence B3 is the chosen boundary. This simple idea of the SVM can be generalized and adapted to suit much more complex and non-linear problems.

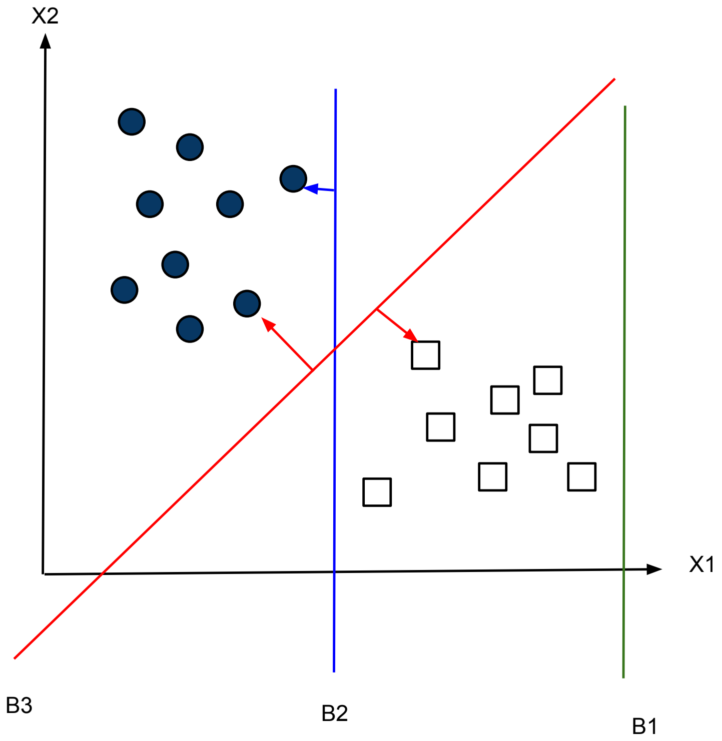


Figure 4.7: An SVM with 2 classes. B1 does not separate the classes. B2 separates the 2 classes but with a smaller margin than B3

4.5 Commcercial BCI Devices

As introduced earlier, the two main BCI devices available in the market are the EPOC device by Emotiv, and the NeuroSky devices (MindSet, MindWave and Brainband). The two brands have spiked the attention of many researchers in various fields due to their relative ease of use, portability and affordability.

The Emotiv EPOC device uses 14 wet electrodes for brain sensing

Whereas in this project the NeuroSky devices were used, it is also important to briefly introduce the Emotiv Epoc as well. The EPOC neuroheadset is made up of 14 saline sensors that are positioned according to the 10-20 positioning system (Figure 4.2) and two reference electrodes. Although the EPOC wet electrodes need some set up time, the device offers high-resolution EEG signals. The Epoc also communicated wirelessly over a Bluetooth connection and includes a gyroscope for better positioning information for cursor and camera control [Emotiv].

A wide range of applications are developed for the EPOC including games, neurofeedback applications and emotion recognition applications. The Epoc can detect various facial expressions, levels of engagement, frustration, mediation, and excitement [Shirazi et al.]. The EPOC comes with a control panel software for connecting the device to the computer as well as training the device to be able to recognize certain brain commands. Figure 4.8 shows an Emotiv EPOC device. Some EPOC applications will be illustrated in chapter 5.

NeuroSky MindSet and BrainBand use one dry electrode

In this project, the two main hardware devices used were: NeuroSky[1] MindSet and MyndPlay[2] BrainBand powered by the NeuroSky chip. Both devices operate in the same manner described in this section with the exception that the MindSet is fitted with a headphone and speaker which connects to a phone or PC via Bluetooth. The BrainBand also communicated with the PC or phone via Bluetooth connection but it does not have a headphone or speaker. Figure 4.9 shows the MindSet (left) and BrainBand (right).

[1] www.neurosky.com
[2] http://myndplay.com/products.php?prod=7

Figure 4.8: Emotiv EPOC Technologies [2012 (accessed October 19, 2012]

NeuroSky devices provide EEG band powers as well as the raw signal via Bluetooth connection

The NeuroSky MindSet and BrainBand are both portable, commercial, research-oriented EEG measuring devices. Both devices are equiped with one dry electrode to be placed on the subject's forehead in the location Fp1 defined in figure 4.2. The MindSet also has three reference electrodes on the left ear, whereas the BrainBand has one large electrode placed at the back of the ear also as a reference. Both devices are fitted with the NeuroSky chip which filters and preprocesses the EEG aquired signal prior to sending it via Bluetooth to the PC or smart phone. The EEG processing protocols are not opensource. However, as stated in the NeuroSky white paper NeuroSky [2009], an FFT is done on the raw signal giving the band powers which are then scaled using a propriraty algorithm to produce output which is only relative to each other. The ouput data is presented in the list below.

1. *Raw EEG signal*: Returns the raw EEG data sampled at 512 Hz

Figure 4.9: Left: NeuroSky MindSet[1], Right: MyndPlay BrainBand[2]

2. *Attention ESense Meter*: An Integer value between 0 and 100 indicating the attention of the subject

3. *Meditation ESense Meter*: An Integer value between 0 and 100 indicating the meditation of the subject

4. *NeuroSky EEG band values*: Delta, Theta, Low-High Alpha, Low-High Beta, Low-Mid Gamma values sampled at 1 Hz

5. *Blink Strength*: Returns an integer value between 0 and 255 indicating the blink strength

6. *Signal Level*: Returns an integer value between 0 [correct signal] and 200 [misplaced device].

Advantage of NeuroSky devices include their portability and embedded filtering

NeuroSky's products make it feasible for researchers and academics to explore the world of BCI due to its portable nature and PC connection. It is embedded with a noise cancellation, signal amplification, filtering protocols which remove the artifacts that were mentioned earlier in this thesis 4.2.1 including muscle, pulse and electrical devices noise artifacts [NeuroSky, a].

The NeuroSky MindSet research tools allow researches and academics develop a wide range of applications addressing many paradigms. NeuroSky is also collaborating with over 300 univeristies around the world in research and development of new hardware and software related to BCI and the

NeuroSky products, as well as conducting novel research in the enhancement of classroom experience and the development of educational aids and research in the medical and healthcare field [NeuroSky, b].

4.6 Weka: Open-source Classification Software

The Weka[3] machine learning software will be used for feature selection and classification in the work presented in this thesis. Weka (Waikato Environment for Knowledge Analysis) is a popular suite of machine learning software written in Java, developed at the University of Waikato, New Zealand. Weka is free software available under the GNU General Public License.

Weka supports many classification and feature selection algorithms and allows parameter pruning

Weka can deal with datasets in a variety of formats, the most popular of which are CSV (Comma Separated Values) and ARFF (Attribute-Relation File Format). Weka includes a collection of machine learning algorithms for data mining tasks implemented in Java. It contains tools for data preprocessing, classification, regression, clustering, association rules, and visualization. From the supplied machine learning algorithms, the Weka J48 implementation of decision trees as well as the Weka implementation of Naïve Bayes and Bayesian Network algorithms, were used in this project. In addition, a set of preprocessing attribute selection filters were applied to the datasets used in this work which will be discussed in detail in chapter 6. Finally, Weka can be used through its Graphical User Interface (GUI), through the command line or through Java code.

Chapter Summary

This chapter has introduced the basic building blocks of BCI systems. The different signal aquisition and processing techniques were discussed. The features of EEG signals were introduced and the various classification algorithms used in EEG classification were explained. In addition, an overview of the BCI hardware devices available in the market and used in this project was given. Finally, the open-source

[3] `www.cs.waikato.ac.nz/ml/weka/`

classification software used in this project was introduced.

Chapter 5

Related Work

In this chapter, previous work in the field of EEG task classification will be discussed. An overview of applications implemented using commercial BCI will also be introduced.

5.1 Task Classification using EEG Signals

The Keirn and Aunon EEG dataset was collected from four subjects during five tasks

Most work mentioned in this section is based on the popular Keirn and Aunon [Keirn and Aunon, 1990] , [Keirn, 1988] EEG dataset. This dataset was collected in strictly controlled lab conditions with minimal noise where four subjects were asked to perform a set of five tasks. The first task was a baseline task in which the subjects were asked to relax as much as possible. Secondly, the letter task, in which the subjects were asked to mentally compose a letter to a friend. The third task was a math task in which subjects were asked to solve non-trivial multipliction problems. The fourth task was a visual counting task in which subjects were asked to imagine numbers being written on a blackboard sequentially. Finally, geometrical figure rotation, in which subjects were asked to visualize a particular 3D figure being rotated in space. Six electrodes fitted in an electrode cap in positions C3, C4, P3, P4 , O1 and O2 according to the 10-20 electrode placement system (4.1.1) were used for signal acquisition.

ANN classification using the Keirn and Aunon dataset

The Keirn and Aunon dataset was used by many researchers

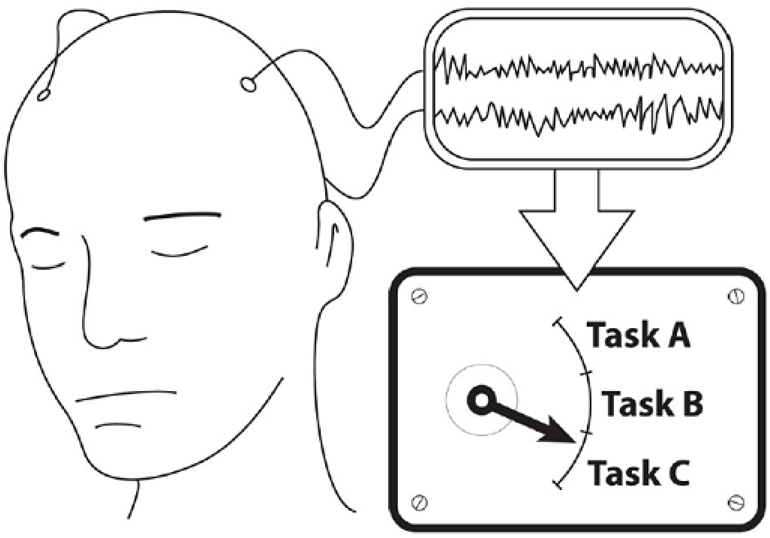


Figure 5.1: A conceptual illustration of a BCI during task classification of EEG signals Lee and Tan [2006]

to apply different feature extraction and classification techniques in an attempt to classify those five mental tasks, or subsets of them. In [Anderson and Sijercic, 1996], data from four subjects from the Keurn and Aunon dataset was used to train a feed-forward Artificial Neural Network (ANN) (4.4.4). The ANN was trained using 10-fold cross-validation. The EEG data was represented using Auto Regressive (AR) models. The output of this work showed that the range of signal classified correctly was between 71% to 38% depending on the subject.

SVM and RBF classification using the Keirn and Aunon dataset

In the research of [Hosni et al., 2007] three of the five tasks from Keirn and Aunon's dataset were used. Namely: the baseline task, the letter task and the math task. Eye blinks were detected and removed using Independent Component Analysis (ICA) that was previously introduced in section 4.2.1. Three different feature extraction techniques were compared in this research paper: Parametric Auto Regressive (AR) modelling, AR spectral analysis and band power differences. Classification was done using Radial Basis Function (RBF) and Support Vector Machines (SVM) and the best classification accuracy recorded was 70%.

In [Palaniappan, 2005], permutations from the baseline task and one other task from the Keirn and Aunon dataset was used. The objective of the study was to classify one more task beside the baseline task with very high accuracy which can be later used by paralyzed patients in controlling a cursor on a screen. For each of the four subjects, the best mental task pair was found. Power band features were extracted and classification was done using an ANN. Classification accuracy reached up to 97.5% given that the most reliable task per subject was used.

Classification of two mental tasks using the Keirn and Aunon dataset

In [del R Millan et al., 2002], a novel method is proposed for applying a simple neural network classifier for online classificaion of five mental tasks. The proposed neural classifier was found to recognize three mental tasks from online spontaneous EEG signals. Classification accuracy was found to be around 70%. However, the rapid responses (every half second) compensate for this modest accuracy. The data was collected from eight subjects using a portable 26-electrode BCI system in which only eight electrodes were used for recording. The mental tasks chosen for this study were were *relax*, imagination of *left* and *right* hand movements, *cube rotation*, and *subtraction*.

Classification of five mental tasks using 8 electrode channels

In [Lee and Tan, 2006] two experiments were carried out using low-cost off-the-shelf EEG equiptment, namely a Brainmaster[4] AT W2.5, which is a PC-based 2-channel EEG system. In the first experiment, a mean classification accuracy of 84.0% was achieved in subjects performing one of three cognitive tasks: rest, mental arithmetic, and mental rotation, while sitting in a controlled environment and without moving. In the second study, conducted in more ecologically valid setting for HCI research, a mean classification accuracy of 92.4% was achieved using three tasks that included noncognitive features: relaxation , playing a PC-based game without opponents, and playing the same game with opponents. Power spectral features were extracted from the signal after preprocessing and noise removal, and classification was done using a Bayesian Network classifier which was previously discussed in subsection 4.4.2.

Using low-cost EEG for classification of three mental tasks

Many more studies on mental task classification using EEGs

[4] http://www.brainmaster.com.

were done. However, we limit this section to the work aforementioned which is closer to the work presented in this thesis.

5.2 Commercial BCI Applications

In this section, applications using commercially available BCIs are reviewed. Most applications mentioned in this section are implemented using NeuroSky or Emotiv devices (discussed in section 4.5) and are targetting mostly healthy users.

5.2.1 Desktop-based Applications

NeuroWander and NeuroBoy games

A BCI game based on the German fairytale of Hansel and Gretel was implemented in [Yoh et al., 2010] using the NeuroSky Mindset. The game is controlled by the user's attention and meditation signals, introduced in section 4.5. Depending on the attention and meditation level, the user can move to the next level of the game. Similarly, NeuroBoy, a game developed by NeuroSky and provided with the Mindset gadget, operates in the same way. In this game, the user takes the role of a boy with psychic powers who moves, floats, and burns object depending on the attention and meditation levels captured from the Mindset, or other NeuroSky gadgets, during play. However, NeuroBoy is quite simplistic and does not provide the user with a meaningful gameplay.

Reading Tutor using NeuroSky Mindset

In [Mostow et al., 2011], a study was implemented using the NeuroSky Mindset in an attempt to know if low-cost commercial BCI can be used in determining when children had reading difficulties. The study was conducted on 10 adult users and 11 school children in which they were presented texts labelled as easy and difficult to be read both, aloud and silently. The Mindset was used in conjunction with Project LISTENs Reading Tutor [Mostow and Beck, 2007]. The Reading Tutor displays text, listens to the users read aloud, and logs detailed longitudinal records the reading sessions. This study has proved that EEG data from the Mindset can discriminate between reading easy and hard sentences reliably with good accuracy, accross different modalities (oral and silent reading), as well as across populations (adults and children). Frequency bands more sensitive to difficulty and to various lexical properties were also identified, which suggests that they can detect changes in cognitive task demands.

In [Shirazi et al.], the Emotiv EPOC was used in annotating video segments to ease in the searching and playback process.This was done by creating an annotation tool called *MediaBrain*. The tool was used in a user study of 11 participants to annotate segments of videos based on the excitement information collected by the headset and by the user's pressing a button when they felt excited. The research concludes that it is feasible to implicitly annotate a video based on excitement information, generate a set of highlights and a video summary.

5.2.2 Mobile-based Applications

NeuroPhone [Campbell et al., 2010] and ThinkContacts [Perkusich et al., 2011] are both mobile phone BCI applications designed to help users with motor disabilities as well as healthy users to dial phone numbers without navigating to their contact list, and only via brain signals.

NeuroPhone Emotiv Application

NeuroPhone [Campbell et al., 2010] is implemented using the Emotiv headset containing 14 electrodes and is based on the P300 EEG signal previously introduced in section 3.2.2. An IPhone application is implemented, which involves a grid containing flashing photos of people from the contact list. When the desired contact flashes, a P300 signal is elicited in the same manner the popular P300 speller operates. EEG signals from the headset are then communicated to the phone and a lightweight classifier recognizes the P300 signal and removes signal noise. The correct phone number of the chosen contact is then automatically dialed. Figure 5.2 illustrates the NeuroPhone application in use.

ThinkContacts NeuroSky Application

ThinkContacts [Perkusich et al., 2011] is implemented using the Minset of NeuroSky, which, relying only on a single

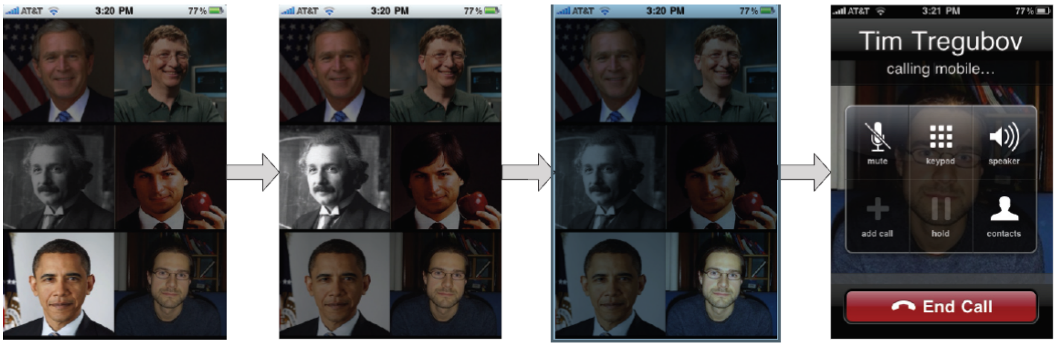


Figure 5.2: NeuroPhone: Snapshots from the iPhone application - The contact 'Tim' is only dialed after flashing Tim's photo as seen in the third snapshot Campbell et al. [2010]

dry electrode, is much simpler than the NeuroPhone application of Emotiv but intends to serve the same purpose. ThinkContacts shows the name and avatar of the contacts in the address book in sequence. Two progress bars at the bottom of the screen are available: one for the current meditation level (left) and one for the current attention level (right), as shown in Figure 5.3. The selection of the desired contact is determined by controlling the attention and meditation levels. If the attention level of user is higher than 70%, the next contact is shown, and if it is lower than 30% the previous contact is shown. To select a contact to be dialed, the meditation level has to be higher than 80%.

An Emotiv mobile application that reconstructs brain images on a Nokia smartphone

The study in [Petersen et al., 2011] attempts to reconstruct a 3D brain image showing the active brain lobes on a Nokia N900 smartphone. The EEG signal from the 14-electrode Emotiv EPOC device is sent via Bluetooth to the phone. A Bayesian approach is applied to reconstruct the neural sources which attempt to distinguish between emotional responses from seeing pleasant, unpleasant and neutral pictures. The neural responses are rendering in a 3D brain model on the phone. The objective of this application is to not only help in the differentiation of emotional responses, but also to provide an intuitive interface for touch-based interaction.

From this brief overview, it is obvious that not so many mobile BCI applications exist. This is due to the novelty of

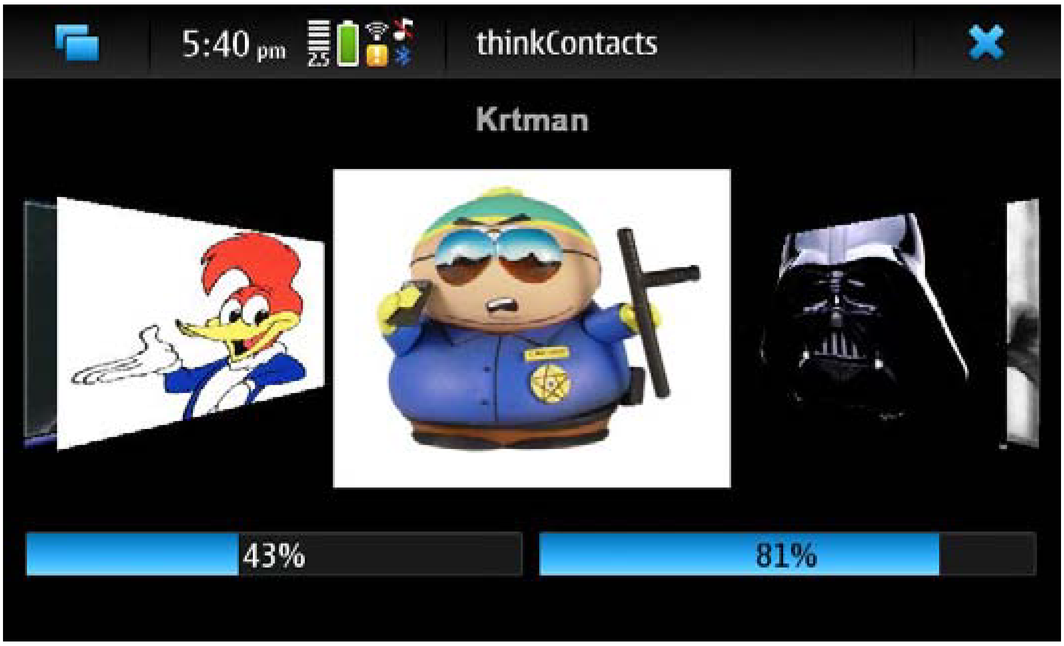


Figure 5.3: A snapshot from the ThinkContacts application [Perkusich et al., 2011]

this modality of BCI applications.

Chapter Summary

In this chapter, an overview of related BCI work was reviewed. Some examples of BCI applications concerned with the classification of mental and cognitive tasks were given in section 5.1. Commercial BCI applications, both desktop and mobile, were reviewed in section 5.2.

Part III

Mental Task Classification

Chapter 6

Data Acquisition and Feature Extraction

This chapter provides a detailed description of the architecture of the implemented system. The chapter begins with an overview of the system architecture, then moves on to discuss each module separately. The design of each module and the challenges encountered will also be discussed.

6.1 System Architecture

According to the basic structure of BCI systems overviewed previously in chapter 4, the system architecture of the work presented in this thesis is divided into four main modules which are illustrated by figure 6.1. The first module, data aquisition, is concerned with aquiring the EEG dataset using the NeuroSky Mindset and Brainband. The second module is the pre-processing, segmentation and noise removal module. The third module is concerned with feature extraction and attribute selection for dimensionality reduction. Finally, the fourth module which is the core of the system, is the classification module. This will be discussed in the the following chapter and will include the various classification algorithms implemented, other signal analysis techniques, and finally the results and evaluation of the system.

The architecture of this system is based on the typical BCI system structure

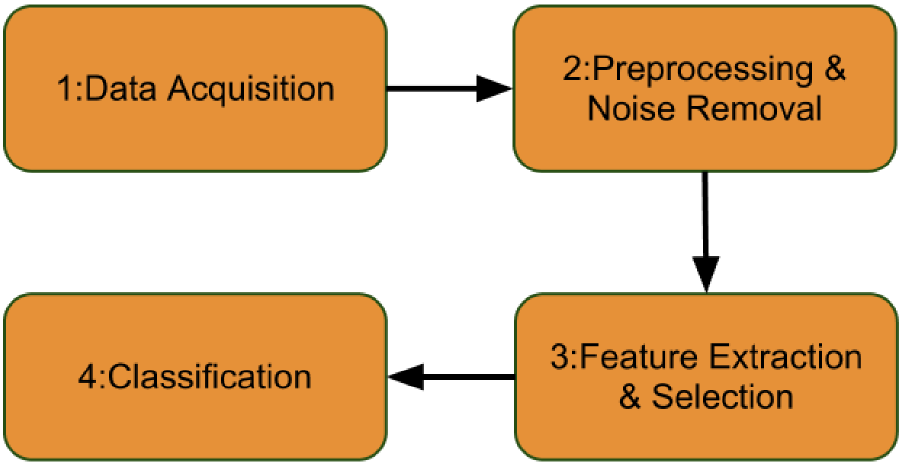


Figure 6.1: System architecture consisting of 4 main modules

6.2 Data Acquisition

The lack of similar work using NeuroSky devices motivated the need for creating a new data corpus of brain signals

In order to investigate the aforementioned research questions 2.4, a large amount of EEG data had to be collected. No previous similar datasets exist, that are using the same type of off-the-shelf commercial BCI targetting mental task classification. Therefore, desiging a user study and creating a new data corpus is considered one of the contributions of this project.

The user study used real day-to-day task recordings

In order to collect a large amount of EEG data, a user study was designed having the objective of being as close to real life as possible. The choice of mental and cognitive tasks to be classified was based on real life tasks that are experienced by normal healthy users in their day-to-day activities. The study consisted of five tasks, each for a duration of five minutes. The tasks are explained in the list below.

1. *Relaxation (Baseline Task)* : Users were instructed to relax as much as possible, for five minutes, in the means they saw best to achieve this purpose.

2. *Reading task*: Users were given a short Arabic drama story of three pages to read.

3. *Listening to audio*: Users were asked to listen to music from a popular online Egyptian radio channel. The music involved pop Arabic songs, advertisements, and some short segments of speech.

4. *Watching a movie*: Users were asked to watch a short Arabic documentary about the brain and how it functions.

5. *Playing Sudoku* : Users were asked to play a game of medium level Sudoku from an online source.

Results of pilot study favoured choosing the Brainband over MindSet since it was more stable

In a pilot study involving three users of different head sizes, usage of both the Brainband and Mindset devices was tested by playing the NeuroBoy game provided by NeuroSky with the Mindset device. By inspecting the $PoorSignal$ value provided from NeuroSky to test the quality of the signal, it was found that the MindSet device is often not stable, especially on users with a smaller skull and hence takes a longer time for the signal to stabilize to get a $PoorSignal$ value of zero. The electrode resting on the forehead often lost contact with the subject's skin in case of subjects who were shorter and subsequently had a smaller sized skull. When the same experiment was repeated with the Brainband, it proved to be much more stable since the band was tied around the head, and hence, well-fitted and did not move or lose contact with the skin. For the aforementioned reasons, the Brainband device was chosen for conducting the user study which is described in detail in the following sections.

Study conditions and duration

The study was conducted in a quiet laboratory with normal lighting conditions and minimal noise from other electric equipment available around the lab. 20 subjects participated in the study after signing a written consent. All subjects were briefed about the purpose of the study and given background information about EEG, how it works, and were aquainted with the Neurosky Brainband device. All 20 subjects were Egyptian and Arabic was their mother-tongue which was particularily important to ensure that no extra mental effort was exerted to understand text or audio in the reading, movie, and audio tests. Additionally, to balance out test sequencing effects, all users started with the relaxation task then the order of the tasks was permuted each time. A summary of user information is given in subsection 6.2.1.

6.2.1 Study Implementation

An Android application was developed for data acquisition

Data aquisition was done through developing a mobile application on a *Samsung Nexus S* smartphone running the Android operating system. The mobile application was designed for two main reasons: To experiment with the Neurosky Andoird API, and evaluate its ease and capabilities, as well as to allow for easy collection of data using a mobile phone, and hence, giving the experimenter more freedom of movement and experience one of the main advantages of all commercial BCI EEG equipment: portability.

Bluetooth was used for communication between the phone and the headset

Communication between the Brainband and the phone was done via a wireless Bluetooth connection and according to the ThinkGear proprietary communication protocol developed by NeuroSky. The application was built using Java (Android SDK) on the Eclipse IDE and using the Android API provided freely on the Neurosky website. The application was installed on the smartphone and it included the following functionalities:

- Entry of subject information (name and age)
- Establishment of a Bluetooth connection between the phone and the Brainband
- Signal the start and end of each test with a 5-minute test timer
- Display and record the raw, attention, meditation, blink and band power data

Android application allowed for signal stabilization before starting the test

Figure 6.2 shows two snapshots from the Android application. The left snapshot shows the start screen with text entry. The start button is not activated except after the stabilization of the signal on a $PoorSignal < 25$ for at least 10 seconds. The snapshot on the right is taking during the second test,

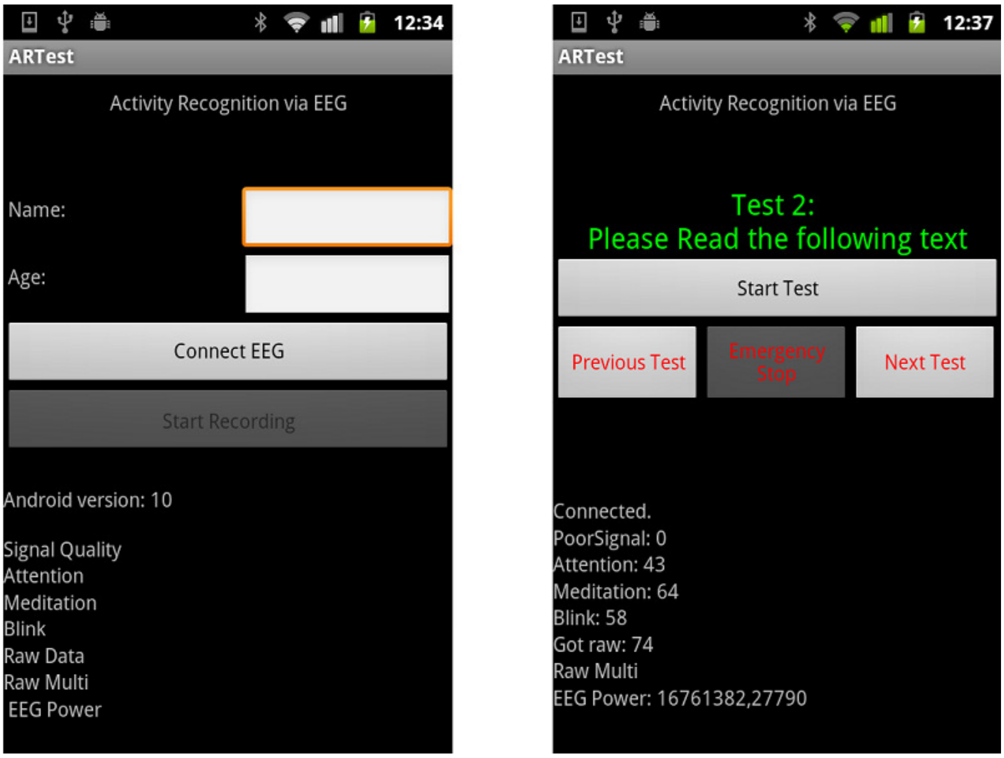


Figure 6.2: Snapshots from the Android Application - Right: Start Screen Left: Reading Test

$reading$. As the snapshot shows, for each test the NeuroSky signal values are shown in real time to the experimenter. The ability to pause, stop, or repeat the test was also implemented in case of erroneous data or misplaced electrodes which would lead to a high $PoorSignal$ level.

Demographics of study participants

Subjects were first asked to fill down an online survey prior to starting the test. The goal of the survey was to collect basic information about the age, gender, handedness, and stress levels of the subjects prior to the test. The following is a summary of the survey data from 20 subjects:

- *Gender:* 12 males and 8 females
- *Average Age:* 23.3 years
- *Handedness:* All subjects were right-handed
- *Previous Usage of a BCI:* 3 subjects used the Neurosky

BrainBand previously

- *Stress Level:* On a Likert scale , 45% of subjects were neutral or not stressed, 55% were stressed during the day the test was conducted

Most subjects closed their eyes and imagined scenery to achieve relaxation

Since the duration of each test was five minutes, each subject took in total between 30 to 45 minutes to complete all 5 tests, the survey and the briefing. After the test, the subjects were asked to continue the survey to gather information about the comfort of the BrainBand as well as the different relaxation techniques the subjects used for the first task. Most of the subjects mentioned that they sat back deeper in the chair, closed their eyes, tried to imagine scenes of the beach and sand, or tried to remember nice memories of the past while listening to the relaxation music played.

Figure 6.3 illustrates a summary of the statistical information gathered from the survey including the gender, previous BCI usage, the stress level that the subjects were feeling on the day of the test, and finally, the perceived comfort level of the brainband (with 5 as least comfortable).

Data collected was saved in timestamped CSV files

The data collected was saved on the SD card of the smartphone in Comma Separated Values (CSV) format. For each subject, five tasks were performed, and for each task, five different CSV files were saved, namely: raw, blink, power, attention and meditation data. Below is a description of the contents of each of these files:

1. *eSense Values: Attention and Meditation:*Values ranging from 1 to 100, at a sampling rate of 1 Hz. Those values are determined via Neurosky proprietary algorithms where values between 40 and 60 are considered 'neutral' or baseline. Values between 60 and 80 mean slightly elevated eSense levels, values between 80 to 100 refer to strongly elevated attention/meditation levels. Similarily, values between 20 and 40 mean slightly lowered levels, and between 1 and 20 indicate strongly lowered levels. Finally, a zero eSense value means the signal cannot be calculated reliably due to background noise.

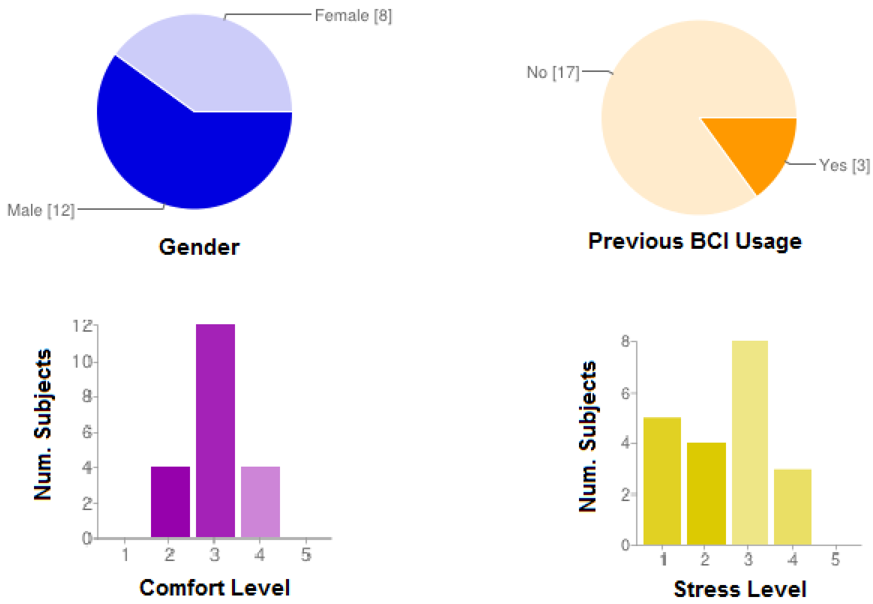


Figure 6.3: Summary of Statistical Survey Data - Top Left: Gender Information - Top Right: Previous BCI Usage - Bottom Left: BrainBand comfort level (5 as least comfortble) - Bottom Right: Stress Level on Study Day (5 as most stressed)

2. *Neurosky Power Values:*A series of eight 3-byte long values ranging from 0 to 2^{24} provided at 1 Hz. These values are: delta (0.5 - 2.75Hz), theta (3.5 - 6.75Hz), low-alpha (7.5 - 9.25Hz), high-alpha (10 - 11.75Hz), low-beta (13 - 16.75Hz), high-beta (18 - 29.75Hz), low-gamma (31 - 39.75Hz), and mid-gamma (41 - 49.75Hz). These values have no units, therefore are only meaningful compared to each other and to themselves, to consider relative quantity and temporal fluctuations.

3. *Blink:*Blink strength value is a 1 Byte value ranging between 1 and 255 provided whenever a blink is detected. The value has no unit and only indicated the relative strength of the blink.

4. *Raw Wave:* A 16-bit value provided at 512 Hz sampling rate. Values for the Mindset communications protocol lie in the interval between -2048 to 2047. Typically in EEGs, time-frequency transforms are used to change the raw signal to the frequency domain, to extract the different EEG power values mentioned in section 3.2.1

Timestamp and Poor Signal levels were recorded in the collected dataset

All files were also timestamped with the date and time of signal collection, and the value of $PoorSignal$ at each second was also noted. As explained previously in section 4.5, the $PoorSignal$ value lies between 0 and 200 and is provided at 1 Hz. Any poor signal value above zero is an indication that there is noise altering the signal. Before the start of each test, 10 seconds were given for the signal to be stable and with a zero or low (<25) poor signal value. The collected data was then downloaded on a PC, and sorted into directories to prepare for the data preprocessing module discussed in section 6.3

Figure 6.4: Subject seated wearing the Brainband device during the study

6.3 Pre-processing and Noise Removal

After the data was collected and organized into folders, the data needed to be manually manipulated to prepare it for the forthcoming feature extraction and classication modules.

Preliminary visual data inspection let to discovery of unusual data fluctuations

By visual inspection of the data, it was noted that some ubnormal data fluctuations were visible in five subjects, where the expected power data values were very high during most of the duration of the tests. This inspection has led to the conclusion that thorough inspection and artifact removal need to be done on the data prior to further processing and this is discussed in the following section.

6.3.1 Identification and Removal of Corrupt Data

All data inspection and analysis in this module were done using Matlab[5] R2011b licensed software. Matlab scripts were written to extract subject data from the CSV files into Matlab structs to be able to manipulate the data easily.

In the beginning, a set of basic scripts were executed to be able to visualize subject data from different tests. It was noticed that data from subjects 1 through 5 had unnatural peaks and ranges in the NeuroSky power band values. This can be seen in figure 6.5 which represents the delta band NeuroSky power value from Subject 2 during the reading test. As shown in figure 6.5, delta values greater than 2^{24} were clipped. After thorough analysis of the dataset, it was concluded that the first 5 subjects who performed the user study did not have the electrode on the Brainband properly adjusted on the $Fp1$ position (explained in figure 4.2), hence, data from those five subjects was considered corrupt and wes removed from the dataset.

Misadjusted scalp positions or excessive motor movements led to infinite impedence and data corruption.

6.3.2 Noise removal by signal averaging

By applying the same set of basic visualisation Matlab scripts, it was found that power values in the rest of the subjects in some seconds also were clipped due to reaching the maximum value. However, since this noise was minimal, instead of discarding the whole dataset, the noise was rectified by implementing a noise removal script. The script took the average of surrounding signals to compensate for the corrupted clipped signal at particular points. Figure 6.6 shows data from subject 6 before and after the clipping artifact removal by averaging the previous and next values.

[5] http://www.mathworks.com/products/matlab/

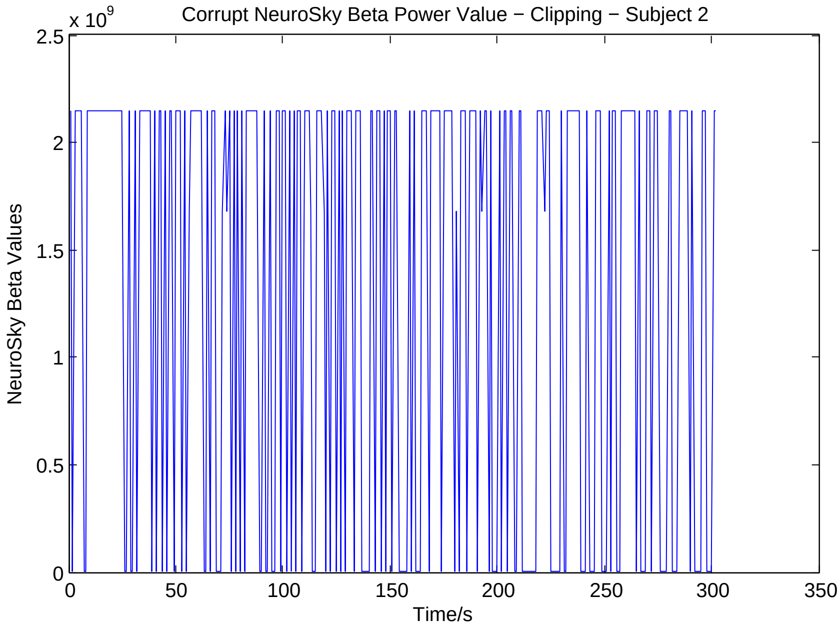


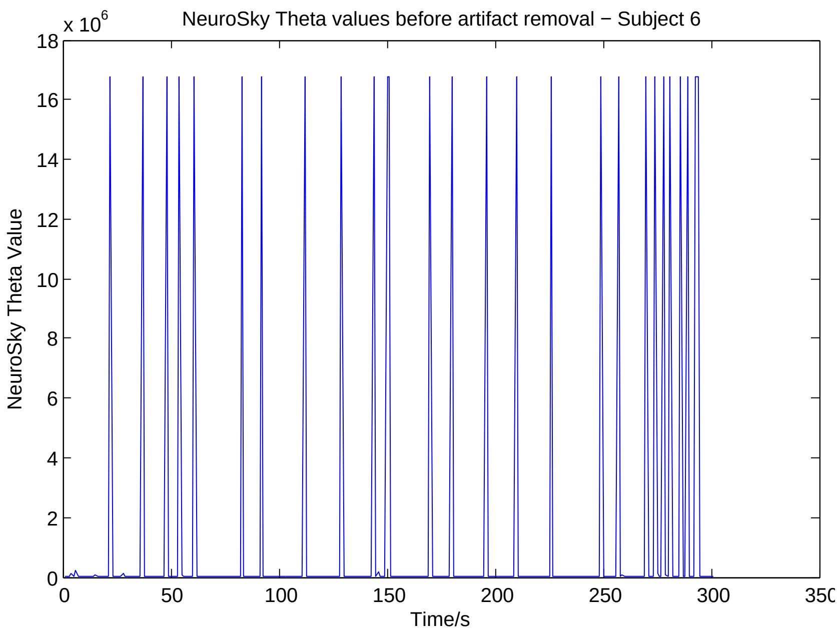
Figure 6.5: Corrupt NeuroSky Delta Power Value for Subject 2 - Clipping at maximum delta value

6.3.3 Data Segmentation

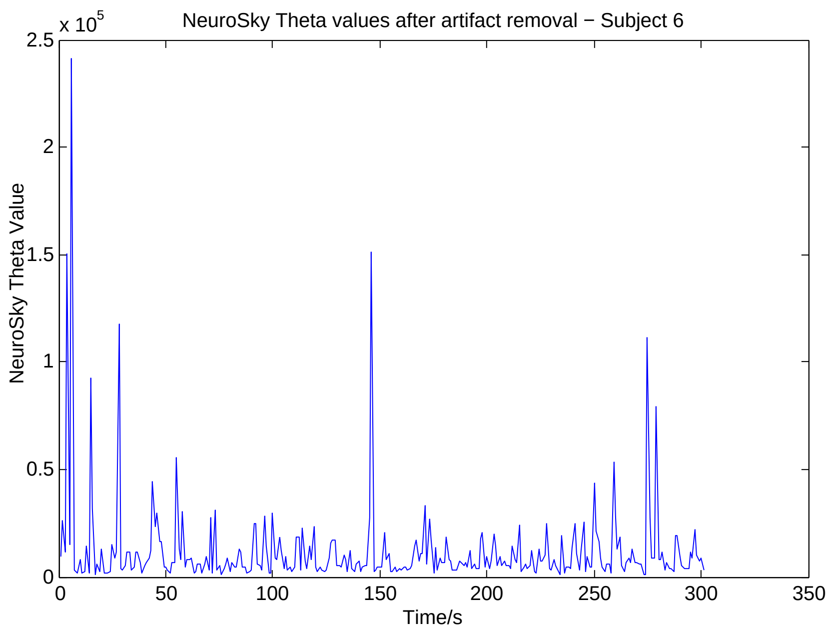
Epoching of EEG data helps to manually extract segments with interesting features

After the process of noise and artifact removal, further preprocessing of the data to prepare the dataset for feature extraction, each 5-minute test was divided into segments and the features were calculated per segment. Segmentation increases the robustness of the system and allows for the development of a more efficient and quick system. Segmentation of EEG data is also known as *epoching* which basically means that the continuous EEG data is chopped into small time periods (e.g. 1000ms). Epoching is used in case of ERP signals to extract segments of the EEG that are known to contain the stimuli.

In this system, epoching is used to prepare the data for the feature extraction step. Each 5-minute continuous signal recording was divided into segments of 10, 5 and 1 seconds. For each of the segmentation options, features were extracted per segment and classification results were then calculated accordingly.



(a) NeuroSky Theta before Artifact Removal - Subject 6



(b) NeuroSky Theta after Artifact Removal - Subject 6

Figure 6.6: Before and after artifact removal

6.4 Post-processing and Feature Extraction

Module Description

The third module in the system architecture described in section 6.1 represents the post-processing and feature extraction step. The set of features extracted from the collected data corpus as well as the feature selection techniques used for dimensionality reduction are described.

6.4.1 The Feature Set

Feature extraction is considered a major and most sensitive step prior to classification. The choice of correct features representing the data is a very critical step in the system. A thorough review of related work has been done to choose the most suitable features for mental task classification.

The development of the final feature set was based on mental task classification work

A preliminary set of features was first chosen to contain all features in table Summary of the features extracted from the collected signals except features number 6 and 29. Those two features were later suggested based on the works of Temko et al. [2010]. In Chapter 7, classification results using the preliminary and final feature sets will be presented.

The final feature set chosen in this work can be divided into: spectral features, NeuroSky features, time-domain features, and finally, cepstral features. Table 6.1 presents a final list of all features chosen to represent the EEG data corpus with a brief description of each feature.

Spectral features

Applying a fast fourier transform (FFT) on the raw signal and bandpassing the delta, theta, alpha, beta and gamma frequency bands (frequency bands previously explained in section 3.2.1), the average of each band per segment is considered as a spectral feature. The ratios between pairs of frequency bands are also calculated. The spectral features are given in features number 7 to 23.

Time-domain features

Features 3 to 6, and 30 to 32, are extracted from the raw time-domain EEG signal. These include the maximum positive, minimum negative and average amplitude of the raw signal per segment, as well as the Root Mean Squared (RMS) value

Num.	Feature	Description
1	Attention Mean	Avg.NeuroSky Attention/segment
2	Meditation Mean	Avg.NeuroSky Meditation/segment
3	Avg. Raw	Avg. of signal raw value/segment
4	Min. Negative Amplitude	Min. raw amplitude/segment
5	Max. Positive Amplitude	Max. raw amplitude/segment
6	Raw RMS	RMS of raw signal/segment
7	Mean Delta	Mean Delta signal/segment (FFT)
8	Mean Theta	Mean Theta signal/segment (FFT)
9	Mean Alpha	Mean Alpha signal/segment (FFT)
10	Mean Beta	Mean Beta signal/segment (FFT)
11	Mean Gamma	Mean Gamma signal/segment(FFT)
12	Delta to Theta	Mean Delta to Theta ratio/segment
13	Delta to Alpha	Mean Delta to Theta ratio/segment
14	Delta to Beta	Mean Delta to Theta ratio/segment
15	Delta to Gamma	Mean Delta to Theta ratio/segment
16	Theta to Alpha	Mean Delta to Theta ratio/segment
17	Theta to Beta	Mean Delta to Theta ratio/segment
18	Theta to Gamma	Mean Delta to Theta ratio/segment
19	Alpha to Beta	Mean Delta to Theta ratio/segment
20	Alpha to Gamma	Mean Delta to Theta ratio/segment
21	Beta to Gamma	Mean Delta to Theta ratio/segment
22	FFT Mean	Mean FFT signal/segment
23	FFT Variance	Variance of the FFT signal/segment
24	NS Delta Mean	Mean NeuroSky Delta/segment
25	NS Theta Mean	Mean NeuroSky Theta/segment
26	NS Alpha Mean	Mean NeuroSky Alpha/segment
27	NS Beta Mean	Mean NeuroSky Beta/segment
28	NS Gamma Mean	Mean NeuroSky Gamma/segment
29	Cepstral Coeff.	Cepstral coefficient/segment 4.3
30	HJorth Activity	Variance of raw signal eq. 4.2
31	Hjorth Mobility	Mobility of raw signal eq. 4.3
32	Hjorth Complexity	Complexity of raw signal eq. 4.4

Table 6.1: Summary of the features extracted from the collected signals

of the raw signal. The RMS, also known as the quadratic mean, is a statistical measure of the magnitude of a varying quantity. It is especially useful when variates are positive and negative. For a raw discrete signal $X = x_1, x_2, \ldots, x_n$, the raw RMS is calculated as:

$$Raw_{rms}(X) = \sqrt{x_1^2 + x_2^2 + \ldots + x_n^2} \tag{6.1}$$

NeuroSky features

Features 1,2 and 24 to 28, were extracted from NeuroSky signals. These are the average attention and meditation values per segment and the average NeuroSky power band values per segment, namely Delta, Theta, Alpha, Beta and Gamma. Altothough NeuroSky power values can be considered from the spectral features, they are not represented in normal power units (dB). Instead, they are represented via the NeuroSky patented algorithms in values that are only comparable to one another.

The calculation of features per segment was done by Matlab scripts and the output was saved per subject in ARFF files which will be explained and discussed in the following section.

6.4.2 ARFF Files

Description of ARFF files

After the noise removal and feature extraction steps were accomplished, the dataset had to be saved in the Attribute-Relation File Format (ARFF) file format that is supported by the classification tool Weka. ARFF files are ASCII files with a list of instances sharing common attributes. The file is divided into two main parts: header and data.

Each row of the ARFF file consists of an instance with attributes separated by commas

The header contains the name of the relation as well as the names of the attributes and their data types. An attribute can be numeric, string, nominal or date. The Data section of the ARFF contains the instances in the dataset in rows. Attribute values are separated by a comma and by convention, the last attribute is usually the labelled class or category that the instance falls in. By the creation of the ARFF files, the dataset is then ready to proceed to the last and final module of the system, the classification module which is discussed in the following chapter.

Chapter Summary

This chapter explained the system architecture and described three of the system's four modules. First, the data aquisition steps were illustrated, followed by the preprocessing and noise removal module. Finally, the features set chosen to represent the data corpus was discussed as well as post-processing steps.

Chapter 7

Data Classification and Evaluation

This chapter presents the fourth and final module of the system architechture illustrated in 6.1, data classification and evaluation. The chapter first starts with a section about the visual and statistical analysis of the data. The second section describes the details of building the different classifiers for different task combinations, classification algorithms and segmentation options. The chapter also explains the problem of sensor drifting and time dependency and how this affected the classification results. Finally, the results of all classifiers is presented, analysed and evaluated.

7.1 Visual and Statistical Data Analysis

Intra-subject data analysis highlighted the stark differences between subjects

In order to get a complete overview of the collected EEG data, ensure its consistency and get a feel about its statistical properties, a set of Matlab scripts were created to inspect and analyse the dataset prior to classification.

A collection of scripts were written to produce intra-subject comparisions within each of the subjects. For each subject the average of the eSense meter values (Attention and Meditation) for each of the five tests was computed. The Neu-

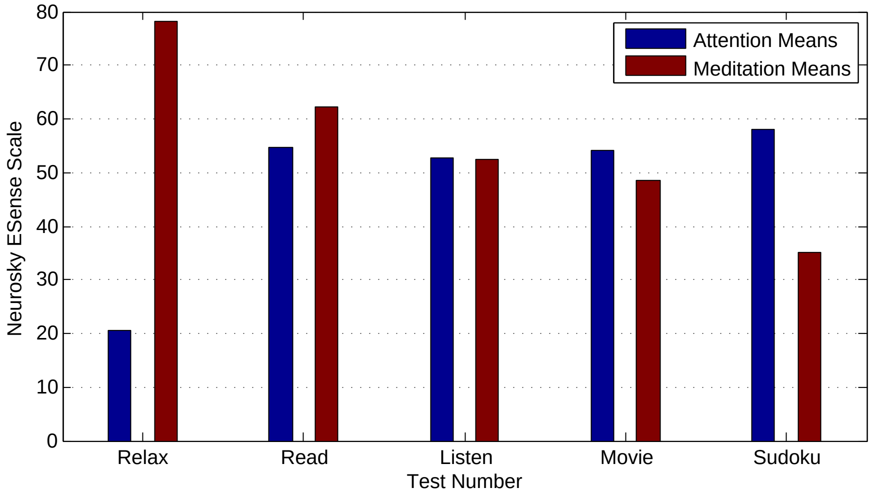


Figure 7.1: Comparison of Attention and Meditation means during the 5 tests - Subject 6 (Male, 23 years)

roSky power values for each of delta, theta, alpha, beta and gamma bands were plotted and visually inspected. As an example, the intra-subject comparisons for subject 6 data is discussed below.

The differences between attention and meditation values were plotted for comparison

Figure 7.1 shows the NeuroSky eSense meter (Attention and Meditation) averages during each of the five tests. As the figure shows, for tests 2, 3 and 4, namely the $Reading$, $Listening$ and $Movie$ tests, the mean of both the attention and meditation was in the neutral range, between 40 and 60 units. For the relaxation test, as expected the meditation mean was close to 80 which is a highly elevated mediation level. For the sudoku test, the mean attention of subject 6 is almost 60, which is in the slightly elevated level indicating that the user was concentrating on solving the sudoku puzzle during the test time.

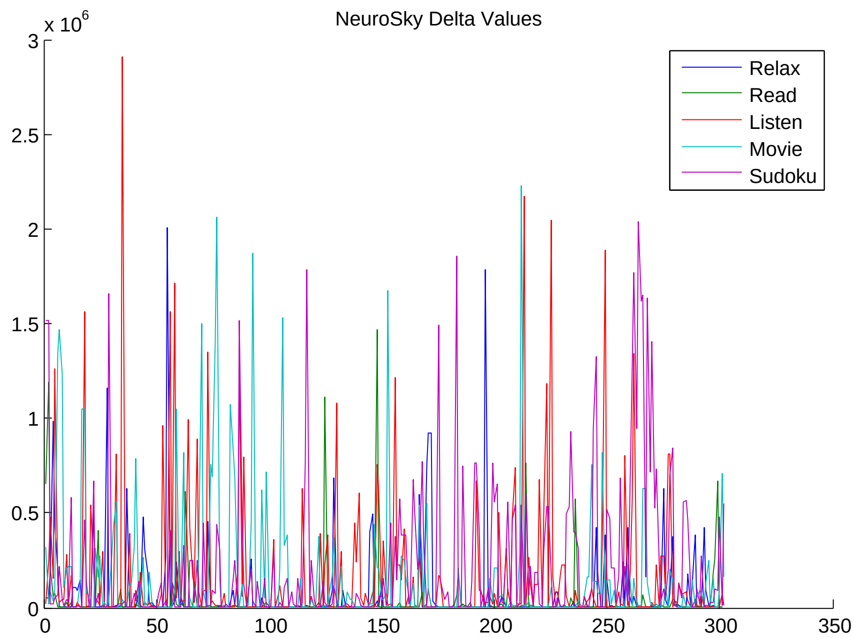


Figure 7.2: Comparing NeuroSky Delta values during the 5 tests - Subject 6 (Male, 23 years)

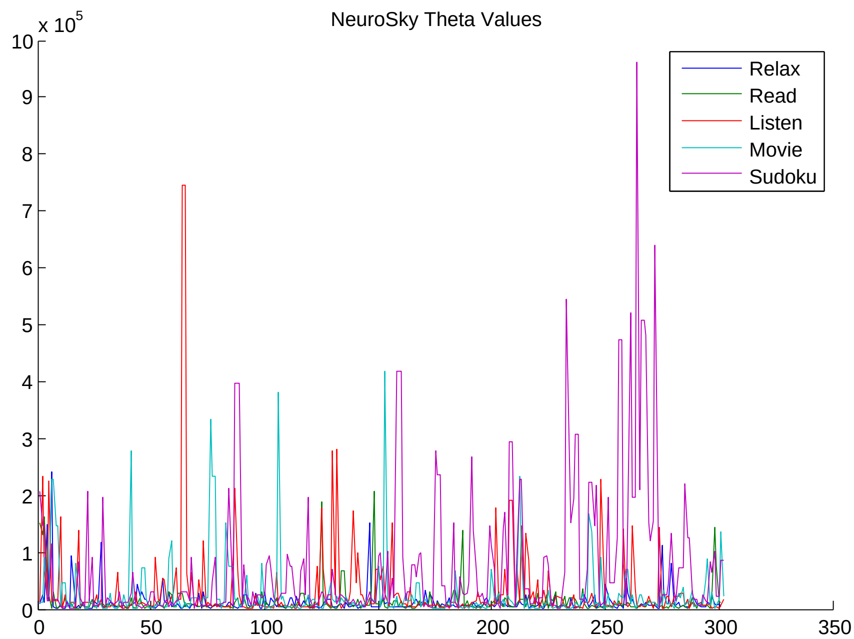


Figure 7.3: Comparing NeuroSky Theta values during the 5 tests - Subject 6 (Male, 23 years)

Visual inspection of delta and theta NeuroSky power values indicate no major differences per task

Figures 7.2 and 7.3 show the NeuroSky power values for the delta and theta frequency bands during the 5 tests. As was presented previously in section 3.1, the delta EEG signal represents deep sleep and the theta EEG signal represents imagination and fantasy. No major visual differences in delta and theta in each of the 5 tasks were found.

NeuroSky Alpha, Beta and Gamma for subject 6

Figures 7.4, 7.5 and 7.6 illustrate the alpha, beta and gamma values over each of the 5 tests for subject 6. Taking a close look on the alpha values, it can be noted that there are higher peaks during test 1, $Relax$ (blue curve), which is consistent with the fact that alpha waves represent meditations and calmness in a concious state as it was previously discussed in 3.2.1.

While some peaks from the $Relax$ test can also be seen in the midrange to high Beta waves, representing thinking or alertness, in figure 7.5, on average, test 5 $Sudoku$ (purple curve) shows consistent higher peaks during the 5 minutes. The cyan curve, $Movie$ test, also shows higher peaks concentrated in the first minute of the test which suggests that the user was more alert and concentrating at that time but then the peaks subsided at the end of the test. The Beta values for the $Listen$ test were also relatively higher during the whole duration of the test.

Gamma values in figure 7.6 were consistently higher for the $Sudoku$ and $Listen$ tests, lower for the $Relax$ and $Movie$ tests, and finally lowest for the $Reading$ test. Gamma waves represent motor activity and higher mental effort.

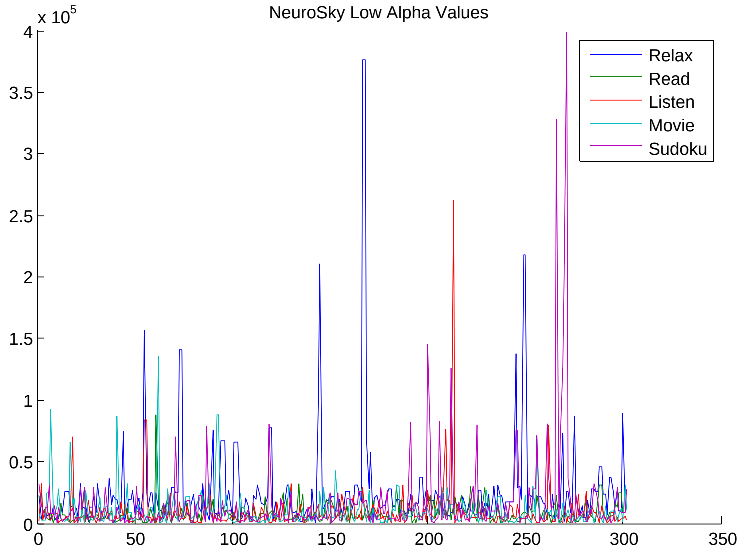


Figure 7.4: Comparing NeuroSky Alpha values during the 5 tests - Subject 6 (Male, 23 years)

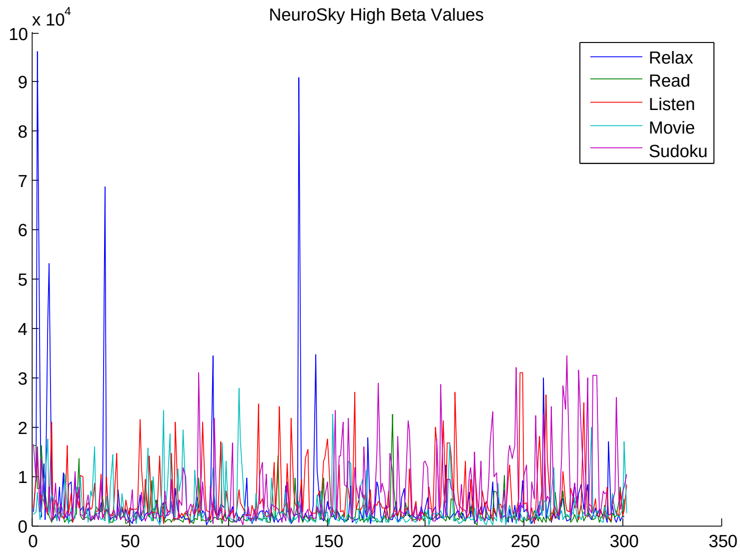


Figure 7.5: Comparing NeuroSky Alpha values during the 5 tests - Subject 6 (Male, 23 years)

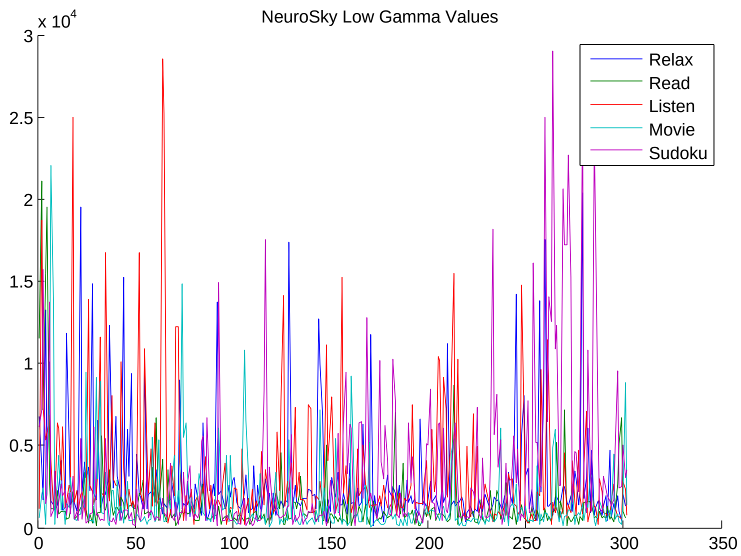


Figure 7.6: Comparing NeuroSky Alpha values during the 5 tests - Subject 6 (Male, 23 years)

Findings from visual inspection: Subject Dependent System

The visual inspection and intra-subject comparison phases were found to be extremely benificial in giving an overview of the subject data and what information can be extracted from it. It also helped in building an understanding on which features would be most useful for extraction which is extensively described in the next module 6.4. However, visual inspection of subjects proved that building a BCI system using NeuroSky, which represents an off-the-shelf commercial device, is extremely subject dependent. This is due to the fact that different subject comparisons even by visual inspection proved to be very different from one another as will be discussed in more depth below.

7.2 Data Classification

In this section the classification process will be explained thoroughly. All classifiers discusse were built using Weka offline classification tool (4.6) and using the generated ARFF files databse that was explained in section 6.4. The classification techniques discussed in section 4.4.4 , namely Naïve Bayes, Bayesian Networks and J48 Decision Trees, were the main techniques used. Both subject dependent and subject independent classification using various segmentation techniques will be discussed. Results of clasification with and without applying dimensionality reduction via attribute selection will also be presented.

7.2.1 Classifier Evaluation Options

In data mining, there are two main ways of evaluating any classification model: using cross-validation or providing a separate test set. An explaination of both classier evaluation options is due before moving on to describe the classifiers used.

Cross-Validation

Stratification ensures that trainins and test sets are balances and representing all classes

Cross-validation is a classification evaluation method that is often used in practical situations with datasets of limited size. The standard way of predicting classifier error rate in this case is to use stratified 10-fold cross-validation. The process of stratification is done to ensure that random sampling is done on the dataset in such a way that guarantees that each class is properly represented in both training and test sets Witten et al. [2011].

Cross-validation is a robust process in case of datasets of limited size

To perform 10-fold cross-validation the data is divided randomly into 10 parts in which each class is represented in approximately the same proportions as in the full dataset. Each part is held out in turn and the classifier is trained on the remaining nine parts; then its error rate is calculated on the holdout test set. This means that the learning procedure is executed a total of 10 times on different training sets and finally, the 10 error estimates are averaged to give an overall error estimate.

Training and Testing Sets

The typical training to testing ratio is 66% to 33%

In case of larger datasets, the most optimal way of testing classifier performance is to divide the dataset into separate training and testing sets. Typically, most of the literature tends to use a two-thirds to one-third training to testing ratio. In this case the classifier is first trained with training data labelled with class values, then tested afterwards with the unseen dataset and the output is one final error rate.

Weka training and testing options

Weka provides both, the cross-validation option, with any arbitrary number of folds, and the separate testing and training files options. It also provides the ability to have one file for the whole dataset and specify a percentage split (with or without perserving order).

Both classifier evaluation options were experimented in the work presented here with different training and testing ratios. For each classifier, the training and testing ratios as well as the algorithms used will be presented.

7.2.2 Classification Algorithms

The two main classification algorithms used in this project were *Bayesian Networks* and *Decision Trees (J48)*. Both classification algorithms were previously introduced in chapter 4. Both algorithms were recommended in the literature and proved promising in case of mental task classification via EEGs.

Bayesian Networks were used due to their probabilistic nature. They were supported in many similar works including Lee and Tan [2006] which is of a similar nature to the work done in this project. Bayesian Networks also consider in dependence of different attributes and hence perform well for multi-class models with correlated attributes.

Bayesian Networks were used due to their good performance in related works

Decision Trees were used mainly because of their simplicity, robustness, and the fact that they are parameter-free. Previous work in the field of EEG classification also implemented decision tree classifiers such as Rong et al. [2007] which also used Weka for classification. The Weka implementation of the C4.5 Decision Tree is named $J48$.

Decision Trees were used due to their simplicity

The Weka implementation of both algorithms was used in all the classification trials recorded in the rest of this chapter. The Weka implementation of the Correlation-based Feature Selection (CFS) and Information Gain (IG) Feature Selection that were introduced in 4.3.1, were also used. For each of the classification trials discussed in the rest of this chapter, the chosen classification technique and feature selection algorithm are mentioned.

7.2.3 Batch Processing

In order to automate the process of creating training and testing ARFF files with arbitrary segment lengths and arbitrary training and testing ratios, the use of Weka through the commandline and through Matlab was crucial. To make this possible, the Weka was ported to Matlab and the free

The use of Weka from Matlab allowed for easier batch processing of the dataset open-source library $Weka2Matlab$[1] was used in conjunction with scripts adapted it to the system.

Using Weka from Matlab allowed for applying feature selection techniques on the training and subsequent testing files automatically. It also facilitated the creation of different ARFF databases with different feature selection and segmentation options, and finally, different classifiers.

7.2.4 Subject Independent Classification

Preliminary trial to classify tasks from all subjects led to expected poor results As a preliminary basic trial, subject independent classification using 10 second segments without feature selection was experimented. The results were not expected to be very promising due to the finding from the preprocessing and visual inspection discussed in 7.1. In this test a Multi-layer perceptron Neural Network (4.4.4) was also used. The results are displayed in table 7.1.

Multilayer Perceptron	Bayesian Network	J48 Decision Tree
34.5%	30.7%	26.6%

Table 7.1: Subject Independent Classification - 10s Segments - No Attribute Selection (31 Features)

Per-subject models are dominant in BCI literature due to the high signal variance

By observing the results of subject independent classification from table 7.1 it can be inferred that a generic system cannot be built. This result was expected since most BCI literature (Lee and Tan [2006], Palaniappan [2005], Keirn and Aunon [1990], Keirn [1988]) constructed per-subject models due to the high variance of EEG signal properties between subjects. Hence any BCI system, whether commercial or medical, needs to be trained by the subjects before any classification or feedback can be correctly received.

[1] http://www.mathworks.com/matlabcentral/fileexchange/21204-matlab-weka-interface

7.2.5 Subject Dependent Classification

After reaching the conclusion that subject-independent classification our current BCI system will not produce any reliable results, subject-dependent classification was attempted using different permutations of task pairs, classifiers, segmentations and attribute selection algorithms. It is important to note that only significant and important classification results are reported here. Many more trials with different task permutations were done, however, only the relevent and useful results are presented below.

It is also important to mention that the testing procedure in the following subsections was done by taking out a separate holdout test set that has not been seen by Weka prior to testing. When feature selection was applied, it was applied on the training set and the chosen set of attributes were selected from the test set. This was done to ensure that no overly optimistic results due to the choice of the best attributes from the whole set to try to keep the results closer to the real-life situation as possible.

Feature selection was done on the training set and imposed on the test set

Two-Task Classification

As a first classification step, the ARFF data was divided into training and testing datasets according to the percentages discussed previously. Features were calculated for 5 second segments as well as for 1 second segments for all the tasks. ARFF files containing permutations of pairs of classes were saved and classified. The most significant results are presented in this section.

Table 7.2 shows results from compairing every task with the baseline $Relax$ task with features calculated for 5 second non-overlapping segments. Separate training and testing files were used with 3 minutes for training (60%) and 2 minutes for testing (40%).

The classifier used in this case was the Naïve Bayes classifier since the comparison is only between 2 classes. No feature selection was done and the features used included all fea-

The simple Naïve Bayes classifier was used to compare class pairs tures in table 6.1 except features 6 (RMS) and 29 (Cepstral Coefficient).

Subject Number	Relax/Sudoku	Relax/Read	Relax/Listen	Relax/Movie
1	100%	100%	50%	50%
2	100%	94%	50%	50%
3	50%	50%	50%	50%
4	50%	44%	50%	54%
5	85%	50%	0%	50%
6	35%	50%	50%	0%
7	50%	100%	65%	50%
8	50%	50%	10%	90%
9	50%	98%	0%	50%
10	50%	50%	50%	64.5%
11	10%	83%	50%	50%
12	52%	67%	50%	50%
13	50%	60%	54%	6%
14	50%	100%	50%	75%
15	50%	50%	52%	0%
Mean	55.5%	69.7%	42%	45.9%

Table 7.2: Comparison of baseline task (Relax) to all other tasks- 5s Segments - No Attribute Selection (31 Features)

Analysis of table 7.2 shows a maximum average of 69.7% accuracy for 2 tasks

By analysing the data in table 7.2 it can be seen that the highest mean classification accuracy was 69.7% in case of comparing $Relax$ versus $Read$ tasks with a segment length of 5 seconds. For most subjects in this category the classification was much higher than chance (50%), however, the results are considered relatively poor for Relax versus Listen and Relax versus Movie which were lower than chance and only slightly better than chance in the case of $Relax$ versus $Sudoku$. This poor classification output might be due to overfitting due to using all 31 features and applying no feature selection. Another reason might be that the training sets are too small due to the segmentation process. The use of a simple classifier, the Naïve Bayes can also be a reason for the low classification accuracies. The next trials attempted to overcome those drawbacks.

In an attempt to enhance the classification accuracy for class pairs, the length of the segments were decreased to 1 second and for each second all features were calculated. A

Bayesian Network classifier was chosen for classification and the output accuracies were calculated. The training set was increased to 3.3 minutes (66.6%) and the test set was decreased to 1.7 minutes (33.3%) to have a better representing and more balanced training corpus.

In addition, the Weka implementation of the Correlation-based feature selection algorith, named $CfsSubsetEval$, was applied as a preprocessing filter (previously introduced in 4.3.1). The algorithm chooses a subset of the features per subject. The results are presented in table 7.3.

Subject Number	Relax/Sudoku	Relax/Read	Relax/Listen	Relax/Movie
1	50%	50%	50%	50%
2	50%	50%	50%	50%
3	54%	77%	50%	63%
4	0%	23%	75%	30%
5	48%	100%	92%	55%
6	50%	50%	50%	50%
7	50%	50%	100%	50%
8	44%	50%	66%	36%
9	81%	50%	50%	100%
10	100%	50%	50%	12%
11	50%	50%	50%	50%
12	50%	50%	100%	100%
13	50%	50%	23%	50%
14	43%	91%	19%	50%
15	50%	50%	50%	100%
Mean	51.3%	56.1%	58.3%	56.4%

Table 7.3: Comparison of the baseline task (Relax) to all other tasks - 1s Segments - CFS feature selection was applied

Using 1 second segments and feature selection enhanced some results but decreased others

Contrary to what was expected, as table 7.3 shows, using 1 second segments, larger testing datasets and feature selection did not enhance the results of the 2 classes that were seen to have highest results in table 7.2. However, using shorter segments and feature selection seemed to balance the results especially for the $Relax$ versus $Listen$ and $Relax$ versus $Movie$ tasks. The overall performance was still not promising however, using the same setting, task permutations of all other task pairs was attempted and the results are shown in table 7.4.

Sub. Num.	Listen/ Movie	Listen/ Sudoku	Movie/ Sudoku	Read/ Sudoku	Read/ Listen	Read/ Movie
1	50%	4%	50%	50%	50%	50%
2	50%	50%	0%	36%	50%	0%
3	81%	84%	50%	22%	8%	50%
4	50%	50%	50%	50%	50%	35%
5	80%	14%	50%	100%	73%	100%
6	50%	0%	0%	0%	50%	81%
7	100%	50%	99%	67%	50%	100%
8	0%	50%	50%	64%	50%	75%
9	50%	50%	50%	50%	50%	50%
10	50%	50%	50%	50%	50%	50%
11	50%	50%	44%	50%	100%	50%
12	50%	50%	50%	68%	99%	50%
13	94%	56%	83%	97%	50%	49%
14	50%	100%	56%	50%	50%	50%
15	0%	100%	40%	100%	50%	100%
Mean	53.6%	50.5%	48.1%	51%	55.3%	59.3%

Table 7.4: Comparison of pairs of tasks without the baseline task (Relax) - 1s Segments - CFS feature selection was applied

Results were higher than chance only for certain subjects

Results from table 7.4 as expected, were only slightly higher than chance with a highest value of 59.3% in case of $Read$ versus $Movie$. The results from the table show that classification is completely subject dependent. Whereas for some subjects the accuracies were all over 50%, for others it went as low as no correctly classified instances.

Three-Task classification

Enhancement 1: Training set increased to 4 minutes (80%)

In order to proceed to compare 3 task permutations, some enhancements to the classification procedure had to be made to attempt to get more promising results. First, the training samples were increased to 4 minutes (80%) instead of 3.3 minutes (the typical 66.6%) and respectively the testing datasets were decreased to 1 minute (60 samples in case of 1 second segments).

Enhancement 2: Applying feature selection on the whole dataset

The second enhancement was that feature selection was applied to the whole dataset prior to division into training and testing sets. While this second enhancement can lead to over optimistic results, it would give an overview of the best features from the feature set that represent the whole data. It might also enhance the accuracy of the investigated classifications and allow for better understanding of the behaviour of the dataset. However, when building applications based on the results of this system, this altration has to be kept in mind.

Given that there are 5 tasks, the number of 3-class permutations is 10 trials. While this number is quite large for recording, the best scoring trials are displayed in this subsection. First, the best scoring tasks from the 5 second segments of pairwise classification (table 7.2) were chosen and combined for 3-class classification using 5-second segments, namely $Relax$, $Read$ and $Sudoku$. The results are shown in table 7.5 with two classifiers, Bayesian Networks and J48 deicision trees, with and without feature selection where *AS* refers to *Attribute Selection*.

Sub. Num.	Bayesian No AS	Network AS	J48 No AS	Trees AS
1	90%	100%	60%	58.3%
2	33%	55%	20%	20%
3	50%	20%	20%	20%
4	25%	100%	20%	100%
5	20%	63.3%	20%	60%
6	93%	100%	20%	100%
7	60%	60%	20%	100%
8	60%	100%	93%	50%
9	60%	90%	20%	100%
10	20%	60%	20%	20%
11	60%	60%	20%	50%
12	100%	100%	100%	60%
13	97%	97%	60%	60%
14	48%	100%	60%	60%
15	60%	60%	60%	20%
Mean	58.4%	77.6%	40.8%	59.8%

Table 7.5: Comparison of classification of 3 tasks: Relax, Read, Sudoku with and without CFS feature selection using Bayesian Networks and J48 Decision Trees - 5s Segments

Feature selection enhanced the results by approximately 20%

By inspecting the results in table 7.5, it can be noted that feature selection enhanced the classification results by almost 20% in both Bayesian Networks and J48 tree classifiers. The Bayesian Network classifier appears to have a better performance and outputs an average accuracy of 77.6% with feature selection.

Combining highest scoring class pairs to get 3-class 1-second classifications

Since the preliminary 3-class results seemed promising, more 3-class permutations using 1 second segments and feature selection to the whole dataset were attempted. Deciding to put aside the baseline *Relax* task, the highest 4 scoring permutations from 7.4 were permuted to give the following 3-class classifications presented in table 7.6

Subject Num.	Sudoku Read/Movie	Read/ Listen/Sudoku	Movie/ Read/Listen	Listen/ Movie/Sudoku
1	33 %	96%	44%	67%
2	67%	99%	70%	46%
3	24%	53%	35%	45%
4	67%	59%	84%	80%
5	33%	63%	60%	68%
6	51%	59%	26%	43%
7	33%	29%	63%	33%
8	64%	86%	64%	33%
9	54%	29%	54%	39%
10	67%	26%	90%	67%
11	33%	63%	62%	17%
12	37%	74%	91%	62%
13	14%	86%	46%	33%
14	33%	68%	97%	67%
15	0%	66%	62%	17%
Mean	41%	56.1%	63.2%	47.8%

Table 7.6: Classification of three tasks using CFS feature selection - 1s segments

Poor results led to investigating other evaluation options

The classification trials were repeated using the rest of 3-class permutations, 4-clsas and all-task permutations with changing the segmentation options and the classification algorithms. However, the very low to poor results for some subjects led to the suspicion that the data might either be too little for training, or suffered from time-dependent noise.

To first overcome the problem of having a small training dataset, the classification trials with 3,4 and 5 classes using Weka's cross-validation evaluation option were attempted. The results were unusually always between 97% to 100%. These abnormally high results turned out to be due to the Weka's randomization process which is applied to the dataset prior to cross-validation.

Using Weka CV produced unusually highly accurate results due to order randomization

The randomization filter in Weka randomizes the order of the instances in the ARFF file so that the training and testing sequences do not come from the same sequential recording second. These results strengthened the belief that there might be time-dependent noise being incrementaly added to the signal and causing this effect. This hypotheses will be tested and discussed in the following subsection (7.2.6).

Weka randomizes the instance orders prior to cross-validation

7.2.6 Time Dependency

As explained previously in chapter 3,EEG signals are considered weak signals that are highly susceptible to noise. In addition, any data collected from sensors usually suffers from what is called *sensor drifting*. This phenomenon was previously noted in activity recognition works using different sensors like accelorometers as well as in EEG Hennighausen et al. [1993].

Electrode Drift Problem

Electrode drift and time dependency in EEG data has been previously mentioned and addressed in previous EEG works such as Lee and Tan [2006]. However, since no prior similar study has been done with particularily the NeuroSky Brainband device, there was no way to ensure that this problem will be experienced prior to doing the previously mentioned tests.

It is common in EEG signals to change over time due to many reasons including: slight shift in electrode positioning over time due to subject movement, gel dryness over time in wet electrode BCI systems and additional signal noise over

Time dependency is dealt with using stochastic filters like the Kalman filter time due to the device itself. Dealing with time dependency and sensor drifts in comonly treated by applying stochastic filters to the dataset prior to training the classification model. However, due to the time constraint in this work, an alternative way to try to decrease the effect of time-dependency on classification.

The third minute was used for testing to try balance out the time dependency

In order to decrease the effect of time-dependency, an alteration was made to the dataset. The test set was changed to be the middle minute (minute 3) of the 5 minute per-subject recording instead of the last minute. This particular minute was chosen so as to have a balanced number of instances that occured chronologically before and after the test set so as to balance out the added time-dependent compoenent. In the following subsections, results from classifying pairs of tasks, 3 tasks, and all the 5 tasks are shown. Comparisons between using the third and fifth minutes for testing are also presented to show the enhancement of this alteration.

Two-task Classification

Table 7.7 shows the classification results with Bayesian Netowork classifier for the baseline tasks $Relax$ with all other tasks when using the third minute as the testing minute in order to reduce the effect of the time dependent noise.

Three-task Classifications

It is important to also mention that 3-class classifications were also attempted, however, not with all task permutations due to timing contraints, the baseline task was left out and the 4 other classes were permuted. Table 7.9 presents the results.

Five-task Classification

An all class classification trial was attempted, despite knowing that the results will not be very promising. However, the purpose of this trial was to show the difference and enhanced accuracy upon using the third minute for testing instead of the fifth minute. As table 7.10 shows, using the third minute for testing gave an average results of 38.7%

Sub. Num.	Relax/Sudoku	Relax/Read	Relax/Listen	Relax/Movie
1	100%	100%	76%	100%
2	100%	34%	100%	100%
3	54%	100%	83%	83%
4	98%	95%	100%	100%
5	50%	63%	100%	80%
6	50%	59%	70%	100%
7	100%	50%	50%	100%
8	88%	100%	100%	23%
9	100%	100%	100%	100%
10	100%	98%	100%	100%
11	66%	88%	100%	55%
12	50%	50%	100%	96%
13	50%	66%	50%	50%
14	77%	100%	100%	100%
15	92%	50%	100%	73%
Mean	78.3%	76.8%	88.6%	84%

Table 7.7: Comparison of classification of baseline task (Relax) to all tasks using the third minute for testing and applying CFS feature selection- 1s segments

with an increase of 13% over using the fifth minute (25.1% average). In addition, in some subjects the enhancements was close to 40% increase in the classification accuracy. This is a strong support for the hypothesis aforementioned which is that the EEG signals are time-dependent.

7.3 Discussion

The many attempts and the varying classification accuracies illustrated throughout this chapter only show the complexity of EEG data and prove the need for complex classification systems in order to achieve high and reliable accuracies for real life applications. In this section, the results presented previously will be analysed and concluded.

From tables 7.2,7.3,7.4, 7.7 and 7.8 it can be seen that classification results for 2 classes are relatively good. For particular classes the classifier performance is better and this can be due to one of the following:

Sub. Num.	Listen/ Movie	Listen/ Sudoku	Movie/ Sudoku	Read/ Sudoku	Read/ Listen	Read/ Movie
1	100%	100%	0%	50%	100%	100%
2	100%	100%	50%	100%	95%	50%
3	96%	100%	88%	100%	100%	65%
4	100%	83%	100%	100%	80%	50%
5	100%	50%	91%	50%	100%	50%
6	100%	50%	50%	50%	99%	100%
7	50%	92%	98%	56%	100%	50%
8	50%	64%	88%	100%	86%	50%
9	50%	100%	78%	100%	100%	50%
10	82%	100%	100%	90%	100%	100%
11	100%	100%	100%	100%	100%	50%
12	97%	60%	100%	50%	98%	37%
13	53%	76%	100%	41%	50%	100%
14	68%	100%	100%	8%	100%	60%
15	0%	53%	95%	60%	50%	94%
Mean	76.4%	81.8%	82.5%	70.3%	90.5%	67.1%

Table 7.8: Classification of pairs of tasks without the baseline task (Relax) using the third minute for testing and applying CFS feature selection - 1s segment

1. Signal aquisition during the test of this subject encountered less noise (less movement artifacts)
2. Certain tasks are more separable than others depending on the brain activity associated with this task

It is impossible to assume a one-to-one matching between mental tasks and brain lobees

In chapter 3, the different cerebral cortex lobes were presented and the functionalities of each lobe were explained. The tasks presented in this study cannot be all referred to a certain brain lobe because it is not a one-to-one mapping. However, certain sensory tasks can be majorly associated to particular brain lobes. For example, listening, as a sensory task, is associated with temporal lobe activity. Nevertheless the interpretation of what is being listened to, language understanding, is associated with frontal lobe activity. This might be an indication of why certain tasks or task pairs gave better classification results using the single NeuroSky frontal lobe electrode.

1. *Relaxation:* When trying to meditate, the concentration

Subject Num.	Read/ Listen/Sudoku	Listen/ Movie/Sudoku	Movie/ Read/Listen	Sudoku/ Read/Movie
1	69%	64%	93%	65%
2	57%	51%	33%	33%
3	67%	100%	58%	57%
4	39%	52%	61%	58%
5	33%	33%	84%	33%
6	67%	67%	100%	67%
7	47%	34%	38%	33%
8	67%	61%	41%	67%
9	65%	54%	49%	100%
10	82%	67%	73%	89%
11	100%	100%	84%	86%
12	67%	76%	62%	34%
13	0%	8%	48%	66%
14	40%	48%	79%	39%
15	0%	19%	32%	32%
Mean	53.3%	55.6%	62.3%	57.2%

Table 7.9: Classification of three tasks using the third minute for testing and applying CFS feature selection - 1s segments

in the relaxation process itself leads to high frontal lobe activity [Banquet, 1973]. However, since in this study the subjects were asked to relax and not necessarily concentrate on meditating, it cannot be claimed for sure that high frontal activity existed. It is also important to note that since comparing tasks to the baseline task (tables 7.2, 7.3 and 7.7) led to high classification results, this means that there was a relatively high difference in brain activity (in this frontal lobe activity) between the baseline task $Relax$ and all other tasks.

2. *Reading:* Associated with many parts of the brain. The occipital lobe is responsible for the vision of the letters whereas the frontal and parietal lobes are responsible for interpreting what has been read. Since the subjects were asked to concentrate on the story being read and since the story was dramatical, heightened emotional activity associated with the frontal lobe should exist [Wren, 2001].

Sub. Num.	Test using minute 5	Test using minute 3
1	12%	44%
2	42%	43%
3	0%	43%
4	33%	40%
5	0%	20%
6	17%	58%
7	33%	40%
8	34%	58%
9	13%	42%
10	57%	55%
11	12%	59%
12	66%	54%
13	18%	0%
14	4%	25%
15	33%	0%
Mean	25.1%	38.7%

Table 7.10: Comparison of classification of the five tasks using fifth minute for testing versus third minute for testing and applying CFS feature selection - 1s segments

3. *Listening:* The temporal lobe is responsible for the sense of hearing, hence temporal lobe activity should be highly present when listening to music [Stewart et al., 2006]. Since in this study, the subjects were instructed to deal as if they are listening to a radio show in the car during the $Listen$ test, it can be assumed that they did not particularly pay attention to the content of the show and hence the frontal lobe activity would be dminished during this task.

4. *Movie:* Watching a movie stimulates many parts of the brain, the first of which is the occipital lobe that is responsible for perceiving visual stimuli. Contemplative movies stimulate the frontal lobe due to concentration and high cognitive activity which is persumed to happen in that lobe [Hasson et al., 2008]. Since the movie presented in this study was a documentary, it is assumed that frontal and occipital lobes activities were present during this test. However, this assumption only holds if the subjects were interested in the documentary material. Although the subjects were instructed to concentrate on the movie, it can never be

ensured that they did.

5. *Sudoku:* Brain puzzles all require memory, concentration, and high cognitive load which are all functions of the frontal lobe [Sobolewski et al., 2009]. Since all subjects attempted to finish the sudoku puzzle, it can be assumed that high frontal lobe activity was present during this test.

Two-task classification of tasks in different lobes is easier

From the above list and the results presented in this chapter it can be concluded that NeuroSky Brainband can differentiate between 2 mental tasks with an accuracy between 67.1% to a maximum of 90.5% in case time dependency is considered. It can also be noted that differentiation between tasks that are supposedly happening in different brain lobes seemed to be easier. For example, the highest classification accuracies were for the $Read$ versus $Listen$ (90.5%) where the reading is assumed to happen in the frontal lobe and the listening in the temporal lobe and yet consistently lower for $Read$ versus $Sudoku$ tests (51% to 70.3%).

Differentiation between the baseline task $Relax$ and the rest of the tasks reched an average between 76.8% and 88.6% which shows that there is elevated brain activity that is sensed by the NeuroSky sensor when performing any mental task compared to the meditative state.

Figure 7.7 highlights the performance enahcements of using 5 second segments, 1 second segments and third versus fifth minute for testing in the $Relax$ versus other tasks classification. Figure 7.8 shows the difference between using the third versus the fifth minute for testing in 2-class pairs.

Figure 7.9 shows classification averages between three classes using third versus fifth minutes for testing. Classification between 3 classes gave the highest result of 62.3% when differentiating between $Movie$, $Read$ and $Listen$ as shown in table 7.6 which can also be due to the fact that those three tasks are happening majorly in three different brain lobes. The $Read$ task is associated with frontal lobe activity, the $Listen$ with temporal lobe activity. Whereas concentration in the $Movie$ task will lead to heightened frontal lobe activity, visual stimuli increase the occipital lobe activity.

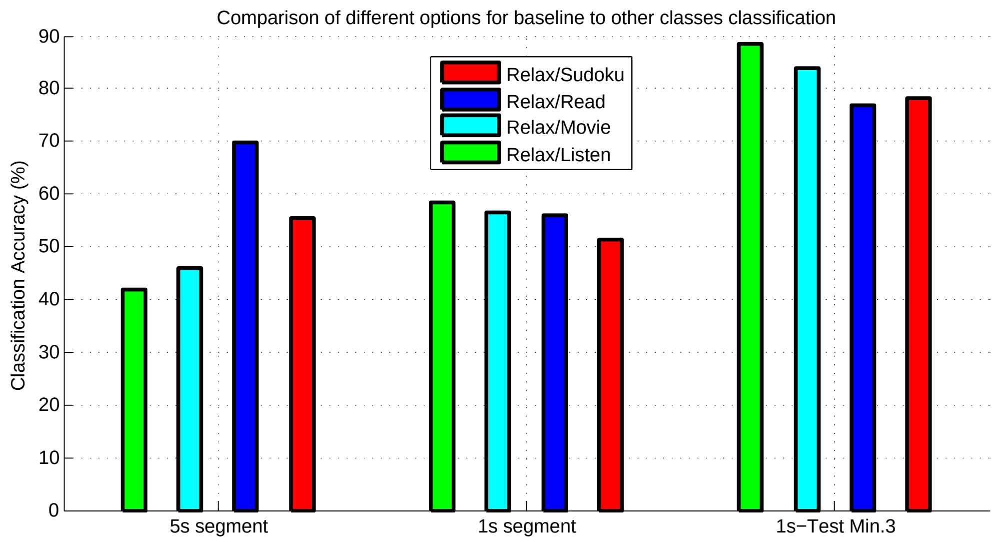


Figure 7.7: Comparing the mean accuracy of classification of the baseline task (Relax) to all other tasks in three settings present on the x-axis: using 5s segments, using 1s segments, and using 1s segments with the third minute for testing

Nevertheless, this cannot be assumed true for all users.

It can also be concluded that NeuroSky is unable to accurately differentiate between 4 or more tasks due to its limited capabilities and single electrode as well as due to the complexity of the chosen tasks. However, a wide range of applications can still utilize task classification by NeuroSky using 2 or 3 tasks and more complex classifiers can be build in order to help classify more mental tasks.

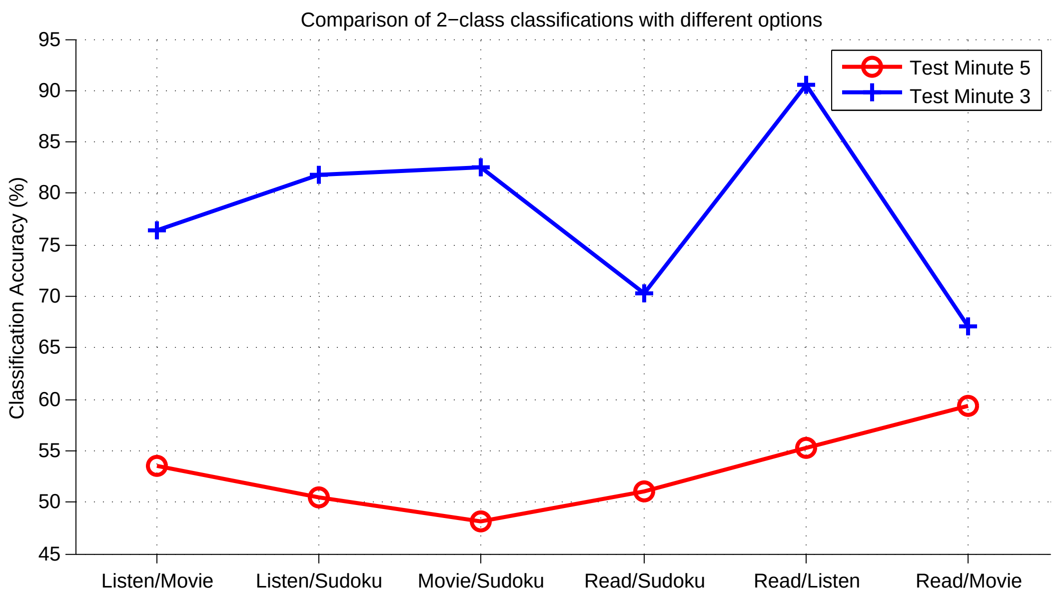


Figure 7.8: Comparing the mean accuracy of classification of pairs of tasks when using the fifth minute for testing (red) versus when using the third minute for testing (blue)

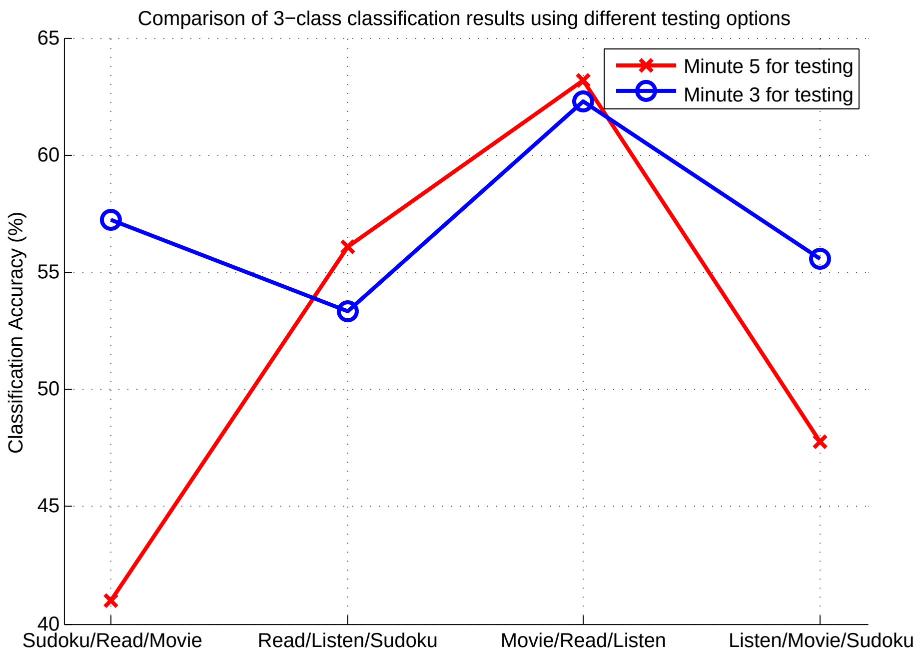


Figure 7.9: Comparing the mean accuracy of classification of three tasks (tasks on x-axis) when using the fifth minute for testing (red) versus when using the third minute for testing (blue)

Chapter 8

Limitations and Shortcomings

The work presented in this thesis was subject to many limitations imposed by the hardware and the software used, the data aquisition techniques and the system evaluation process. This chapter aims to highlight those limitations and shortcomings and introduce ways in which those shortcomings can be avoided in future works.

8.1 Device Limitations

In chapter 4, the various types of BCI systems were introduced and the differences between different EEG devices was highlighted.

The EEG device used in this work, a Myndplay Brainband device which is powered by a NeuroSky chip, was a single dry-electrode EEG device with only one EEG channel and a baseline electrode placed on the left ear. Whereas this made the device more portable, the setup faster, and the cost of the device lower than traditionally expensive devices, this meant that the device only collects EEG signals from one location, the FP1.

Only one electrode reduces the dimensionality of the EEG data collected

As was explained before, each of the brain lobes is responsible for certain mental and cognitive tasks. Neurons firing from all brain lobes during certain tasks generate an electric potential large enough to be picked up by EEG electrodes. However, collecting signals from one electrode location leades to the inability to identify more than 2 mental states with high accuracy. This limited data also did not make it possible to compare the brain lobes' responses to different mental tasks and limited the nature of the identifiable tasks to frontal lobe related tasks such as concentration or meditation.

Additionally, the device could operate for almost 8 hours without the need to recharge. However the power of the device decreased by time which could be one of the reasons for the corrupt signals collected from some users that was observed in the signals and discussed in 6.3.1.

8.2 Data Acquisition and Filtering Limitations

During the data aquisition procedure there were some limitations that can be avoided if similar experiments are to be repeated in the future. The first of which, in this study the device has been kept open between each of the tasks and during the whole duration of the study. This led to incrementally increasing the noise (7.2.6 added to the signals being collected and did not provide equal basis for comparison between different tasks balancing all variables.

EEG collection sessions should be repeated on separate days

Another limitation of the data aquisition process was that the whole session was done for each user on the same day and only for a duration of 5 minutes for each task. It would have been better to do the sessions on separate days repeating the tasks again, as well as taking longer recordings to provide a better training corpus for the classifier. Taking the same test on separate days or repeating certain tasks will help give an idea about the relation of the EEG signals to the state of the user that day and normalize this effect.

8.3 Classifier Limitations

Limitations of Weka include the imposed randomization during cross-validation

Classification in this thesis was done using Weka open-source classification tool. While Weka provides many options for filtering, visualization of the data and classification algorithms, it can be limited when it comes to customizing evaluation schemes. This problem was faced when attempting to use cross-validation instead of a separate test file where Weka imposes the application of a randomization filter which randomly changes the order of the instance as a pre-step to cross validation. Randomization in case of this project, would lead to inaccurate and overly optimistic results. The EEG datasets collected are time-dependent and highly correlated and randomizing the order would lead to using highly correlated data in the training and test sets of the cross-validation folds producing overly optimistic and false results.

In addition, as explained in subsection 7.2.5 in order to try and enhance the classification accuracy due to the limited training set size, feature selection was applied on the whole dataset. This is a limitation in the classification process since it would mean that the test set was seen and the features best representing the training and testing sets were used. In practice, this might lead to over optimistic results.

As a final point in the shortcomings of the classification process, all classifications investigated in this thesis were done in an offline mode. This was due to two main reasons: the use of Weka did not promote online classification, and secondly, building an online classification system is a complex task that needed this prior research. An online mental task classification system would be a good application to the findings of this thesis.

Finally, the knowledge of those limitations and shortcoming can help to avoid them in future work or enhancements of this system or similar commercial BCI systems, which is one of the contributions this thesis intended to provide.

Chapter 9

Conclusion and Future Work

9.1 Conclusion

This project was mainly intended to investigate the feasibility of using commercial BCI in mental task classification. After a thorough literature review about the biomedical background of the brain and EEGs (3) and review of the basic blocks for building a BCI system (4) as well as a review of relevent related work (5), a study was designed to collect a data corpus of EEG signals during five mental tasks.

The designed study used the NeuroSky Brainband device to collect data from 20 users with an average age of 23.3 years. The five mental tasks chosen were: $Relax$, $Read$, $Listen$, $Sudoku$ and $Movie$, each task was performed for a duration of five minutes. An Android application was developed to collect the EEG signals from the study participants via wireless Bluetooth connection between the Brainband and the phone. The complete details of the study conditions and collected dataset were discussed in chapter 6.

The collected database of raw and filtered EEG signals is considered a valuable contribution of this thesis since no similar dataset existed previously using the NeuroSky Brainband device. The corpus of collected EEG signals are saved in CSV file format and contain timestamped data for each subject performing five minutes of each task. The dataset will be released online for use in future studies.

After the data acquistion phase, the dataset went through a series of preprocessing filter for artifact removal. This was followed by the feature extraction and segmentation steps. Different segmentation options were implemented where features were calculated for ten, five and one second segments and saved in the format supported by the classification tool, Weka. Prior to the classification step, visual and statistical inspection and analysis of the data was done and discussed in section 7.1 to get an overview of the properties of the collected signals.

Finally, the classification step was presented in chapter 7 using mainly Bayesian Networks and Decision Trees. Classification results for comparing the baseline task with other tasks were first compared using different segmentation and feature selection options. This was followed by the clasification of pairs of tasks without the baseline task, then combinations of three tasks were classified. Finally, the classification of all the five tasks was also attempted.

Classification results using a random order of instances for the training and testing sets suggested that a time-dependent component might be present that is hindering higher classification results. Due to timing constraints of the projects, this problem was tackled by choosing different test datasets from the middle of the data corpus to minimize the time-dependent component. The third minute of the signal recording for each task was then chosen for testing instead of the fifth minute and subsequently, all classification trials attempted previously, were attempted again using the new setting.

Results using the new setting were found to be higher than the previous trials in most cases, which proved that a time-dependent component exists. The new classification results were also recorded in chapter 7.

This project attempted to answer the research questions that were presented in 2.4. Below are the findings that answer those aforementioned questions:

- In answer to the research question of whether off-the-shelf BCI can be used to classify mental tasks, the system presented in this thesis proved to be able to classify to an acciracy between 67% to 90% between two mental tasks. For three mental tasks the classification accuracy averagr decreases to 41% to 63.2%. For a higher number of mental tasks and with the given classifiers, it is not possible to get reliable results.

- In answer to the research question of what information can be extracted from single dry electrode BCI systems, the system presented provided many answers. First, information about the different EEG frequency bands extracted from the $Fp1$ position (previously explained in figure 4.2) gave information about the cognitive state of the system users as presented in 7.1. The system also gave insight about the fact that it is easier to differentiate between mental tasks having different brain lobe acitivites. For example, the ghiest accuracy of 2 task classification was achieved when a frontal lobe task ($Read$) was classified against a temporal lobe task ($Listen$). Although this point needs further investigation, this research can assume the hypothesis that NeuroSky Brainband can differentiate between tasks happening in two or more different brain lobes.

- In answer to the research question of what limitations and challenges face commercial BCI systems, the system presented in this thesis provided a thorough overvoew of commercial BCI challenges in 1. In addition, the limitations and shortcomings of the data aquistion device and classification process were presented in 8.

To conclude, this thesis aimed to contribute to the ongoing research in the world of BCI. It focussed on the use of off-the-shelf, portable and affordable BCI systems and investigated the possibility of extending its uses beyond gaming and into mental task classification. The thesis proved that classification of 2 or 3 mental tasks is feasible using the NeuroSky Brainband device and shed light on the challenges and limitation that were faced. Finally, the future work proposed in this field is presented in the following section.

9.2 Future Work

There are many opportunities and possibilities for future work based on the findings of this system.

First, the investigation of the effect of different brain lobe activities on classification accuracy using the NeuroSky Brainband can be further researched. The limitations of the data aquistion process covered in 8 can be avoided and longer training data can be aquired on separate session days and with reseting the sensor between tasks.

The segmentation process can also be enhanced to attempt to overcome the time-dependency problem as well as to produce a more correlated and balanced dataset. This can be done by segmenting the data into overlapping windows instead of non-overlapping ones as suggested in previous research such as Lee and Tan [2006]. In addition, stochastic filtering can be applied to decrease from the time-dependency effects.

Other classifiers can also be used to enhance the classification process. Research has shown that the performance of Artificial Neural Networks (ANN) and Support Vector Machines (SVM) perform particularily well in mental task classification. The performance of those classification algorithms using the NeuroSky dataset collected would be interesting to investigate.

Whereas the future work suggested above is tackling enhancements of the current system, possibilities of using the system to build applications already exist. A suggested application would be using NeuroSky Brainband's mobile Android API for developing a mobile application for tracking the user's cognitive load. The user can be provided statistical information about his brain activity at the end of the day including how many hours of concentration, and information about the nature and duration of the cognitive and mental tasks performed throughout the day.

Applications that involve NFT (Neurofeedback Training) and rehabilitation of users suffering from strokes or brain injuries can also be implemented to help in tracking their mental progress post brain surgery or injury. The simple setup and ease of use of commercial BCI devices such as the NeuroSky Brainband can be utilized by developing rehabilitation applications that do not need strict monitoring from medical staff.

Another field of application that can be investigated utilizing the findings of this project, is the safety and security field. This classification results presented in this project suggest that EEG signals are extremely subject dependent. This promotes the research of using BCI in identifying users via concentration or relaxation brain-prints for security applications.

To conclude, there is a wide range of applications that can be implemented that utilizes the portable, affordable and usable off-the-shelf BCI systems available in the market nowadays in a variety of fields. Those fields include self-tracking, rehabilitation, security and brain training, as well as the already flourishing field of gaming and entertainment.

Part IV

Bibliography

Bibliography

C.W. Anderson and Z. Sijercic. Classification of eeg signals from four subjects during five mental tasks. In *Solving Engineering Problems with Neural Networks: Proceedings of the Conference on Engineering Applications in Neural Networks (EANN96)*, pages 407–414. Turkey, 1996.

K. Ansari-Asl, G. Chanel, T. Pun, et al. A channel selection method for eeg classification in emotion assessment based on synchronization likelihood. In *Eusipco 2007, 15th Eur. Signal Proc. Conf*, 2007.

J.P. Banquet. Spectral analysis of the eeg in meditation. *Electroencephalography and clinical neurophysiology*, 35(2):143–151, 1973.

Biopac. `http://www.biopac.com/`.

A.R.F. Cabrera. Feature extraction and classification for brain-computer interfaces. *Brain*, 2009.

A. Campbell, T. Choudhury, S. Hu, H. Lu, M.K. Mukerjee, M. Rabbi, and R.D.S. Raizada. Neurophone: brain-mobile phone interface using a wireless eeg headset. In *Proceedings of the second ACM SIGCOMM workshop on Networking, systems, and applications on mobile handhelds*, pages 3–8. ACM, 2010.

M.A.G. Correa and E.L. Leber. Noise removal from eeg signals in polisomnographic records applying adaptive filters in cascade. 2011.

J. del R Millan, J. Mouriño, M. Franzé, F. Cincotti, M. Varsta, J. Heikkonen, and F. Babiloni. A local neural classifier for the recognition of eeg patterns associated to mental tasks. *Neural Networks, IEEE Transactions on*, 13(3): 678–686, 2002.

Emotiv. Emotiv epoc headset. `http://emotiv.com/`.

M. Fatourechi, A. Bashashati, R.K. Ward, G.E. Birch, et al. Emg and eog artifacts in brain computer interface systems: A survey. *Clinical neurophysiology*, 118 (3):480–494, 2007.

B. Graimann, B. Allison, and G. Pfurtscheller. *Brain-Computer Interfaces: Revolutionizing Human-Computer Interaction*. Springer Verlag, 2011.

M. Grierson. Composing with brainwaves: minimal trial p300b recognition as an indication of subjective preference for the control of a musical instrument. In *Proceedings of International Cryogenic Materials Conference (ICMC08)*, 2008.

M.A. Hall. *Correlation-based feature selection for machine learning*. PhD thesis, The University of Waikato, 1999.

AlZubi Hamzah S. Preprocessing and feature extraction for asynchronous multi-class noninvasive brain computer interface based on eeg signal. Master's thesis, Masdar Institute of Science and Technology, 2011.

U. Hasson, O. Landesman, B. Knappmeyer, I. Vallines, N. Rubin, and D.J. Heeger. Neurocinematics: The neuroscience of film. *Projections*, 2(1):1–26, 2008.

HealthPage.org, Jun 2010.

E. Hennighausen, M. Heil, and F. Rösler. A correction method for dc drift artifacts. *Electroencephalography and Clinical Neurophysiology*, 86(3):199–204, 1993.

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

S.M. Hosni, M.E. Gadallah, S.F. Bahgat, and M.S. AbdelWahab. Classification of eeg signals using different feature extraction techniques for mental-task bci. In *Computer Engineering & Systems, 2007. ICCES'07. International Conference on*, pages 220–226. IEEE, 2007.

Brain Master Technology Inc., 2001.

S. Kanoh, K. Miyamoto, and T. Yoshinobu. A p300-based bci system for controlling computer cursor movement. In *Conference proceedings:... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Conference*, volume 2011, page 6405, 2011.

M. Kaur, P. Ahmed, and M.Q. Rafiq. Technology development for unblessed people using bci: A survey. *International Journal of Computer Applications*, 40 (1), 2012.

Z.A. Keirn. Alternative modes of communication between man and machine. *Master's thesis, Purdue University*, pages 35–56, 1988.

Z.A. Keirn and J.I. Aunon. A new mode of communication between man and his surroundings. *Biomedical Engineering, IEEE Transactions on*, 37(12): 1209–1214, 1990.

I. Koprinska. Feature selection for brain-computer interfaces. *New Frontiers in Applied Data Mining*, pages 106–117, 2010.

D.J. Krusienski, E.W. Sellers, F. Cabestaing, S. Bayoudh, D.J. McFarland, T.M. Vaughan, and J.R. Wolpaw. A comparison of classification techniques for the p300 speller. *Journal of neural engineering*, 3:299, 2006.

I. Kuzovkin. Pattern recognition for non-invasive eeg-based bci.

J.C. Lee and D.S. Tan. Using a low-cost electroencephalograph for task classification in hci research. In *Proceedings of the 19th annual ACM symposium on User interface software and technology*, pages 81–90. ACM, 2006.

J.R. Millán, M. Franzé, J. Mouriño, F. Cincotti, and F. Babiloni. Relevant eeg features for the classification of spontaneous motor-related tasks. *Biological cybernetics*, 86(2):89–95, 2002.

J. Mostow and J. Beck. When the rubber meets the road: Lessons from the in-school adventures of an automated reading tutor that listens. *Scale-Up in Education*, 2:183–200, 2007.

J. Mostow, K. Chang, and J. Nelson. Toward exploiting eeg input in a reading tutor. In *Artificial Intelligence in Education*, pages 230–237. Springer, 2011.

NeuroSky. About us:brainwave technology. `http://neurosky.com/AboutUs/BrainwaveTechnology.aspx`, a.

NeuroSky. Academics:university. `http://www.neurosky.com/Academics/University.aspx`, b.

NeuroSky. Brain wave signal (eeg) of neurosky,inc. White paper, December 2009. URL `http://neurosky.com/Documents/Document.pdf?`

DocumentID=77eee738-c25c-4d63-b278-1035cfa1de92. Available online (22 pages).

Junior Tennis Association of Queensland Inc. *Understanding the Science*, 2004 (accessed October 19, 2012). URL http://www.jtaqld.com/jtahealthscience.htm.

M. Othman, A. Wahab, and R. Khosrowabadi. Mfcc for robust emotion detection using eeg. In *Communications (MICC), 2009 IEEE 9th Malaysia International Conference on*, pages 98–101. IEEE, 2009.

R. Palaniappan. Brain computer interface design using band powers extracted during mental tasks. In *Neural Engineering, 2005. Conference Proceedings. 2nd International IEEE EMBS Conference on*, pages 321–324. IEEE, 2005.

M.B. Perkusich, T.S. Rached, and A. Perkusich. Thinkcontacts: Use your mind to dial your phone. In *Consumer Electronics (ICCE), 2011 IEEE International Conference on*, pages 105 –106, jan. 2011. doi: 10.1109/ICCE.2011.5722486.

M. Petersen, C. Stahlhut, A. Stopczynski, J. Larsen, and L. Hansen. Smartphones get emotional: mind reading images and reconstructing the neural sources. *Affective Computing and Intelligent Interaction*, pages 578–587, 2011.

J. Rong, D. Dou, G. Frishkoff, D. Tucker, R. Frank, and A. Malony. A semi-automatic framework for mining erp patterns. In *Advanced Information Networking and Applications Workshops, 2007, AINAW'07. 21st International Conference on*, volume 1, pages 329–334. IEEE, 2007.

J. Saab, B. Battes, and M. Grosse-Wentrup. Simultaneous eeg recordings with dry and wet electrodes in motor-imagery. 2012.

A.S. Shirazi, M. Funk, F. Pfleiderer, H. Glück, and A. Schmidt. Mediabrain: Annotating videos based on brain-computer interaction.

R.C. Smith. *Electroencephalograph based brain computer interfaces*. PhD thesis, University College Dublin, 2004.

R. Sobolewski, R.B. Reilly, S. Finnigan, P. Dockree, K. O'Sullivan, and I.H. Robertson. Monitoring of cognitive processes in older persons. In *Neural Engineering, 2009. NER'09. 4th International IEEE/EMBS Conference on*, pages 132–135. IEEE, 2009.

S.S. Stevens, J. Volkmann, and EB Newman. A scale for the measurement of the psychological magnitude pitch. *The Journal of the Acoustical Society of America*, 8(3):185–190, 1937.

L. Stewart, K. Von Kriegstein, J.D. Warren, and T.D. Griffiths. Music and the brain: disorders of musical listening. *Brain*, 129(10):2533–2553, 2006.

Robert Stufflebeam. Neurons, synapses, action potentials, and neurotransmission. `http://www.mind.ilstu.edu/curriculum/neurons_intro/neurons_intro.php`, Jan 2002.

A. Subasi and E. Erçelebi. Classification of eeg signals using neural network and logistic regression. *Computer Methods and Programs in Biomedicine*, 78(2): 87–99, 2005.

Zayin Technologies. *Emotiv EPOC neuroheadset review: The telepathy device you were looking for (or not)*, 2012 (accessed October 19, 2012). URL `http://www.zayintech.com`.

A. Temko, G. Boylan, W. Marnane, and G. Lightbody. Speech recognition features for eeg signal description in detection of neonatal seizures. In *Engineering in Medicine and Biology Society (EMBC), 2010 Annual International Conference of the IEEE*, pages 3281–3284. IEEE, 2010.

A. Vallabhaneni, T. Wang, and B. He. Braincomputer interface. *Neural Engineering*, pages 85–121, 2005.

J. van Erp, F. Lotte, and M. Tangermann. Brain-computer interfaces for non-medical applications: How to move forward. *Computer*, 2012.

I.H. Witten, E. Frank, and M.A. Hall. *Data Mining: Practical Machine Learning Tools and Techniques: Practical Machine Learning Tools and Techniques*. Morgan Kaufmann, 2011.

S. Wren. The brain and reading. topics in early reading coherence. 2001.

Y. Yang and J.O. Pedersen. A comparative study on feature selection in text categorization. In *MACHINE LEARNING-INTERNATIONAL WORKSHOP THEN CONFERENCE-*, pages 412–420. MORGAN KAUFMANN PUBLISHERS, INC., 1997.

Myeung-Sook Yoh, Joonho Kwon, and Sunghoon Kim. Neurowander: a bci game in the form of interactive fairy tale. In *Proceedings of the 12th ACM* *international conference adjunct papers on Ubiquitous computing*, Ubicomp '10 Adjunct, pages 389–390, New York, NY, USA, 2010. ACM. ISBN 978-1-4503-0283-8. doi: 10.1145/1864431.1864450. URL `http://doi.acm.org/10.1145/1864431.1864450`.