

https://dspace.org/

Electroencephalography brain computer interface using an asynchronous protocol

Khoza, Phumlani Rueben Nhlanganiso

2016

http://hdl.handle.net/10539/21648

Khoza, Phumlani Rueben Nhlanganiso (2016) Electroencephalography brain computer interface using an asynchronous protocol, University of Witwatersrand, Johannesburg, <http://wiredspace.wits.ac.za/handle/10539/21648>

Downloaded from a DSpace repository. https://dspace.org/



Electroencephalography Brain Computer Interface Using An Asynchronous Protocol

A dissertation submitted to the Faculty of Science,
University of the Witwatersrand, in fulfillment of the
requirements for the degree of Master of Science.

Phumlani R. Nhlanganiso Khoza

Supervised by: Benjamin Rosman, Hima Vadapalli, and Pravesh Ranchod.

School of Computer Science and Applied Mathematics
The University of the Witwatersrand
Johannesburg

October 31, 2016

Declaration

I declare that this project is my own, unaided work. It is being submitted as partial fulfilment of the Degree of Master of Science at the University of the Witwatersrand, Johannesburg. It has not been submitted before for any degree or examination in any other University.



Phumlani R. Nhlanganiso Khoza
October 31, 2016

Abstract

Brain Computer Interface (BCI) technology is a promising new channel for communication between humans and computers, and consequently other humans. This technology has the potential to form the basis for a paradigm shift in communication for people with disabilities or neuro-degenerative ailments. The objective of this work is to create an asynchronous BCI that is based on a commercial-grade electroencephalography (EEG) sensor. The BCI is intended to allow a user of possibly low income means to issue control signals to a computer by using modulated cortical activation patterns as a control signal. The user achieves this modulation by performing a mental task such as imagining waving the left arm until the computer performs the action intended by the user. In our work, we make use of the Emotiv EPOC headset to perform the EEG measurements. We validate our models by assessing their performance when the experimental data is collected using clinical-grade EEG technology. We make use of a publicly available data-set in the validation phase.

We apply signal processing concepts to extract the power spectrum of each electrode from the EEG time-series data. In particular, we make use of the fast Fourier transform (FFT). Specific bands in the power spectra are used to construct a vector that represents an abstract state the brain is in at that particular moment. The selected bands are motivated by insights from neuroscience. The state vector is used in conjunction with a model that performs classification. The exact purpose of the model is to associate the input data with an abstract classification result which can then used to select the appropriate set of instructions to be executed by the computer. In our work, we make use of probabilistic graphical models to perform this association.

The performance of two probabilistic graphical models is evaluated in this work. As a preliminary step, we perform classification on pre-segmented data and we assess the performance of the hidden conditional random fields (HCRF) model. The pre-segmented data has a trial structure such that each data file contains the power spectra measurements associated with only one mental task. The objective of the assessment is to determine how well the HCRF models the spatio-spectral and temporal relationships in the EEG data when mental tasks are performed in the aforementioned manner. In other words, the HCRF is to model the internal dynamics of the data corresponding to the mental task. The performance of the HCRF is assessed over three and four classes. We find that the HCRF can model the internal structure of the data corresponding to different mental tasks.

As the final step, we perform classification on continuous data that is not segmented and assess the performance of the latent dynamic conditional random fields (LDCRF). The LDCRF is used to perform sequence segmentation and labeling at each time-step so as to allow the program to determine which action should be taken at that moment. The sequence segmentation and labeling is the primary capability that we require in order to facilitate an asynchronous BCI protocol. The continuous data has a trial structure such that each data file contains the power spectra measurements associated with three different mental tasks. The mental tasks are randomly selected at 15 second intervals. The objective of the assessment is to determine how well the LDCRF models the spatio-spectral and temporal relationships in the EEG data, both within each mental task and in the transitions between mental tasks. The performance of the LDCRF is assessed over three classes for both the publicly available data and the data we obtained using the Emotiv EPOC headset. We find that the LDCRF produces a true positive classification rate of 82.31% averaged over three subjects, on the validation data which is in the

publicly available data. On the data collected using the Emotiv EPOC, we find that the LDCRF produces a true positive classification rate of 42.55% averaged over two subjects.

In the two assessments involving the LDCRF, the random classification strategy would produce a true positive classification rate of 33.34%. It is thus clear that our classification strategy provides above random performance on the two groups of data-sets. We conclude that our results indicate that creating low-cost EEG based BCI technology holds potential for future development. However, as discussed in the final chapter, further work on both the software and low-cost hardware aspects is required in order to improve the performance of the technology as it relates to the low-cost context.

Contents

1 Introduction **10**
1 Problem Definition . . . 10
1.1 Importance of the Problem . . . 10
2 Proposed Solution . . . 11
2.1 Technical Prerequisites . . . 11
2.2 The Contribution of this Work . . . 12
3 Conclusion . . . 13
4 Organization of the Dissertation . . . 13

2 Theoretical Foundation **14**
1 Introduction . . . 14
2 Neuroscience Theory . . . 14
2.1 Brain Anatomy and Physiology . . . 15
2.1.1 Anatomy of the Neuron . . . 17
2.1.2 Electrophysiology of the Neuron . . . 18
2.2 Computational Neuroscience . . . 19
2.2.1 Action Potential Modeling . . . 22
2.2.2 Pragmatic Considerations . . . 23
2.3 Electroencephalography . . . 24
2.3.1 Overview of EEG . . . 25
2.3.2 Manifestation of EEG . . . 26
2.3.3 EEG Measurement Standardization . . . 28
2.3.4 EEG Interpretation . . . 28
3 Signal Processing Theory . . . 30
3.1 Feature Vectors . . . 31
3.2 Frequency Decomposition . . . 32
3.2.1 Fourier Transform . . . 32
3.2.2 Fast Fourier Transform . . . 33
3.3 Signal Sampling . . . 33
3.3.1 Nyquist-Shannon Theorem . . . 34
3.3.2 Spectrogram Representation . . . 34
4 Machine Learning Theory . . . 36
4.1 Function Approximation and Classification Techniques . . . 36
4.1.1 Function Approximation Techniques . . . 37
4.1.2 Classification Techniques . . . 38
4.2 Motivation for Machine Learning Application . . . 40
4.3 Probabilistic Graphical Models . . . 40
4.3.1 Model State-Variables Specification . . . 41

4.3.2 Generative versus Discriminative Models . . . 42
4.3.3 Model Representation and Properties . . . 43
4.3.4 Model Parameter Estimation . . . 43
4.3.5 Model Querying and Inference . . . 44
5 Discussion and Conclusion . . . 45

3 Related Research **46**
1 Introduction . . . 46
2 EEG Classification . . . 47
2.1 Synchronous EEG Classification Protocol . . . 47
2.2 Asynchronous EEG Classification Protocol . . . 50
3 Application Domains . . . 51
3.1 Text Input . . . 52
3.2 Rehabilitation and Robot Control . . . 54
4 Discussion and Conclusion . . . 56

4 Research Framework Overview **58**
1 Introduction . . . 58
2 Research Motivation . . . 58
3 Research Hypothesis . . . 59
4 Research Methodology . . . 60
4.1 System Design and Development . . . 61
4.1.1 System Design Considerations . . . 61
4.1.2 System Components . . . 62
4.2 Data Collection . . . 68
4.3 System Performance Evaluation . . . 68
4.3.1 Command Inference . . . 69
4.3.2 Classification Performance Metric . . . 70
4.3.3 Control Performance Metric . . . 71
5 Experimental Equipment . . . 72
5.1 Experimental Equipment Limitations . . . 72
6 Discussion and Conclusion . . . 74

5 Investigation Procedure **75**
1 Introduction . . . 75
2 Data Collection Methodologies . . . 75
2.1 Emotiv Headset Collected Data . . . 75
2.1.1 Phase 1: Training without control feedback . . . 76
2.1.2 Phase 2: Discontinuous task execution without control feedback 76
2.1.3 Phase 3: Continuous task execution without control feedback . . 77
2.2 BCI Competition Sourced Data . . . 78
2.2.1 Continuous task execution without control feedback . . . 78
3 Machine Learning Models . . . 79
3.1 Hidden Conditional Random Fields . . . 80
3.2 Latent Dynamic Conditional Random Fields . . . 82
4 Conclusion . . . 84

6 Results and Analysis **86**
1 Introduction . . . 86
2 Visualization Tools . . . 86

2.1 EEG Time Series . . . 86
2.2 Temporally Sequenced Spatio-Spectral Topographical Maps . . . 87
2.3 Temporal Classification Performance Plot . . . 87
2.4 Segmentation and Labeling Performance Plot . . . 88
2.5 Confusion Matrix . . . 88
3 Experiment Outcomes . . . 89
3.1 Data Irregularities . . . 89
3.2 Classification Results . . . 90
3.3 Sequence Segmentation and Labeling Results . . . 91
4 Analysis . . . 96
4.1 Neuroscience Correspondence . . . 96
4.1.1 Hardware Limitations . . . 97
4.1.2 Inter-day Variations . . . 98
4.1.3 Inter-subject Variations . . . 98
4.1.4 Control Feedback Effects . . . 100
4.2 Model Performance Analysis . . . 100
5 Discussion and Conclusion . . . 102

7 Research Conclusion **104**
1 Research Findings Discussion . . . 105
1.1 Key Research Findings . . . 105
1.2 Unanticipated Outcomes . . . 106
1.3 Conclusion Concerning Performance . . . 106
2 Proposed Future Work . . . 107
2.1 Realtime Cortical Activation Visualization . . . 107
2.2 Automated Mental Task Selection . . . 107
2.3 Control Feedback Effect Assessment . . . 108
2.4 Minimal Montage and Channel Selection . . . 109
2.5 Unsupervised Feature Learning . . . 109
2.6 Intra-subject Hierarchical Command Inference . . . 110
2.7 Inter-session Intra-Subject Transfer Learning . . . 110
2.8 Inter-subject Transfer Learning . . . 111
3 Conclusion . . . 111

8 References **113**

List of Figures

2.1 The cortical area function-demarcated model of the human brain. The labels indicate the function of the corresponding area. Varying amounts of neural tissue are dedicated to performing functions, as illustrated in the image. 15
2.2 Cortical homunculus model for neural-tissue-to-body-part mapping. Varying amounts of neural tissue within a cortical region are dedicated to a particular body part. The spatial distribution of the tissue has a direct influence on the measurement of EEG activity corresponding to a particular region. 17
2.3 Schematics of biological neurons. Cell bodies that contain the nucleus are depicted using black bulbous dots. The long projections emanating from the cell body represent the axon of the neuron that transmits the action potential of that neuron to other neurons it is connected to. The short projections represent the dendrites of the neuron that receive action potentials from other neurons connected to it. 18
2.4 Measurement of neuron potential difference. Depending on the stimulation induced on a neuron that has synaptic coupling with another neuron, a measurable change in the potential difference is produced between two spatially separated locations in the receiving neuron. Charge is a conserved quantity, so the spatial distribution of the potential difference has to reflect the conservation. In the image, the potential differences are measured with respect to ground at two locations and differences in the plotted profiles illustrates the conservation of charge. 19
2.5 Neuro-transmiter mediated inter-neuron electric action potential transfer. The axon terminal is connected to the neuron that is the source of the action potential. The recipient of the action potential is neuron with the dendritic extension. . . . 20
2.6 Schematic depicting long-range information processing in the nervous system. Not all nerves are directly connected to the brain with a single long-range connection. The spinal cord provides the relay mechanism that transmits action potentials to and from the brain. Concerning motor function, the connection that emanates from the brain terminates at a corresponding muscle. Concerning sensory function, the connections that emanate from sensory units such as tactile sensors on the skin terminate in the brain. 21
2.7 Schematic of the Hodgkin-Huxley model for action single neuron action potential. The model is based on biophysical modeling with certain physical attributes of a neuron represented by an electrical circuit that performs an analogous role in an electrical circuit. 22

2.8 The samples of main types of artifacts in human EEG. 1 - Electrooculographic artifact caused by the excitation of eyeball's muscles. 2 - Electrode's artifact caused by bad contact between P3 electrode and skin. 3 - Swallowing artifact. 4 - Common reference electrode's artifact caused by bad contact between reference electrode and skin. 26
2.9 EEG signal decomposition into primary EEG bands. The frequency decompositions of the EEG signal represent the current categorization scheme. All except the gamma waves have been attributed to specific brain activiy. 27
2.10 General montage template using the 10-20 electrode placement system. The topographical layout of the nodes is intended to provide full coverage on the scalp. 29
2.11 Illustration of the proportion marked 10-20 electrode placement system. The relative demarcation is designed such that differences in head sizes have a minimal influence on inter-subject comparative measurements. 30
2.12 Illustration of the effect of the time-to-frequency domain transformation of a signal. 31
2.13 Illustration of the decomposition of a signal into its constituent signals. Points of similarity or dissimilarity are easily identifiable in the frequency domain. . . . 32
2.14 Discrete sampling of a continuous signal, the sampling rate is usually a fixed value during sampling. The sampling rate determines how many samples are measured within a given time interval, the higher the sampling rate, the better the signal is approximated. 35
2.15 Graphical representation of a signal spectrogram. Along the vertical axis are frequency values, along the horizontal axis is time, and the colour bar represents the magnitude of the frequency component. With the three pieces of information it is possible to visualize the temporal evolution of the frequency components. . . 36
2.16 Function approximation with noisy data. The vertical lines on the data points represent error or variation in the measurement. 37
2.17 Two class animal classification with a decision boundary between the sets. 39
2.18 Graphical illustration of the structure of probabilistic graphical models. 42
2.19 Bayesian and Markov probabilistic graphical models. 43
2.20 Probabilistic graphical model and associated disjoint conditional probability distributions. 44

3.1 Orthosis position actuation. 55

4.1 System design schematic with signal acquisition for on-line mode data-flow. . . . 62
4.2 System design schematic for off-line mode data-flow. 63
4.3 Device status monitor GUI depicting device status information. 64
4.4 Classifier GUI depicting model classification and command inference results. . . . 65
4.5 Text input GUI with predictive text functionality. 67
4.6 Temporal classification performance visualization tool. 68
4.7 Snap-shot of temporally-sequenced spatio-spectral topographic maps. The row labels in the image represent mental tasks, and the column labels represent frequency bins that are spaced in 2Hz intervals from 2Hz to 26Hz. 69
4.8 Control model as defined on the text input interface. 70
4.9 Emotiv EPOC EEG headset. 73

5.1 The evaluated discriminative probabilistic graphical models have blue labels. The CRF is shown for comparison purposes, and both the HCRF and LDCRF have been used in the experiments. 80

6.1 Emotiv EEG time series visualization program snap-shot. 87
6.2 Confusion matrix results for sub2 corresponding to row two in Table 6.1. 89
6.3 Effect of sampling rate and contact quality variations on the EEG measurements. Variations of this kind can be three orders of magnitude higher than the normal EEG activity, and negatively influence the classification performance of the models. 90
6.4 Confusion matrix results obtained using sub1 test data. The LDCRF doesn't perform well on the word generation mental task. The motor-imagery tasks appear to be discriminable, and the structure they encode has been learnt by the LDCRF. 92
6.5 Temporal classification performance on Emotiv data for sub1. The plot represents the results obtained from the test data. The results obtained from sub2 are similar in form. The misalignment in the inferred and actual class labels indicated misclassification. 93
6.6 Confusion matrix results obtained using subj2 test data. The results indicate that both the first and third mental tasks are inferred with 100% accuracy when the actual class label coincides with any of those classes. Using Figure 6.7, it is clear that the classification error observed with the second mental task results is due only to the first two clusters in the data, beyond that the LDCRF produces error-free classification performance. 94
6.7 Temporal classification performance on BCI Competition data for subj2. The plot represents the results obtained from the test data. The misalignment in the inferred and actual class labels indicated misclassification 94
6.8 Confusion matrix results obtained using subj3 test data. The results indicate that the third mental task is inferred with 100% accuracy when the actual class label coincides with that class. A large proportion of the misclassification is attributable to the second mental task, in which the structure appears to not being learnt well by the LDCRF. 95
6.9 Temporal classification performance on BCI Competition data for subj3. The plot represents the results obtained from the test data. The misalignment in the inferred and actual class labels indicated misclassification 96
6.10 Temporal layout of classification results. The figure indicates that the model consistently infers the same class label for almost all data samples. It should also be noted that only two of the total three classes are inferred throughout the whole process. 99
6.11 Confusion matrix results obtained from the cross-subject validation. The results indicate that the model doesn't infer one of the classes at all, either correctly or incorrectly. The 100% true positive accuracy noted in one of the classes could be spurious. 99
6.12 In some samples there were similarities in the measurements, as highlighted with ovals in the image. The samples were collected in one session and from one subject. Each mental task was performed independently of other mental tasks, consequently they do not share the same timestamp. It is a matter of coincidence that they appear in the same frame on the TeSSTMaps visualization tool. Thus the degree of similarity depicted here is a gross underestimate of the true amount of similarity between the sequences. All the tasks were mental, and that means that the motor tasks were imagined. 101

List of Tables

5.1 Emotiv Data: Frequency bands and channels used for HCRF and LDCRF state representation. 77
5.2 BCI Competition Data: Frequency bands and channels used for HCRF and LDCRF state representation. 79

6.1 HCRF classification using Emotiv data. Each row represents an indepedent trial. 91
6.2 LDCRF classification using Emotiv data. 92
6.3 LDCRF classification using BCI Competition data. 93

Acknowledgements

First, I would like to thank Shun and Pravesh for grooming me over the past couple of years, and for supporting me in pursuing this madness, which on occasion results in wild bursts of energy and lethargy. I'd also like to thank Benji for going the extra mile in supervising this project, and in particular for reviewing numerous late-night-typed documents that somehow managed to carry vast amounts of grammatical and spelling errors. I'd also like to thank Hima for providing additional supervision in pursuing this experiment. Lastly, I would like to thank Adam Pantanowitz and Professor David Rubin from the School of Biomedical Engineering for their generous help. Without their Emotiv EPOC headset, this project would have ground to a halt. I hope this is the starting point for productive future collaboration. From whom I inherited blood:

```
..ripples tell tales from distant ages,
and winds blow across forgotten lands,
so flows the great current,
and in the darkest night it rages.

Grains define isiqu,
and they never grow weak,
for the earth remembers that which arose from it.

The first ray of dawn illuminates,
and the wise stand at the edge of the hills,
but only to witness the gushing rivers,
and though them life flows..

                                                -Nhlangothi
```

Abbreviations

BCI: Brain Computer Interface
CRF: Conditional Random Fields
CSP: Common Spatial Pattern
DFT: Discrete Fourier Transform
EEG: Electroencephalography
EMG: Electromyography
FFT: Fast Fourier Transform
fMRI: functional Magnetic Resonance Imaging
GUI: Graphical User Interface
HCRF: Hidden Conditional Random Fields
HMM: Hidden Markov Model
LDCRF: Latent Dynamic Conditional Random Fields
LDA: Linear Discriminant Analysis
kNN: k Nearest Neighbour
MDA: Minimum Distance Analysis
MRI: Magnetic Resonance Imaging
SDK: Software Development Kit
SMR: Sensory-Motor Rhythms
SNR: Signal-to-Noise Ratio

Chapter 1

Introduction

1 Problem Definition

Life is filled with unpredictable events that can lead to loss of motor control either through tissue damage which is induced by injuries, degeneration of a tissue group, or other ailments. Amongst injuries, one of the most common causes are motor vehicle accidents. There are various causes for degeneration of neural tissue, some common examples include: Parkinson's disease, Huntington's disease, and amyotrophic lateral sclerosis [Purves et al., 2004]. Other causes include ailments that influence motor control, but do not necessarily manifest as neural tissue degeneration. A prime example of such an ailment is stroke [Purves et al., 2004]. Loss of motor control leads to a considerable degradation of the quality of life. This manifests itself under various guises, among which we have a considerable or total loss of capacity to be able to tend to one's basic needs such as being able to feed oneself, or to communicate with other people.

The problem under consideration in this work concerns the severing of the neuro-muscular connection, which in some cases this amounts to dismemberment of an associated appendage. Although it is possible to continue living, the experience does carry important implications concerning the quality of life. One of the most common problems associated with motor-control loss is the disconnect from society that the afflicted person feels. This disconnect often leads to depression for the afflicted person [Taylor 2006].

1.1 Importance of the Problem

The effects of the aforementioned conditions go beyond the person who is afflicted, since the condition becomes a problem within the family. In other cases it becomes a problem to society at large. Within the family unit, a care-giver is required to attend to the needs of the afflicted person. Since there is considerable concentration of wealth in South Africa (Gini coefficient of 0.65 as measured in 2011 [Statssa 2011]) as well as in many other countries in the world, this leads to the creation of an additional non-negligible burden on the family. This usually calls for greater resource requirements while the means of obtaining those resources are decreased, as in the case of an economically active person who becomes disabled.

Within a larger societal context, the effects are compounded by the number of people who then have to be allocated state disability grants (1 127 285 people as measured in South Africa

in August 2014 [Sassa 2014]). Although disability grants are based on prudent government policy, the financial resources that are required are substantial (R17 450 371 800 per annum as measured in South Africa in August 2014 [Sassa 2014]). These resources could be used for further advancement of society if people with disabilities could be given the opportunity to contribute to society, which in itself is a fulfilling experience, as it forms a part of purposeful living.

2 Proposed Solution

This work is aimed at providing an alternative channel for communication that does not depend on neuro-muscular connections, although there are advances in tissue engineering which could prove to be the ultimate solution for some of these problems by recreating neuro-muscular connections [Purves et al., 2004]. Applying those advances to solve problems at large-scale at present is not possible for technical and financial reasons [Purves et al., 2004]. In contrast we propose a low-cost solution that can be used in everyday situations by the average person.[1]

We propose a low-cost and non-invasive brain-computer interface (BCI).[2] Non-invasive BCI technology does not require the insertion of measurement devices into body tissue. In the context of EEG, the measurements are obtained from placing electrodes on the scalp of the subject. The application of the BCI could for instance be to allow a user to perform text input on a computer by performing certain mental tasks. The tasks could include imagining moving one's left arm or performing arithmetic operations. Different mental tasks would be mapped to different instructions that a computer would execute. This would result in text being typed out without the use of standard peripheral devices such as a keyboard or mouse.

2.1 Technical Prerequisites

In order to achieve the above stated objective, various technical challenges have to be addressed. The first challenge pertains to what it is that has to be measured in order to make the solution work. The answer to this is provided by neuroscience theory. As discussed in Section 2 of Chapter 2, the brain exhibits non-uniform anatomical structure and neurophysiological behaviour across various regions. Of particular importance is the electrophysiology of the brain, as this provides the electrical signal that can be measured and used as a correlate signal conveying information about the underlying activity in the brain. This signal is measured as electric potential differences between electrodes placed at different locations on the scalp. The essential insight here is that the time series of the electric potential differences encodes structure concerning which cortical regions are activated, and in what temporal order. The spatio-temporal dynamics in this signal are what is required in order to proceed to the next step leading to the

[1] The price of a clinical-grade EEG machines go into the region of ZAR100 000. In contrast, the Emotiv EPOC headset we use in this work costs ZAR10 000, although some of the newer models from the same manufacturer cost less than the Emotiv EPOC.

[2] Non-invasive BCIs such as those based on electroencephalography are easier to use since they do not require direct contact with internal tissue. This stands in contrast to invasive BCIs such as those based on electrocorticography (ECG). Although invasive BCIs have higher spatial resolution and lower noise in the measured signal, the risk they pose with regards to infection or inflammation responses rules them out for this work.

proposed solution.

The second challenge pertains to how the measured signal should be transformed in order to reveal the structure encoded therein. Theoretical insights from both neuroscience and signal processing motivate that the transformation should reveal frequency domain structure, as opposed to purely temporal domain structure (see Section 3 of Chapter 2). The resulting frequency distribution information also exhibits complex spatio-temporal dynamics. Thus in order to derive meaning from this information, a model that can represent the dynamics is required. Unfortunately, one of the major limitations of current neuroscience theory is that a general theory of how the brain works is still not available, so interpreting the data becomes a substantial problem.

The third challenge pertains to finding an appropriate model that addresses the limitation of neuroscience theory outlined above. This limitation makes it challenging to derive a mathematical model that can represent the dynamics of the observed data, since there is currently no sound theoretical justification for the observed data. In order to circumvent the problem, we recast the mathematical modeling problem into a machine learning problem (see Section 4 of Chapter 2). When the problem is expressed within the machine learning paradigm, the main objective then becomes automatically learning the association between input data and output data. This means that the input data, which is frequency distribution information, is associated with an abstract class, and this class represents a label of some predefined mental activity. It is important to note that the reformulation of the initial problem is not intended to provide epistemological insight pertaining to how the brain works, but instead is for providing a mapping between the input data and an abstract classification label. Once the mapping has been learnt for each subject, the labels produced by the model can then be used to invoke the execution of a predefined set of instructions on a computer.

2.2 The Contribution of this Work

The main contribution of this work concerns investigating the feasibility of creating a low-cost BCI, using the low-cost Emotiv EEG sensing technology (see Section 5 of Chapter 4). The chosen hardware is not useful in clinical applications due to the low resolution and the quality of the signals it provides, but it may be used for BCI applications [Emotiv 2014]. Our particular focus is on the machine learning models that can be used to facilitate acceptable performance levels of the proposed BCI. We perform classification and sequence labeling using probabilistic graphical models. A broad outline of these models is provided in Section 4.3 of Chapter 2, and a detailed description is provided in Section 3 of Chapter 5. The probabilistic graphical models that we use have been shown to provide superior performance in various application domains, including EEG related domains [Quattoni et al. 2007, Morency et al. 2007, Saa and Çetin 2012, Saa and Çetin 2013]. The models also have various theoretical properties, such as learning temporal dynamics and circumventing unjustifiable statistical independence assumptions, that make them especially suited to this application domain. The use of low-cost EEG hardware and probabilistic graphical models in the context of BCI technology is our main contribution.

3 Conclusion

This chapter presents the argument that the problem of motor-function loss leads to the degeneration of the quality of life for an afflicted individual, and that the repercussions are also significant at a societal level. The socio-economic implications carried by this problem are substantial. In the second section, the proposed solution that may prove to be applicable to a wide range of people in the economic spectrum was introduced.

There are various technical challenges that have to be addressed in this regard, and some of the solutions to these challenges are found within the theoretical constructs of neuroscience, signal processing, and machine learning. The essential insight concerning the proposed solution is that an electrical correlate signal can be measured on the scalp of the person using the BCI, and the signal can be meaningfully interpreted such that the intentions of the user are translated into the execution of computer instructions. All the required theoretical constructs are developed and discussed in the subsequent chapters.

4 Organization of the Dissertation

In the chapters that follow, the theoretical foundation of the proposed solution is first outlined in Chapter 2. This is achieved by discussing the required key theoretical results from neuroscience, signal processing, and machine learning. In Chapter 3, various studies relating to the subject of this work are reviewed. In that chapter, we note the dominant use of clinical-grade EEG hardware in BCI technology research field. In this work we take a different approach by seeking to make the technology accessible to people of average means. This is achieved by using low-cost EEG technology.

In Chapter 4, the procedure that was followed to perform the investigations is outlined. It is in that chapter this the research hypothesis and the research methodology are presented. This work has a considerable empirical research component, and so a detailed account of the tools which were developed in order to enable experimentation is presented. Chapter 5 then outlines the experimental protocol that was followed to test the performance of the system. An outline of the data collection methodologies and experiment paradigms is provided. The concluding section provides a technically oriented description of the probabilistic graphical models that we use in this work.

In Chapter 6, the results obtained from the experiments are presented. This presentation begins with an outline of the visualization tools that assist in extracting useful information from the experimental data. The chapter ends with a discussion pertaining to the data and some of the irregularities that were encountered during the experimental sessions. After the results have been presented, we turn to the analysis of the results in Chapter 6, in which the research hypothesis of this work is tested. In that chapter, the correspondence between the data and neuroscience is also investigated. The chapter ends with an analysis of the performance of the machine learning models that were used in this work. Lastly, Chapter 7 provides a discussion of the research findings and possible avenues for future work are outlined.

Chapter 2

Theoretical Foundation

1 Introduction

This chapter discusses the theoretical results that are required to implement the proposed solution. The discussion begins with basic insights from neuroscience that have led to the current model of the brain and how it functions from a computational perspective. It is in this section that the foundation of electroencephalography (EEG) is outlined (see Section 2). Subsequently we introduce some basic concepts from signal processing theory that are required when data with complex structure has to be processed (see Section 3). The discussion concludes with a discussion on machine learning and some of the main results from the field that are used in this work (see Section 4).

2 Neuroscience Theory

Decades of brain research have produced important results concerning both the structure and function of brain regions. Lesion studies have revealed that brain functions seem to be related to very specific brain regions. Although the regions do depend on other regions for input of information, some functions are dominated by certain regions in the sense that if the region is damaged, then the corresponding function is lost [Purves et al., 2004]. This discovery formed the basis of the model that is presented in Figure 2.1. The model represents current knowledge on matters of anatomical definition and physiological processes in the brain.

Although a general theory of brain functions is still not available, the study of neuroscience has undergone substantial transformations and has made considerable progress. Some of the major contributions have revealed deep anatomical information, with one iconic example being the neuron staining technique introduced by Santiago Ramón y Cajal over a century ago. Cajal's technique provided the first insights into the structure of the atomic units of the brain which are the neurons (see Figure 2.3).

With the neuron having been isolated, further research would be carried out which not only unveiled more delicate anatomical definitions, but gave rise to the study of the physiology of the neuron itself. Most of the studies pertaining to neuron physiology involved investigating electrical properties of the neuron. Regularities in electrical response of the neuron would soon lead to what we now call computational neuroscience, as discussed in Section 2.2.

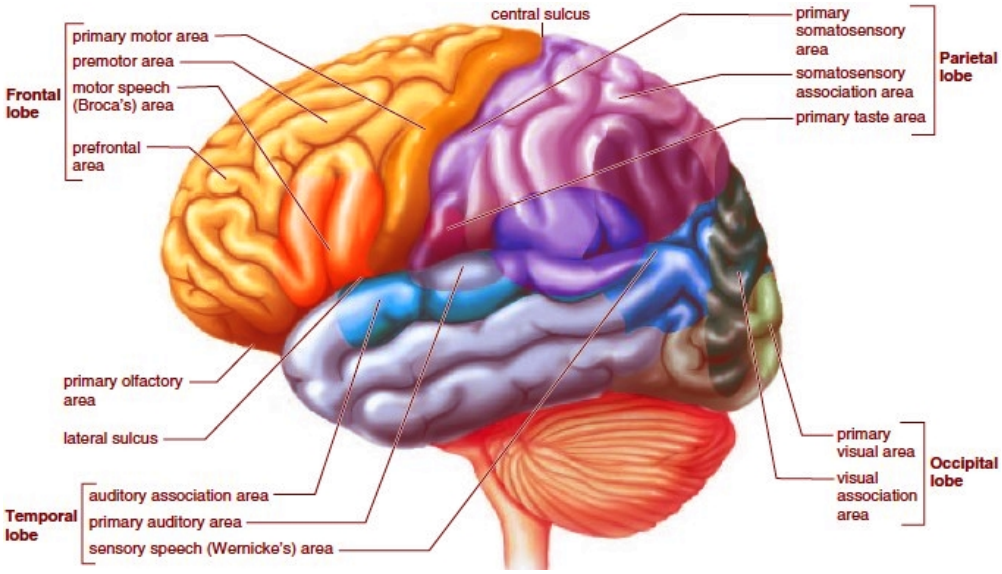


Figure 2.1: The cortical area function-demarcated model of the human brain. The labels indicate the function of the corresponding area. Varying amounts of neural tissue are dedicated to performing functions, as illustrated in the image.

Source: Neurophysiology [KIN 2016].

Progress in related fields such as physics have provided tools that allow for safer and more effective investigation of both anatomical and physiological details. In the context of anatomy, magnetic resonance imaging (MRI) has proven indispensable when financial constraints are not prohibitive. The primary benefit of using MRI is very high spatial resolution. In the context of physiology, functional magnetic resonance imaging (fMRI) proves to be indispensable, but the latency which can vary up to six seconds is not negligible. EEG, which is the sensing modality used in this work, offers very low spatial resolution. However the latency is low, and this makes it ideal for investigating physiological processes that vary in the timescales of the order of hundreds of milliseconds.

In subsequent subsections, a deeper exposition of neuroscience principles is provided. The exposition begins with anatomical and physiological concepts about the brain, in which the structural and functional details are uncovered (see Section 2.1). This has been the dominant view on the subject for over a century. The next subsection concerns the computational neuroscience perspective, in which the brain is viewed as a computational system which performs information processing via physiological processes that are supported by the underlying anatomy (see Section 2.2). The formulation of neuronal activity in terms of a mathematical model is one of the fundamental results from the field. EEG measurement forms the basis of this work, so a detailed account of its underlying principles is important. Thus the exposition ends with a discussion of EEG, in which the source of the phenomenon is revealed and the methodology for measurement standardization is outlined (see Section 2.3).

2.1 Brain Anatomy and Physiology

The study of brain anatomy concerns the structural make-up of the brain. The brain has numerous structures, most of which we exclude from our discussion. Our discussion focuses primarily

on the neocortex. The neocortex is a multi-layer neural tissue that represents the outermost surface of the brain. Only mammals have the neocortex, and it is thought to have appeared recently in evolutionary terms. Current knowledge suggests that the neocortex is responsible for higher-order cognitive capabilities such as language processing, planning, mathematical reasoning, and other capabilities of comparable complexity.

Brain anatomy is important because it helps us better understand the brain both in anatomical definition and physiological functioning. Anatomical definition defines the possibilities for physiological processes. For example, the neocortex has a non-trivial structure. There are multitudes of cortical folds which enable higher neural tissue densities to be packed into the relatively small volume provided by the human skull. The effects of cortical folds go beyond just neural tissue densities, with the measurable consequences being non-negligible. There is research that is aimed at functional localization, an example being by Brett et al. [2002] on the brain warping problem which is one of that fundamental problems in the research area.

The folding reorients patches of the neural tissue and with it, the orientation of electrical properties such as electric dipoles. When EEG measurements are performed, the apparent source of the EEG activity may be very misleading. This happens because the summation of electric dipoles is not independent of the geometry of the dipole configuration. Source localization is one of the difficulties that are impeding the development of EEG based brain imaging technology. The problem is aggravated when inter-subject comparative investigations are performed. This follows from the observation that individual brains exhibit different patterns in the structure of the cortical folds [Brett et al. 2002]. The result being that neural activation from the same cortical areas may have varying apparent source localities, when the orientation of the cortical folds is not taken into account.

In contrast to brain anatomy, brain physiology concerns the study of the physical processes that allow the brain to function properly and they manifest in various forms. The processes are also mediated by various mechanisms which include chemical messengers and electric potentials. In this work, the focus is exclusively on electrically mediated processes, as it is from these processes that the EEG signal arises (see Section 2.3).

Brain physiology is important because it provides the theoretical framework used to characterize the normality of certain processes in the brain. There are various physiologically inspired tests, some of which are based on EEG. For example, EEG tests are used to determine if a patient has epilepsy. Based on the spatial distribution and regularities in the EEG signal, a neurologist is able to make inferences concerning normality.[1] This topic on EEG is developed further in Section 2.3.4. A quantitative study of physiological processes naturally leads to the computational neuroscience perspective, because this shift in perspective involves information processing (see Section 2.2).

An important neurophysiological observation is that varying amounts of neural tissue within the cortex is dedicated to the control of certain parts of the body (see Figure 2.2). Furthermore, neural tissue corresponding to different parts of the body is localized around certain cortical subregions within the motor cortex. The result being that if there are electrodes positioned to measure EEG signals from those regions, then detecting the EEG activity from those high neural

[1] It is worth stating that abnormal EEG activity does not necessarily imply clinically abnormal brain function. The converse argument also holds. Thus the challenge that the neurologist faces is not trivial.

density subcortical regions is easier because of the higher levels of activity.[2] The next subsections describe basic anatomy and electrophysiology of the neuron, which is the fundamental building block of the brain. Section 2.1.2 naturally leads us to the information processing perspective from computational neuroscience (Section 2.2). This perspective provides the philosophical foundation assumed in this work, and marks the end of our discussion of neuroscience.

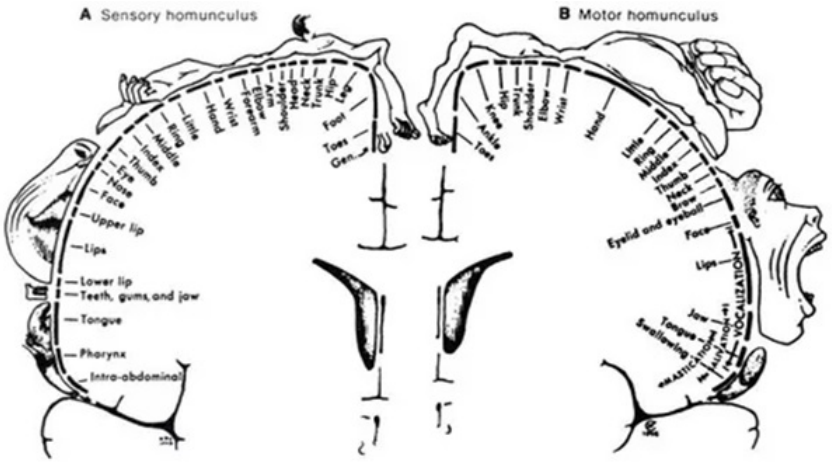


Figure 2.2: Cortical homunculus model for neural-tissue-to-body-part mapping. Varying amounts of neural tissue within a cortical region are dedicated to a particular body part. The spatial distribution of the tissue has a direct influence on the measurement of EEG activity corresponding to a particular region.

Source: The cerebral cortex of man: a clinical study of localization of function [Penfield and Rasmussen 1950].

2.1.1 Anatomy of the Neuron

Different parts of the nervous system have different kinds of neurons, and the neurons have different functions within the nervous system. Anatomy of the neuron is the study concerned with the structural composition of the neuron, and thus provides a static view of the neuron. Three of the anatomical signatures of neurons are depicted in Figure 2.3. Neurons have three primary structures. The bulbous black object represents the cell body, and the spiny projections represent either axons or dendrites. Vital functions that keep the cell alive are performed in the cell body. The cell nucleus is also located in the cell body. Action potentials to be communicated with other neurons propagate via the axon.[3] Action potentials received from other neurons propagate via the dendrites, and are relayed to the cell body of the receiving neuron.

The anatomical properties of the cell are important because they define the possibilities for the electrophysiological processes (see Section 2.1.2). For example, the bi-lipid layer on the neuron provides the electrical insulation that is required for propagating the action potential along

[2]To use music as an analogy, these regions are the band-members who play the loudest.
[3]The action potentials are the electrical potential signals that are communicated between neurons.

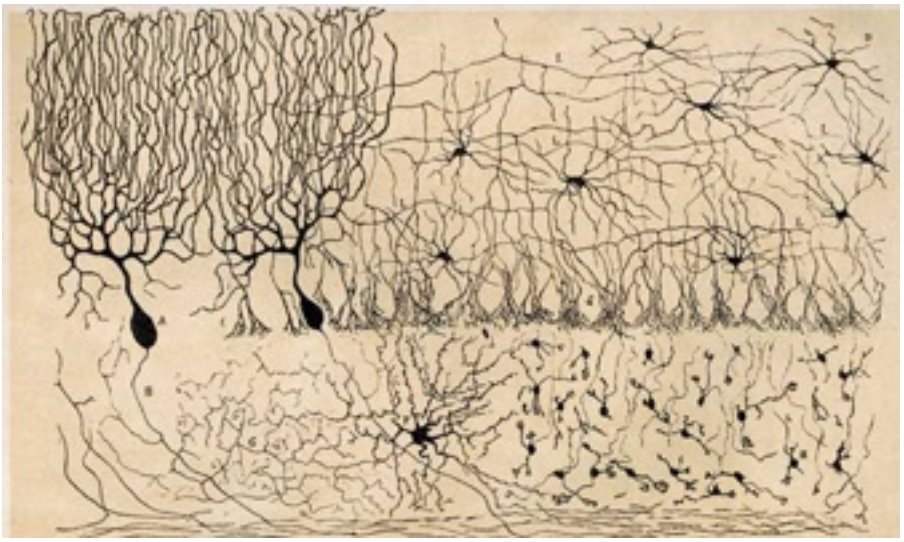


Figure 2.3: Schematics of biological neurons. Cell bodies that contain the nucleus are depicted using black bulbous dots. The long projections emanating from the cell body represent the axon of the neuron that transmits the action potential of that neuron to other neurons it is connected to. The short projections represent the dendrites of the neuron that receive action potentials from other neurons connected to it.

Source: Neuron doctrine [Wikipedia neuron 2015].

the axon. If there are structural weaknesses on the bi-lipid layer, then information transmission is compromised.

2.1.2 Electrophysiology of the Neuron

Electrophysiology of the neuron is the study concerning the electrical properties of the neuron. Electrophysiology provides a dynamic view of the neuron since the electrical properties vary with time, giving rise to the differential equation formulation presented in Equations 2.1 of Section 2.2.2. In this work we focus on the dynamical properties.

What is typically measured is the electric potential difference between two reference points on the neuron. Neurons have intrinsic electrical properties because they maintain potential differences between the inner and outer environments of the cell. The processes involved in maintaining the potential differences are active ion-exchange mechanisms. The ion-exchange mechanisms are mediated by ion-channels which are located on the cell membrane of the neuron.

One of the major advances in the subject came by way of neuron excitation experiments. A classic setup of these is depicted in Figure 2.4. Those experiments suggested that the encoding scheme that the brain uses, which manifests through the discharge of action potentials, is determined by anatomical structures and well-defined physical processes [Dayan and Abbott 2001]. Even though different brains induce variations to the mechanism owing to differences in how each brain has developed, the underlying principle is invariant.

This leads to the view that there is a particular encoding mechanism that the brain uses to process and transmit information. This idea is further developed in Section 2.2. If those information processing mechanisms can be understood, then they can be directly used to encode,

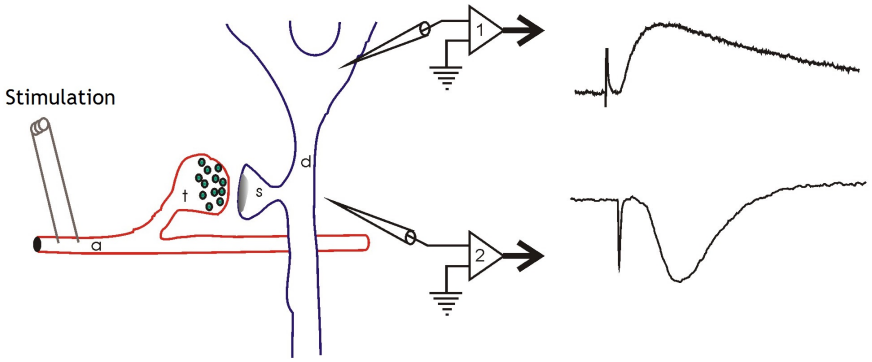


Figure 2.4: Measurement of neuron potential difference. Depending on the stimulation induced on a neuron that has synaptic coupling with another neuron, a measurable change in the potential difference is produced between two spatially separated locations in the receiving neuron. Charge is a conserved quantity, so the spatial distribution of the potential difference has to reflect the conservation. In the image, the potential differences are measured with respect to ground at two locations and differences in the plotted profiles illustrates the conservation of charge.

Source: Electrophysiology [Wikipedia electrophysiology 2016].

decode, or transmit neural information beyond the nervous system. In this work, the objective is to leverage this notion in order to create the bridge between mental tasks and the execution of computer instructions.

Action potentials generated in a neuron are propagated through the neuron's axon. The destinations of the action potential are the synaptic junctions in which the dendrites from other neurons make connections. It is worth noting that the connection is not a trivial physical connection: there are spaces at the synaptic junctions. The transmission of action potentials is mediated by neuro-transmitters which traverse the space between the synaptic junctions of different neurons (see Figure 2.5). Section 2.2 provides a quantitative framework which is used to describe the creation and propagation of the action potentials.

2.2 Computational Neuroscience

Mathematical modeling has been successfully applied in various fields, and has produced resounding success in physics. Circa 1950, a new paradigm drawing from biophysics emerged in neuroscience. The paradigm led to the creation of the Hodgkin-Huxley model which formed the basis for the biophysical modeling of action potentials. Computational neuroscience pertains to the application of quantitative techniques to describe the physical processes in the brain in terms of information processing. Within this paradigm, the primary focus is on information processing and the anatomical structures that make it possible.

Computational neuroscience is important because through the use of quantitative techniques, it provides a framework that allows for the systematic study of information processing in the brain. It also shows that the brain employs definite and unambiguous mechanisms for information processing. The main task in computational neuroscience then is uncovering those

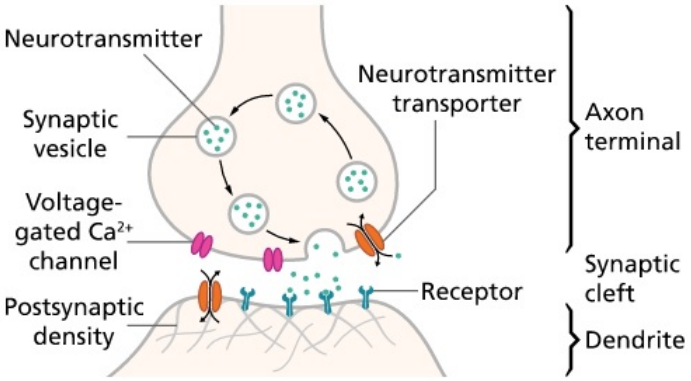


Figure 2.5: Neuro-transmiter mediated inter-neuron electric action potential transfer. The axon terminal is connected to the neuron that is the source of the action potential. The recipient of the action potential is neuron with the dendritic extension.

Source: Action potential [Wikipedia action potential 2016].

mechanisms that underlie information processing.

This philosophical basis is pivotal to this work, because it justifies the search for a general trend in cortical activity that is associated with certain mental tasks. The general idea being the following. Certain mental tasks give rise to increased or decreased neural activity in certain cortical regions. If the neural activity in those cortical regions is largely invariant whenever a particular mental activity is performed, then the observables (i.e. through EEG signals) associated with neural activity would also be largely invariant. Hence, the cortical activity would also exhibit a general trend associated with the performance of particular mental tasks. If the mental tasks invoke neural activity differently, then it follows that in the absence of informational degeneracy, the general trends observed at the cortical level would also be different. Thus the cortical activity trends would be distinguishable from each other.

There are various models which focus on different problems in the field. On one side of the spectrum, we find chemical intracellular processes that regulate the cross-membrane potential differences and regulate firing rates for a single neuron (see Figure 2.4). On the other side of the spectrum, we find inter-neuron processes that communicate information within small regions of neural tissue (see Figure 2.5), and across long-range pathways between cortical regions and sensory or motor extensions of the nervous system (see Figure 2.6).

The primary means by which information is communicated between neurons is through action potentials.[4] The action potentials propagate through neuron axons. In this work, we focus exclusively on aspects related to inter-cortical action potentials, the propagation of those potentials, and the measurable effects that they produce. Below, we discuss the biophysical model which describes the propagation of action potentials.

[4]There are other mechanisms, such as those mediated by chemical messengers like hormones, which are involved in the communication of information. But for practical reasons, which pertain to what we measure in the study, they are not discussed in this work.

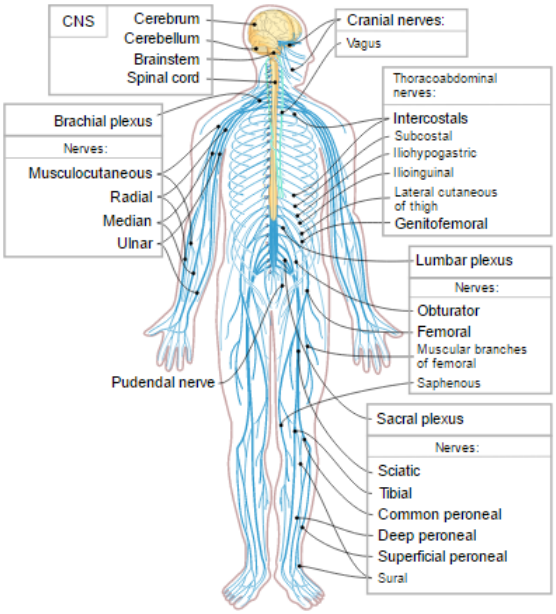


Figure 2.6: Schematic depicting long-range information processing in the nervous system. Not all nerves are directly connected to the brain with a single long-range connection. The spinal cord provides the relay mechanism that transmits action potentials to and from the brain. Concerning motor function, the connection that emanates from the brain terminates at a corresponding muscle. Concerning sensory function, the connections that emanate from sensory units such as tactile sensors on the skin terminate in the brain.

Source: Spinal cord [Wikipedia spinal cord 2016].

2.2.1 Action Potential Modeling

Various efforts have been made in the computational neuroscience field regarding the physical processes that lead to the observed neuronal excitation. The Hodgkin-Huxley model [Hodgkin and Huxley 1952] provides early attempts based on biophysical insights to model the electrical response of a neuron given its internal state and its environmental conditions (see Figure 2.7). In the Hodgkin-Huxley model, the action potential which propagates through the neuron's axon is of central importance. The essence of the idea is that a neuron can be considered as an electrical circuit. The physical attributes of the cell are then modeled as electric circuit components (see Figure 2.7).

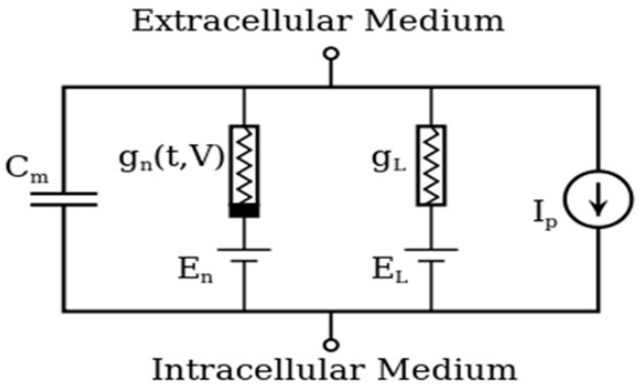


Figure 2.7: Schematic of the Hodgkin-Huxley model for action single neuron action potential. The model is based on biophysical modeling with certain physical attributes of a neuron represented by an electrical circuit that performs an analogous role in an electrical circuit.

Source: Hodgkin and Huxley model [Wikipedia Hodgkin-Huxley 2016].

For example, the bi-lipid layer is represented as a capacitance C_m. Voltage-gated and leak ion channels are represented as non-linear conductances $g_n(t, V)$ and g_L, respectively. The electrochemical gradients driving the flow of ions are represented by batteries E_n and E_L. And the ion pumps and exchangers are represented by current sources I_p. The main task is reduced to modeling the electric potential difference V between the intracellular and extracellular media. This electric potential difference affects the current state of the neuron, and determines the properties of the action potential when it arises.

In the original Hodgkin-Huxley model, synaptic interactions are not considered. However, in the extended Hodgkin-Huxley model the synaptic interactions are also part of the model [Mirowski 2006]. This makes the model useful for representing the dynamics of neural activity within a network of neurons. This is a significant improvement on the original model because information processing is defined at the network level, and not at the individual neuron level. A single neuron cannot extract visual information from the retina on its own: a network of neurons is required to accomplish such a task.

Both the original and extended versions of the Hodgkin-Huxley models are expressed using a set of coupled non-linear differential equations, as illustrated Equation 2.1. The variables in Equation 2.1 have a subscript i which indexes the neuron being considered in the network. The set variables of primary interest are the V_i. As can be seen in the model, these depend on other variables which have their own dynamical properties as well. The variables n_i, m_i, h_i represent dimensionless quantities which range between 0 and 1, representing the potassium

channel activation, sodium channel activation, and sodium channel inactivation. The voltage-dependent parameters α_j and β_j represent the flow rates for the different ion-channels. The parameters $\bar{g}_{Na}$, $\bar{g}_K$ and $\bar{g}_L$ represent the maximum values for the conductances associated with the sodium, potassium and leak ion-channels. $G_{synapse[j,i]}$ specifies the synaptic coupling between neurons i and j. The original model can be recovered if the subscript i is dropped and the terms involving synapses are set to make a zero contribution [Mirowski 2006].

$$
\begin{aligned}
&C\frac{\partial V_i}{\partial t} + g_i(V_i - E_i) = I_i^{applied}(t) \\
&g_i = \bar{g}_{Na} m_i^3 h_i + \bar{g}_K n_i^4 + \bar{g}_L + \sum_{j \neq 1} G_{synapse[j,i]} \\
&E_i = \bar{g}_{Na} m_i^3 h_i E_{Na} + \bar{g}_K n_i^4 E_K + \bar{g}_L E_L + \sum_{j \neq 1} G_{synapse[j,i]} E_{j,i} \\
&\frac{\partial m_i}{\partial t} = \alpha_m(V_i)(1 - m_i) - \beta_m(V_i) m_i \\
&\frac{\partial n_i}{\partial t} = \alpha_n(V_i)(1 - n_i) - \beta_n(V_i) n_i \\
&\frac{\partial h_i}{\partial t} = \alpha_h(V_i)(1 - h_i) - \beta_h(V_i) h_i.
\end{aligned} \tag{2.1}
$$

The computational neuroscience perspective offers invaluable insight owing to its quantitative formulation, as it shows that there is some underlying natural order in how neurons process information. In this work we also adopt a quantitative formulation of brain activity. However, the standard neuroscience representation is not useful for BCIs which employ non-invasive measurement techniques. Three criticisms that lead us to depart from the standard computational neuroscience approach are presented below.

2.2.2 Pragmatic Considerations

The current models make no allowance for purposefully guided neural activity. What is meant is that the role of thinker is not explicitly recognized, and purposeful thinking is consequently not recognized. In our work, the intent of the user is the central concern. The approach is akin to the kinematical description of moving bodies in classical mechanics, in which the evolution of the system is completely determined by intrinsic kinematics of the moving body.

Given appropriate initial conditions and considering stochastic effects, this evolution can in principle be determined within the framework of the model.[5] By and large, the theory provides a description of how a neuron or a group of neurons react to given electrical input and environmental conditions. The question concerning whether the brain fundamentally operates a highly sophisticated input-output mapping is currently within the confines of speculative philosophy. The question need not detain us here. In this work, the intentions of the user are fundamental. Consequently, the computational neuroscience approach in its current form cannot be fully adopted.

[5]There is a formulation of the models that is based on stochastic calculus, but that doesn't detract from the argument put forward.

The next consideration is that the current models are specified in the language of non-linear differential equations. Currently there is no known general computational framework for solving general non-linear differential equations.[6] This short-coming in current analysis techniques makes working with non-linear differential equations challenging from a mathematical perspective. There are numerical techniques that yield numerical solutions to these problems. However, those techniques are not panacea, because they are prone to convergence failure. The problem is further complicated by the fact that the equations appearing in Equations 2.1 have to be solved for each of the approximately 100 billion neurons in the human brain [Purves et al., 2004, pg. 29]. All these problems render the model computationally infeasible. These problems are especially severe in the context of wearable EEG technology, despite the appeal of the models being theoretically rigorous. We seek quantitative techniques which provide aggregate ensemble information.

Taking the lead from physics, we seek a framework akin to thermodynamics for the description of the behavior of gaseous substances. With the understanding that the sought framework may give rise to phenomenological observables, we purposefully refrain from making any inferences about the underlying dynamics within the neural tissue.[7] In short, the approach we seek makes no epistemological claims about brain function.

Lastly, our measurement process does not allow us to directly measure the action potentials that are communicated between neurons. We also cannot determine the other parameters that appear in the extended Hodgkin-Huxley model provided by Mirowski [2006] (see Equations 2.1). These limitations follow from the fact that our technique is non-invasive, and measures aggregate activity on the scalp. The EEG signal that we measure is a spatially aggregated and attenuated correlate signal of the underlying action potentials (see Section 2.3). Thus our approach is not directly compatible with the approach relying on action potentials, or spatially aggregated spike-trains, as used in the electrocorticographic measurement process.[8]

Our approach leads us to employ signal processing (see Section 3) and machine learning (see Section 4) as surrogates to standard computational neuroscience. Both topics are discussed and their use motivated in the indicated sections.

2.3 Electroencephalography

This subsection provides an overview of EEG-related concepts. First, the manifestation of EEG is discussed. Thereafter, EEG measurement standardization and interpretation of EEG data are discussed.

[6] This stands in contrast to linear differential equations which in principle can be solved when the roots of the characteristic equation are obtained.

[7] Taking the lead from physics again, we note that the observable defined as pressure is not a fundamental entity in the same way as energy of the gas molecules. It is however very useful when gas behaviour is investigated.

[8] Electrocorticographic measurements are performed intracranially, thus there is physical contact with the neural tissue. The spatial resolution of the measurement is higher than that of EEG, but it is vastly lower than that achieved when thin electrode filaments are directly inserted into the tissue.

2.3.1 Overview of EEG

EEG is a non-invasive tool for recording electrophysiological activity in the brain. EEG activity is a measure of the difference in electric potentials as measured between two different regions on the scalp. This is achieved by placing electrodes at different locations on the scalp, and measuring the resultant potential difference providing outputs as depicted in Figure 2.8. EEG is important because it provides an alternative to the costly and bulky fMRI technology. An additional advantage of using EEG over fMRI is the signal latency. EEG provides neurophysiological information at a temporal resolution of milliseconds, and this stands in stark contrast to the temporal resolution of fMRI which can go up to six seconds. This means that short-lived dynamics of physiological processes with corresponding electrophysiological correlates can be detected when EEG is used. Low latency is particularly important in situations such as BCI control applications, where relatively short reaction times are required or are desirable. Another benefit of low latency is that the lag between signal detection and task performance is also low, and that makes signal and task synchronization easier to determine. However, all of these benefits should be weighed against the downside of low spatial resolution, which is also a characteristic feature of EEG.

EEG is used in a variety of contexts. In the clinical context, EEG is used to infer clinical brain function abnormalities such as epilepsy. A newer application of EEG is emerging in BCI technology. In BCI technology, the EEG provides a control signal for a computer program which interprets the EEG signal, then executes the appropriate predefined set of instructions. The latter context is the subject of this work.

EEG measures the difference in electric potentials between two different locations on the scalp as it varies in time. Recording electrodes usually consist of small metal cups or disks that are attached to the scalp so that they make good mechanical and electrical contact [Fisch 1999, pg. ix]. The measurements are presented as multi-channel recordings as depicted in Figure 2.8. There are two technologies for measuring EEG activity: wet-electrode, and dry-electrode technologies. The most common and currently most reliable is the wet-electrode technology. Typically saline solution is required to increase electrical conductivity of the skin, as this helps with obtaining higher signal quality.[9] Dry-electrode technology is by-and-large still under development. The signal quality offered by dry-electrode technology is still inferior to that of wet-electrode technology. However, there are problems associated with using wet-electrode technology (see Section 5). The EEG sensor used in this work is based on wet-electrode technology.

EEG is highly susceptible to environmental disturbances. Variations in EEG data are artificially induced by environmental features such as lighting conditions. Other sources of artificial variations are body-part movements from the person wearing the equipment (see Figure 2.8). For example, the eye-ball has the structure of an electric dipole, so when the eye is moved, a time varying electric potential is produced. Another example is muscle tissue: there is a significant amount of electrical activity that is involved during the contraction of muscles. This activity results in what is called electromyography (EMG), producing signals which are usually 2 orders of magnitude higher than EEG. When the nature of the measured phenomenon is taken into consideration (i.e. potential differences in the range micro-volts up to a few milli-volts), it becomes apparent why these environmental features can easily induce variations in the EEG

[9] Saline solution is a mixture of sodium chloride and distilled water.

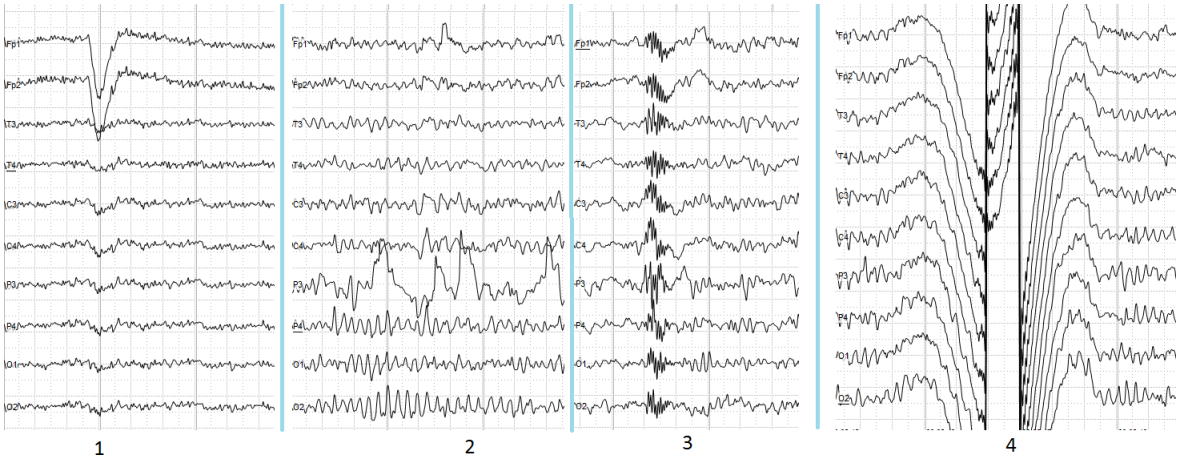


Figure 2.8: The samples of main types of artifacts in human EEG. 1 - Electrooculographic artifact caused by the excitation of eyeball's muscles. 2 - Electrode's artifact caused by bad contact between P3 electrode and skin. 3 - Swallowing artifact. 4 - Common reference electrode's artifact caused by bad contact between reference electrode and skin.

Source: Electroencephalography [Wikipedia electroencephalography 2016].

data. Unfortunately there is currently not much that can be done to handle these sensitivities, apart from creating a controlled environment. Thus we see that signal sensitivities present themselves as serious problems when working with EEG.

In addition to the above sensitivities, EEG data is also inherently noisy and exhibits substantial loss of signal fidelity. Usually, the noise manifests in the EEG data as artifacts. The noise is primarily due to the sensing technology being used. In general this is a universal feature of all measurements. The outlined sources of sensitivities and noise hold back the development of BCI technologies that are robust. Other limiting factors are attributable to the cost of the system. Thus most of the work on BCI technology is still confined to laboratory settings. One of the aims of this work is to investigate the feasibility of using low-cost EEG sensing technology in BCI applications. A further discussion pertaining to signal fidelity is provided in Section 5.

2.3.2 Manifestation of EEG

The source of the EEG is electrical activity generated by nerve cells in the cerebral cortex in response to cortical processing and various kinds of input, including that from pacemakers of rhythmical activity in the depth of the brain [Fisch 1990, pg. ix]. These fluctuating potentials summate and conduct to the scalp where they can be recorded as the scalp EEG [Fisch 1990, pg. ix].

Since each neuron makes a contribution to the total EEG that is measured, it follows that neurons located in neighbouring regions to the neurons of interest create an interference pattern. Roughly speaking, this can be considered as a form of noise because the signal of interest is contaminated by the signal from neighbouring regions.[10] The noise is accentuated

[10]In strict terms, labeling the interference as noise is inappropriate since what we consider as interference is

when the region of interest is located within a cortical fold, and it is surrounded by other regions which are also activated when a certain task is performed. In addition to noise, there is signal attenuation that occurs while the EEG signal is in transmission towards the scalp. The attenuation is induced by the dura mater and the bone tissue that forms the cranial cavity. The effect of the mentioned tissue structures is not negligible, and is one of the major reasons why electrocorticography provides better results than EEG.

In broad terms, EEG activity is categorized into five groups according to the frequency of the signals.[11] Decomposing a multi-frequency signal into various constituent signals is the subject of Section 2.13. Figure 2.9 illustrates how the various categories are demarcated. A brief outline concerning the function of each of the rhythms is also provided below.

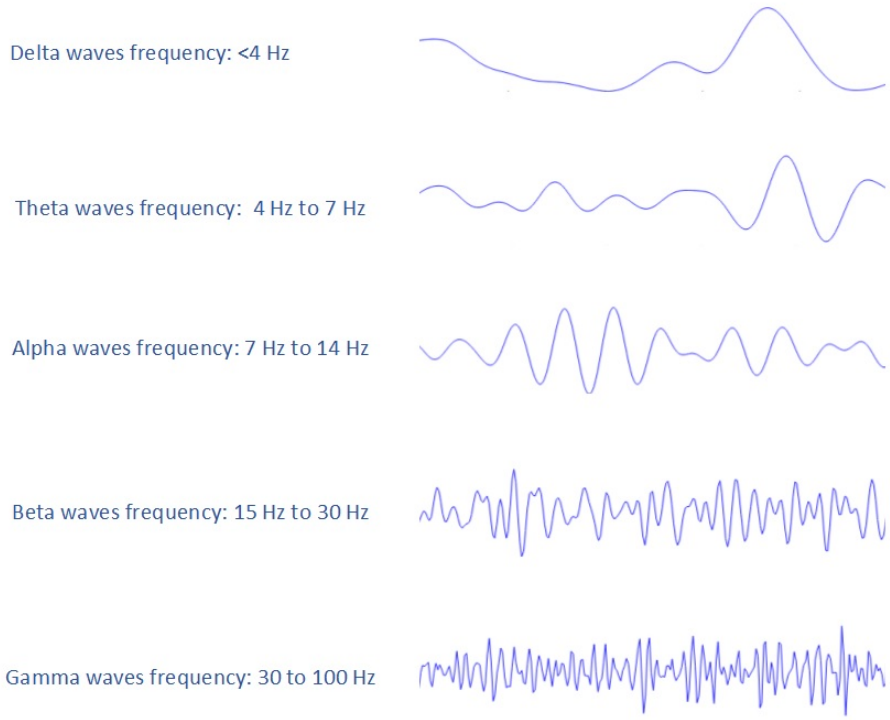


Figure 2.9: EEG signal decomposition into primary EEG bands. The frequency decompositions of the EEG signal represent the current categorization scheme. All except the gamma waves have been attributed to specific brain activiy.

Delta rhythms are slow brain activities typically preponderant only in deep sleep stages of normal adults. Otherwise, they may be indicative of pathologies [Adeli and Ghosh-Dastidar 2010, pg. 75].

Theta rhythms exist in normal infants and children as well as during drowsiness and sleep in

also part of brain physiology. That is to say, we have no epistemological grounds to use labels such as noise and interference in the strict interpretation of the words. In this context, the words are to be interpreted loosely so as to avoid excessive reductionism.

[11]There are also mu and sensory-motor rhythms that are contained within the categories, but we leave them out of the discussion.

adults. Only a small amount of theta rhythms appears in the normal waking adult. Presence of high theta activity in awake adults suggests abnormal and pathological conditions [Adeli and Ghosh-Dastidar 2010, pg. 75].

Alpha rhythms exist in normal adults during relaxed and mentally inactive awakeness. The amplitude is mostly less than 50μV and appears most prominently in the occipital region [Adeli and Ghosh-Dastidar 2010, pg. 75].

Beta rhythms are primarily found in the fronto-central regions with lower amplitude than alpha rhythms. They are enhanced by expectancy states and tension [Adeli and Ghosh-Dastidar 2010, pg. 76].

Gamma rhythms in contrast to the other rhythms, are usually not of much clinical and physiological interest and therefore often filtered out in EEG recordings [Adeli and Ghosh-Dastidar 2010, pg. 76].

In this work, the EEG bands that are used correspond to the alpha and beta rhythms. As discussed in Section 2 of Chapter 5, certain frequency windows are selected from the rhythms and these frequency windows are associated with voluntary movement, and state of concentration. Pfurtscheller et al. [1976] presents findings pertaining to power spectrum changes when the brain is engaged in certain tasks. Pfurtscheller et al. [1976] reports the changes in the power spectrum as indicative of event-related cortical desynchronization of the EEG, and this desynchronization is a result of activation of certain cortical regions.

2.3.3 EEG Measurement Standardization

EEG measurement standardization pertains to the definition of a standard template for conducting EEG measurement experiments. The measurement templates are called montages, and they define the spatial configuration for electrode placement on the scalp. A montage can be defined by selecting a set of electrode placement sites using the general montage diagram in Figure 2.10. The selected set would then constitute a montage.

In order to assist in experimental design standardization, the international 10-20 system for placement of EEG electrodes was developed (see Figure 2.11). The system ensures that experiments or measurements can be reproduced with ease. The standard is essential since there are non-negligible differences in the structure of cortical folds and head size variability between people (see Section 2.1). The 10-20 system ensures that electrode placement locations approximately correspond to the same functional regions even across different individuals. In this work, we use wet-electrodes and the montage used has a fixed electrode placement layout which covers the cortical areas labeled as: AF3, F7, F3, FC5, T7, P7, O1, O2, P8, T8, FC6, F4, F8, and AF4.

2.3.4 EEG Interpretation

Interpretation of EEG data pertains to the extraction of meaningful information from the data. The information may be used for various purposes including clinical diagnosis, and BCI technology. Depending on the context of study, the actual data that is being studied varies according to some transformation that is applied to the raw EEG data. For example, in the clinical context, the data that are studied are the multi-channel voltage recordings. The interpretation of EEG data is of paramount importance when assessing the presence of certain neurological

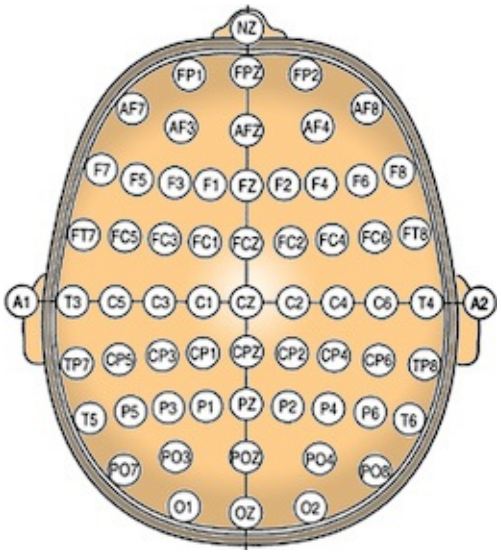


Figure 2.10: General montage template using the 10-20 electrode placement system. The topographical layout of the nodes is intended to provide full coverage on the scalp.

Source: Anatomy and physiology [Neurofeedback 2016].

disorders such as epilepsy. In the context of BCI technology, the data that are usually studied are obtained after some frequency transformation is applied to the data multi-channel voltage recordings. When the concept is applied to BCI technology, it allows us to infer the intentions of the user of the BCI, thereby providing a communication link that bypasses the standard neuromuscular connection.

In clinical applications, the interpretation is currently carried out by skilled personnel who, as a result of extensive training, can notice irregularities which have been established primarily through empirical findings. As can be seen in Figure 2.8, the structure of the EEG data is non-trivial, and extracting irregularities is challenging. In the context of BCI technology, the task is typically carried out by a machine learning inspired model that has been trained and validated with experimental data (see Section 4). There are substantial practical difficulties associated with EEG interpretation. The difficulties arise owing to different factors, one of which being the difficulty of source localization as outlined in Section 2.1. However, other factors are dependent on the context in which the task is performed.

In the clinical context for example, another factor is that the services rendered by skilled personnel are expensive. Additionally, even access to those services is a difficulty in its own right, depending on geographical location. In Africa, this problem is especially severe. According to a World Health Organization (WHO) report from 2004, there were only 111 neurologists practicing in South Africa. There were no practicing neurologists in some African countries such as Botswana, Namibia, and Mozambique.[12] In a 2013 follow up study aimed at assessing progress in the increase of health workers in Africa, WHO found no significant improvements.

In the BCI technology context, the interpretation of the data is typically treated as a machine learning problem (see Section 4.2). There are various difficulties that arise in this

[12]Note that this list is not exhaustive.

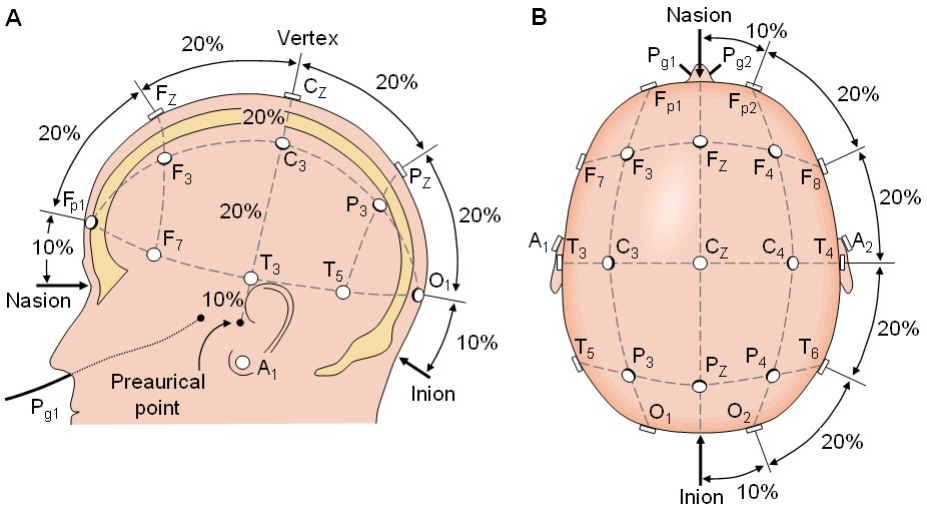


Figure 2.11: Illustration of the proportion marked 10-20 electrode placement system. The relative demarcation is designed such that differences in head sizes have a minimal influence on inter-subject comparative measurements.

Source: American EEG society guidelines for standard electrode position nomenclature [Sharbrough et al. 1991].

context. The difficulties range from finding:

- a transform that makes the underlying structure easily discernable. To this end we use frequency decomposition (see Section 2.13).
- an appropriate set of feature vectors to represent the information. To this end we use frequency values motivated by neuroscience findings as outlined in Section 2.3.2.
- an appropriate model that will be able to effectively learn the structure from the data. To this end we use probabilistic graphical models (see Section 4.3).

The next section provides a discussion of signal processing concepts. These concepts are required for the purposes of signal transformation. In this work, the data that is studied is the power-spectrum of the underlying potential differences which are recorded as a time series.

3 Signal Processing Theory

There are numerous problems in which we find that the solution is readily obtained if the basis of representation is transformed from the default representation. Formally, this procedure is known as changing the basis set. The elements of the set are used to describe an entity of interest within the space defined by the basis set and operations on those basis set elements. For example, sound can be studied using at least two basis sets of representation. One approach would be to study sound in the temporal domain, in which case the fluctuations as they appear in time are pronounced. Frequency information is obscured in this representation. Another approach would be to study sound in the frequency domain, in which case the fluctuations as they appear in frequency are pronounced. Temporal information is obscured in this domain.

Signal processing theory provides a mathematical framework which can be used to induce

the representation basis transformations on the input data. In this work we only consider temporal-to-frequency transformation of the input signal. The sections below provide motivation for using the aforementioned transformation.

3.1 Feature Vectors

Feature vectors are representations of the state of the system being studied. Typically these structures contain descriptors of the data from which they are extracted. For example, these descriptors may be the mean value and the variance of a particular portion of the data. Extraction of feature vectors is the process in which data is processed by some algorithm that takes the input data and computes the descriptors.

One of the basic problems in feature vector extraction is transforming a given input signal from one space of representation into another space in which the structure that is sought can be easily discerned. This concept is useful in the proposed work because the structure that we seek is more easily discerned in the frequency domain than in the temporal domain. Figure 2.12 provides an illustration of this idea using EEG data. It is easier to discern the similarity between the blue and red curves when the frequency basis is used, as opposed to when the temporal basis is used. This observation and other studies in neuroscience (see Section 2.3.2) necessitate that the signal be transformed from the temporal domain into the frequency domain. In the frequency domain, the feature vector can then be simply defined as the energy contributions of certain frequency intervals. The subsection below addresses the question concerning the exact details of how the required transform works.

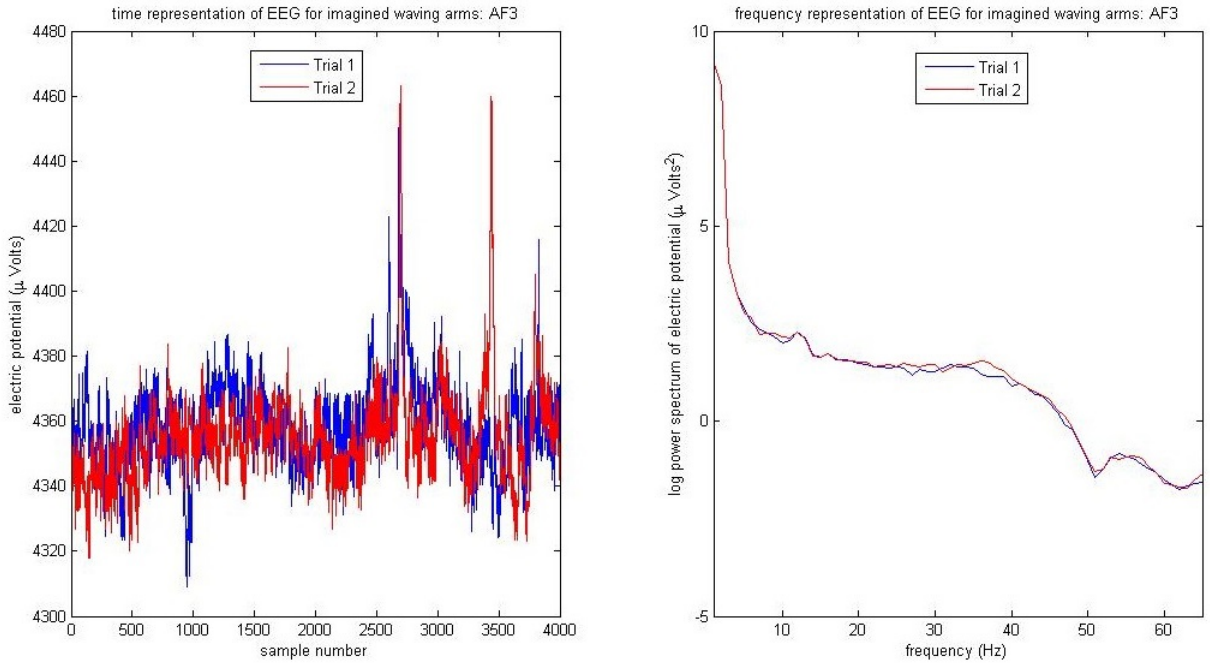


Figure 2.12: Illustration of the effect of the time-to-frequency domain transformation of a signal.

3.2 Frequency Decomposition

Frequency decomposition concerns transforming a signal expressed in the time domain into another signal that is expressed in the frequency domain. A common approach for achieving this objective is the Fourier transform. The standard Fourier transform is defined in a continuous domain, in which the variables can assume any real number. However, there are numerous variations of the standard Fourier transform. Amongst the variations are the Fourier transform defined in discrete domains and the fast Fourier transform with reduced algorithmic complexity defined in discrete domains. The Fourier transform defined in a continuous domain is described next.

3.2.1 Fourier Transform

The core of the idea is that a signal with a non-trivial time signature is a superposition of infinitely many stationary sine and cosine functions. The Fourier transform (FT) then determines what the scaled contributions each of the sine and cosine functions make to the given signal. The basic intuition is depicted in Figure 2.13. The mathematical formulation of the Fourier transform is given below.

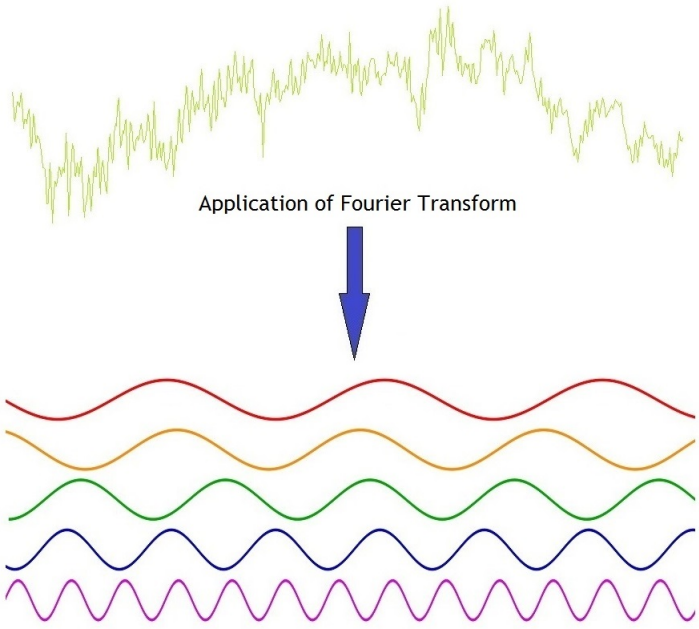


Figure 2.13: Illustration of the decomposition of a signal into its constituent signals. Points of similarity or dissimilarity are easily identifiable in the frequency domain.

Let $f(t)$ represent the continuous time function which describes the given signal. Let $\hat{f}(\xi)$ represent the Fourier transform of the function $f(t)$ at frequency ξ, and $e^{2i\pi\xi} = cos(2\pi\xi) + isin(2\pi\xi)$. The Fourier transform of $f(t)$ is defined as

$$\hat{f}(\xi) = \int_{-\infty}^{\infty} f(t)e^{-2i\pi\xi}dt$$

with the inverse Fourier transform of $\hat{f}(\xi)$ which is equivalent to $f(t)$ is defined as

$$f(t) = \int_{-\infty}^{\infty} \hat{f}(\xi) e^{2i\pi t} d\xi.$$

A common usage of the Fourier transform is in representing a given function by an infinite series of sine and cosine functions. There are theoretical reasons for choosing these two functions as the basis set. One of the reasons is that the functions are mutually orthogonal, and this significantly reduces the complexity of the representation.

Continuing with the above-mentioned $f(t)$, the Fourier transform permits the following representation of $f(t)$:

$$f(t) = \frac{a_0}{2} + \sum_{j=1}^{\infty} a_j cos(jt) + b_j sin(jt)$$

where the a_j and b_j are real numbers defined as follows:

$$a_j = \frac{1}{\pi} \int_{-\pi}^{\pi} f(t) cos(jt) dt$$

$$b_j = \frac{1}{\pi} \int_{-\pi}^{\pi} f(t) sin(jt) dt.$$

The basis functions cosine and sine form an orthogonal basis, and this property simplifies the process of computing the inner products $< f(t), cos(jt) >$ and $< f(t), sin(jt) >$. This follows because the inner product of cross-terms evaluates to zero.

3.2.2 Fast Fourier Transform

In the case in which the signal (i.e. $f(x)$) is sampled at discrete points in time, the resulting formulation of the FT changes slightly. For the purposes of this work, we describe the Fast Fourier Transform (FFT) which we employ as a variation of the discrete Fourier transform (DFT). The FFT algorithm is a performance enhanced version of the DFT. The standard DFT has algorithmic complexity of $O(n^2)$. Whereas the FFT has algorithmic complexity of $O(nlog(n))$ owing to the recursive nature of the algorithm. The performance enhancement makes the FFT better suited to situations in which time-to-compute constraints are considerable. The FFT algorithm outlined in Algorithm 1 is known as the Cooley-Tukey algorithm.

3.3 Signal Sampling

Signal sampling concerns measurement of a continuous signal at discrete points in time, such that the resulting sampled signal is discrete in form. An illustration is provided in Figure 2.14. In many real-world applications, it is sufficient to work with discrete signals. Sampling is mandatory when digital electronics are employed as part of solving the problem. This happens because the input signal has to be converted from analog to digital format. Analog-to-digital converters are known to have physical limitations. Thus real-world sampling is an imperfect process with regards to retaining exact details of the input signal.

Algorithm 1 Outline of the Cooley-Tukey algorithm for computing the FFT.

function FFT($x_{0,s,2s,...,(N-1)s}, N, s, X$):
  if $N = 1$ **then**
    $X_0 \leftarrow x_0$
  else
    $X_0, ..., _{\frac{N}{2}-1} \leftarrow \text{FFT}(x_{0,2s,4s,...}, \frac{N}{2}, 2s, X)$
    $X_{\frac{N}{2}}, ..., _{N-1} \leftarrow \text{FFT}(x_{s,3s,5s,...}, \frac{N}{2}, 2s, X)$
  end if
  for $k = 0 : \frac{N}{2} - 1$ **do**
    $t \leftarrow X_k$
    $X_k \leftarrow t + e^{\frac{-2\pi ik}{N}} X_{\frac{k+N}{2}}$
    $X_{k+\frac{N}{2}} \leftarrow t - e^{\frac{-2\pi ik}{N}} X_{\frac{k+N}{2}}$
  end for
end function

In this work, the signals of interest (i.e. EEG signals) are continuous in time. However, we are forced to sample the signal in the process of analog-to-digital conversion. Sampling rate is one of the fundamental concepts that arise when signal sampling is studied. The sampling rate is the number of measurements of the signal intensity performed per unit time. A visual depiction is provided in Figure 2.14, in which the period T is determined by the sampling rate of the analog-to-digital converter. The exact relationship between the sampling rate (f) and T is $f = \frac{1}{T}$. The Nyquist-Shannon theorem which is described below, provides the mathematical criteria required to make reliable transitions from analog signals to discrete signals. In particular, the Nyquist-Shannon theorem answers the question concerning the selection of the appropriate sampling rate so as to preserve the information contained in an input analog signal.

3.3.1 Nyquist-Shannon Theorem

The Nyquist-Shannon theorem specifies sufficient criteria for signal sampling. Using the appropriate sampling rate is particularly important when analog-to-digital conversion is performed, because discretization introduces errors.

The essential idea behind the theorem is as follows. Suppose the signal $x(t)$ is passed through a low-pass filter such that no frequency components beyond F Hz are present in the resultant signal, then the resultant signal can be reproduced by a discretization scheme for which every two consecutive samples are $\frac{1}{2F}$ seconds apart. It follows that the sufficient sampling rate is $2F$ samples per second. Although it is beyond the control of the user of the digital-to-analog converter, it is important that the sampling frequency doesn't fall below the Nyquist frequency so as to preserve signal fidelity. The variation in the sampling frequency has considerable consequences. This topic is further discussed in Section 3.1 of Chapter 6 in which we discuss the irregularities found in the experiment results.

3.3.2 Spectrogram Representation

The spectrogram provides a visual representation of the frequency spectrum of a signal as it evolves in time. The spectrogram is usually computed using the short-time Fourier transform algorithm. The information provided by the spectrogram is important when the structure that

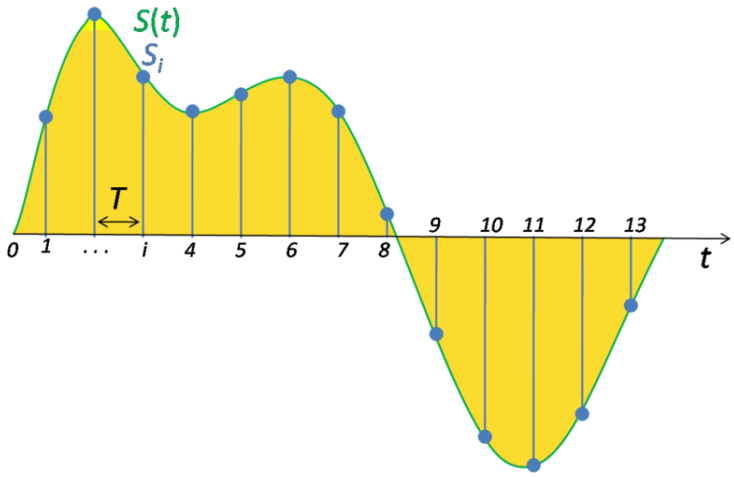


Figure 2.14: Discrete sampling of a continuous signal, the sampling rate is usually a fixed value during sampling. The sampling rate determines how many samples are measured within a given time interval, the higher the sampling rate, the better the signal is approximated.

Source: Sampling (signal processing) [Wikipedia signal processing 2016].

is being sought in the data appears in the frequency domain, and especially between frequency spectrum samples over a certain duration in time. In this work, the spectrogram is used to provide information concerning the temporal evolution of the EEG signals. Samples of the power spectrum of the EEG signal are arranged next to one another according to the temporal order in which they are obtained.

Typically, the spectrogram is presented in a two-dimensional figure. The horizontal axis represents time, and the vertical axis represents the frequency value. In order to depict the magnitude of a frequency component at a particular time, a colour-bar is introduced in order to represent the magnitude of the frequency components. The resulting figure resembles a heat-map with two spatial dimensions. Figure 2.15 provides an example of a spectrogram.

In the context of this work, there are difficulties that arise when the spectrogram is used directly as the state information to be used to determine the mental task being performed by the user. The difficulties are that the spectrogram does not have a fixed structure as time progresses (i.e. the signal is non-stationary). The non-stationarity also appears to be stochastic, in the sense that the temporal variation of the spectrogram does not appear to produce a discernable and predictable pattern. This greatly increases the difficulty of determining which spectrogram is associated with which mental task. To overcome the above-mentioned difficulties, we seek a scheme that would be able to find a representation that would model all the spectrograms that belong to the same mental task. An added constraint is that the scheme should be able to differentiate between spectrograms that belong to different mental tasks. The application of machine learning is appropriate when such solutions are sought.

Briefly, one of the main benefits of using machine learning is that the researcher is freed from finding a mathematical model that would achieve the stated objectives. The procedure that is followed is such that the chosen generic model learns how to perform the association between the given inputs and the desired outputs. The learning process is dependent on the given experiment data. The exact algorithm that is applied is determined by the class of the machine learning technique used. In the next section, a discussion of machine learning is provided.

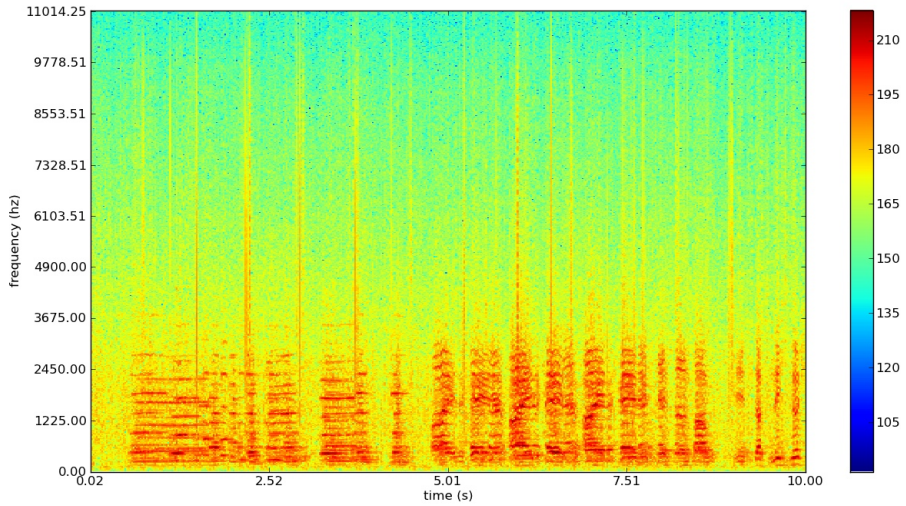


Figure 2.15: Graphical representation of a signal spectrogram. Along the vertical axis are frequency values, along the horizontal axis is time, and the colour bar represents the magnitude of the frequency component. With the three pieces of information it is possible to visualize the temporal evolution of the frequency components.

Source: Spectrograms [Zalkow 2016].

4 Machine Learning Theory

Machine learning provides tools that can be used to enable a computer program to automatically learn to associate certain input signals to certain output signals. The learning algorithm modifies the parameters of a function in order to achieve the required association. These parameters are determined by the kind of machine learning paradigm used for the task, as well as the form of the function. The basic idea behind the development and use of machine learning is that the computer program should be able to perform a particular task, and this happen without a researcher having to explicitly write the instructions that have to be followed. By imposing the constraint on the researcher, we expect that the task should be learned from examples. This is particularly useful in cases in which even the researcher does not know the exact instructions or models that are required for performing the task. An example of such a task is recognizing a particular face in an image which contains many other faces.

4.1 Function Approximation and Classification Techniques

In our discussion, we contrast two different categories of problems that are studied in machine learning. These problems are known in the literature as regression and classification problems respectively. First, we consider problems in which there is noisy data, and we wish to find the model that best describes the data as whole. These problems generally fall under the scope of function approximation techniques. Second, we consider problems in which there is compartmentalized noisy data, and we wish to find the model that best differentiates the data such that it is associated with a particular compartment.

Both categories are described for completeness but for the purposes of this work, the classification techniques are required for the problem at hand. The function approximation is of no significance since we don't have a model that is based on a function that we intend to use

for interpolation or extrapolation purposes. As discussed in Section 3.3.2, the main task is to determine in which class of mental activity does each of the observed spectrograms belong.

4.1.1 Function Approximation Techniques

There are many problems in machine learning, other areas of mathematics and statistics in which finding an approximation of some function is required. What is usually given is the experiment data, and the problem is reduced to finding a function that best models the data. Usually, there are constraints which arise from theoretical insight or tractability considerations that have to be satisfied. The constraints may restrict the functional form of the model being sought, in which case the parameters take on different interpretations. For example, model parameters can be the mean and variance values of a mixture of Gaussian distributions. The main problem then is to learn the best parameter values for the chosen model in order to capture the structure of the given data.

Function approximation is important for various reasons. First, we may be interested in interpolating between given data-points obtained from experiments. Second, we may be interested in extrapolating beyond the data-points obtained from experiments. In both scenarios, the appropriate function would have to be obtained. Following from that, the appropriate parameter values would have to be determined using the experiment data. Usually, the task is formulated as a regression problem.

The basic assumption is that there is an unobservable underlying model which is the generator of the observed data, and the task it to find an approximation of that model. As illustrated in Figure 2.16, we have the observed noisy data (red dots), the assumed unobservable underlying model which is the data generator (purple curve), and the estimated model which is to be used as the learned model (green curve).

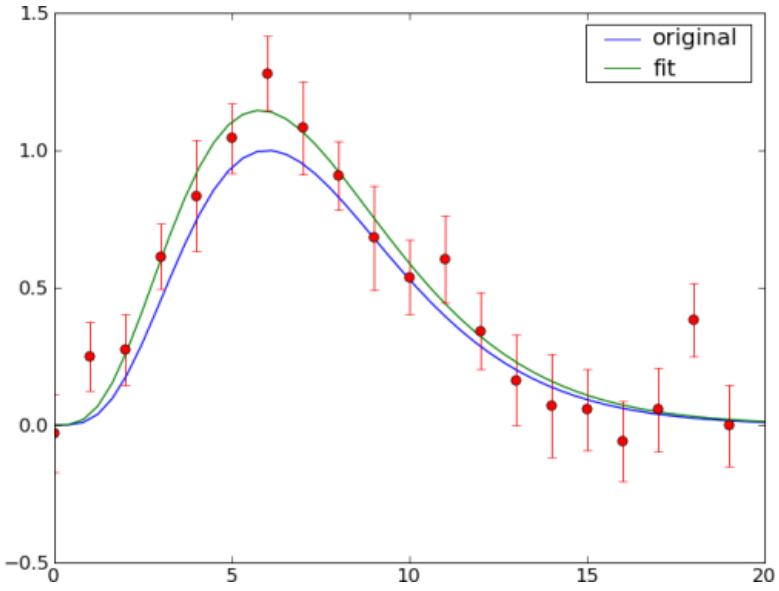


Figure 2.16: Function approximation with noisy data. The vertical lines on the data points represent error or variation in the measurement.

Source: Curve fitting [PyMVPA 2012].

It is important that the approximation procedure which is employed doesn't over-fit the model to the given data.[13] Over-fitting occurs when the model parameters are modified to the extent that there is almost no mismatch between the given data and the model. When over-fitting occurs, the model performs sub-optimally on unseen data because the model would have not discounted the noise in the data. In reality, it is impossible to perfectly reconstruct the true underlying model using the experiment data, which follows from the fact that the exact details of the noise signature are not known. As illustrated in Figure 2.16, there is always some degree of mismatch between the approximation and the true underlying model.

4.1.2 Classification Techniques

In contrast to function approximation problems, there is another category in which the main problem is to determine the association between input data and a set of abstract classes. The main problem then is to learn which data-set elements belong to which class. Similar to function approximation, the given experiment data is also noisy, and the problem is reduced to determining the model that best differentiates the data. There are also constraints pertaining to computational tractability and the functional form of the model. Similar to function approximation, the main problem is to learn the best parameter values for the given model to differentiate between the sets described by the data. Consider the following problem.

Suppose we wish to automatically remove defective products from an automated assembly line. We may attempt to determine a representation that captures the structure pertaining to defectiveness, but that could prove to be a very challenging task. A much more feasible solution is to find a machine learning technique which would automatically extract structure pertaining to defectiveness using input data which may be in the form of images. For example, defectiveness could be associated with scratches in particular regions on the product. The scratches could be inferred if edge detection were to be applied in the data preprocessing stage. In this context, it doesn't matter what shape or size the scratch assumes, the prominence and location of the scratch would be the primary measures. The machine learning technique would then learn the association between the abstract class (i.e. defective, or not defective) and the given input data (i.e. the set of images). The basic assumption is that there exists some well defined decision boundary that demarcates the space separating the elements of one class from the elements from another class. The basis for the differentiation is based on the information provided by the feature vectors. The feature vector is a set of measurements of certain attributes of the entity being studied. Under this formulation of the problem, feature vector classification assumes central importance.

Feature vector classification involves grouping together feature vectors that are similar into abstract classes, as defined by some similarity metric. Figure 2.17 provides a visual representation of a decision boundary between data-set elements that represent dogs (represented by blue plus symbols), and those that represent cats (represented by red asterisks). There are two measurements that are made to form a datum, the first may be the mass of the animal and

[13] Over-fitting can be monitored and corrected for by using cross-validation techniques. Cross-validation entails dividing the data-set into two subsets. One of the subsets is used during the training phase, in which the model is assessed on how well it learnt the implied structure in the subset. The other subset is used during the validation phase, in which the model is assessed on how well it can extrapolate to unseen data, assuming that the implied structure in the data is invariant.

the second may the length of the tongue of the animal.[14] As illustrated in Figure 2.17, the attributes of the feature vector are represented as the horizontal and vertical axes, respectively.

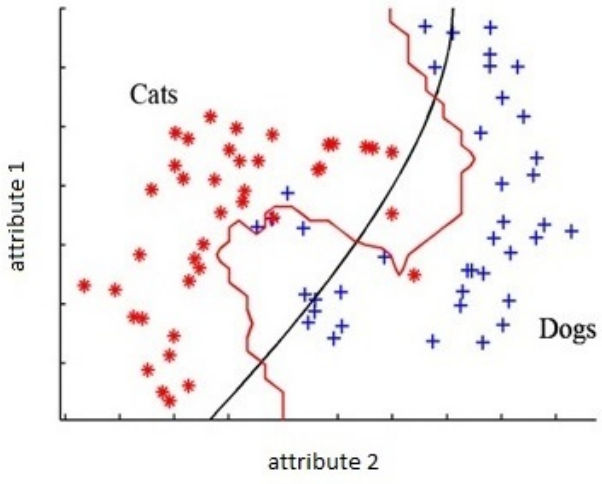


Figure 2.17: Two class animal classification with a decision boundary between the sets.

Source: Classification [Ivan 2016].

Measuring only a few attributes has advantages and disadvantages. One advantage is that the classification problem becomes easier to handle in computational terms. High dimensional feature-spaces entail problems of considerable complexity. By using a low dimensional state-space those problems are circumvented.[15] One disadvantage which leads to misclassification is that with the limited information, it becomes harder to disambiguate elements from different classes. Consider the animal classification problem discussed above. If only mass and tongue length are used as attributes, then the feature vector of a chihuahua would be quite similar to that of a cat, mostly owing to the mass attribute. This problem is accentuated in the case in which the chosen features are also correlated. The dependence between correlated features means that there is a need for discriminative models which do not make independence assumptions on the features.

As illustrated in Figure 2.17, there are at least two (and technically infinite) decision boundaries which are used differentiate the data, and they appear as the red curve and the black curve. The jagged red curve appears to be a good demarcation, but there are challenges that are associated with demarcation of that form. One of the main challenges is that it is not immediately clear how to mathematically specify the curvature. In practice, the black curve would provide a more feasible demarcation since a relatively simple parametric model could be specified to represent the curve.

Similar to function approximation, the problem of over-fitting also occurs in classification

[14]The features are probably correlated, and that introduces some problems, but this does not affect the analysis presented in the argument.

[15]Under-sampling of the state-space is one of the problems, obtaining sufficient data may be costly or practically infeasible. But perhaps a more important problem is the curse of dimensionality: a significant number of machine learning algorithms have an exponential algorithmic complexity on the number of dimensions of the state-space.

if the exact properties of the data-set are learnt. Learning the exact details of the data-set is undesirable because the exact details of noise signature in the data-set are unknown. Additionally, over-fitting can be monitored and corrected for by using cross-validation techniques, see footnote 15 in Section 4.1.1. Classifying data into a set of abstract classes eliminates the need to interpret or treat all vectors that can be sampled as different entities. While this may not be a requirement in problems with very small discrete sample spaces, it becomes a requirement in large spaces and mandatory in continuous sample spaces. Classification is required in the context of EEG data processing because the space of allowed values is continuous in voltage-amplitude, frequency and contribution in energy in each channel over any frequency value.

4.2 Motivation for Machine Learning Application

There are at least two reasons for applying machine learning techniques to sensor data. First, sensor data tends to contain noise that a researcher cannot easily remove. The structure of the noise may be complicated by context or temporal sensitivity. This follows from the fact that physical devices cannot be manufactured to be perfect. Second, the mapping between the input data and the required output data may be complex in its own right. This may occur if the input data has high dimensionality, because this usually makes visualization a very complex task. It may also happen if the structure in the data that is sought requires non-trivial transformations in order to make the structure evident.

When machine learning methods are employed, the task of creating the required association is usually reduced to an optimization problem. The optimization problem is then solved using one of various rigorous techniques that find an acceptable solution in the model parameter space. There are numerous paradigms in machine learning, and consequently this work focuses on a select class of supervised techniques. In particular, we use supervised models which require that the training data have ground-truth labels. In the context of this work, this translates into providing labels which indicate in which class any given data-set element belongs to. The performance of the classifier is then assessed by examining the classification accuracy on previously unseen data, which is also labeled.

In this work, machine learning techniques are used to associate certain EEG measurements with abstract classes which will then be interpreted as certain instructions that the computer must execute. The essence of the idea is to learn to correctly classify the EEG data as it is differentiated by different mental tasks. To this end, we seek to correctly classify the feature vectors. There are various techniques that may be used for this purpose. However in this work we focus on a particular set of probabilistic graphical models (PGMs).

4.3 Probabilistic Graphical Models

PGMs provide a representation in the form of a graph which can compactly represent a high-dimensional distribution with conditional independence relationships between random variables. The formulation of the PGMs makes them versatile in the sense that the representation is declarative. One of the advantages of using a declarative representation is that the knowledge and reasoning are separated. In this approach, we construct a model of the system about which we would like to reason. The model encodes our knowledge of how the system works in a computer readable form. The representation has its own clear semantics, separate from the

algorithms that one can apply to it. Thus it can be manipulated by various algorithms that can answer questions based on the model [Koller and Friedman 2009, pg. 42].

In this work, we report experiment results obtained using Hidden Conditional Random Fields (HCRF) and Latent Dynamic Conditional Random Fields (LDCRF). A description of each of these models and the classification performance that they provided is presented in Section 3 of Chapter 5 and Section 4.2 of Chapter 6, respectively. In the discussion that follows below, we provide a broad overview of the PGM machine learning approach.

4.3.1 Model State-Variables Specification

There are problems in which the researcher can readily specify the feature labels in a datum. This is done to simplify the task that the model has to perform. An example in which such a situation is encountered is as follows. Suppose we wish to create a supervised model that can distinguish between images of smiling people from images of frowning people. One approach would be to present labeled images to the classifier, with each image having a label of either "smile" or "frown". In addition to that, we could also mark the sections of the image in which the lips are, and this would form one region of interest. The region of interest could also be labeled as "up" when there is a smile, or "down" when there is a frown.

The task of the model would then be to learn the association between the: label of the image, label of the region of interest, and the pixel information structure in the region of interest. The structure in the region of interest could be the number of connected pixels which form a concave or convex shape. The model would then have fewer features to handle because the researcher would have reduced the number of possible features that could be learnt in order to perform the classification. For example, the presence of a tree in some of the images would not be part of what has to be learnt, even though it could be used as part of the decision boundary between the images from different classes. This essentially means that the explicit features that would be part of the model would include the regions of interest, and not unspecified aspects of the image.[16]

Unfortunately, there are also problems in which the above mentioned approach fails. The context of this work is one such example. In order to circumvent the difficulty of specifying the explicit state variables, models that have hidden state variables are used. Hidden state variables allow the model to automatically encode the structure in a datum without the researcher having to explicitly specify it.[17] Using the above classification problem as an example, a model with hidden states could automatically discover the significance of the region of interest which corresponds to the lips in an image. In addition to discovering the discussed region of interest, the model would probably also discover that the presence of a tree is useful information for classification, and that this information would also be used for deducing the classification value of the image. In such a scenario, the researcher need not specify the labels "up" or "down". The model would build its own hidden representation of that information. In this work, we focus exclusively on PGMs with hidden states so that we avoid having to specify intermediate structure relevant to our problem.

[16] Note that the state variables can be either discrete or continuous.

[17] Although the researcher could well know it is there.

4.3.2 Generative versus Discriminative Models

There are two common forms in which the models with hidden states appear (see Figure 2.18). The first form is generative, in which the model learns the joint probability distribution between observables and the hidden states. An example of such a model is the hidden Markov model (HMM). One of the benefits of using such models is that samples can be generated from the model itself after it has been trained. One of the disadvantages is that when the joint probability distribution is not required, the generative models do not provide excellent performance. This disadvantage is typically observed in classification tasks, such as the task in this work.

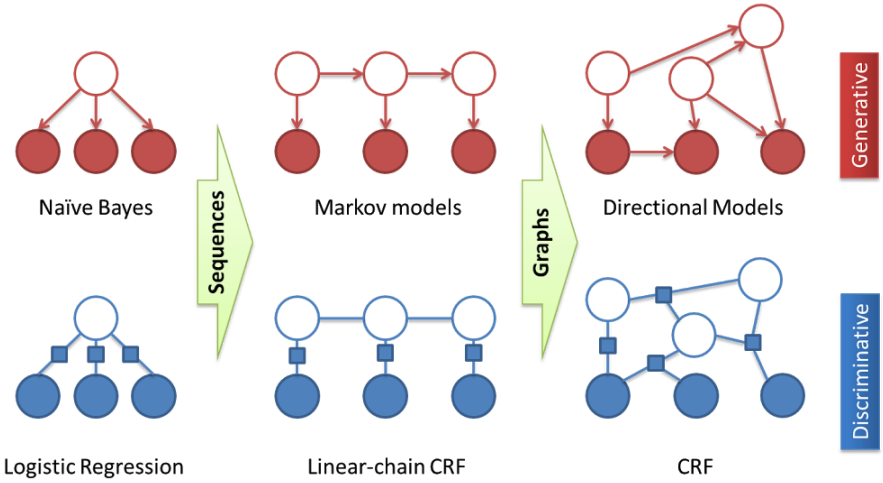


Figure 2.18: Graphical illustration of the structure of probabilistic graphical models.

Source: An introduction to conditional random fields [Sutton and McCallum 2010].

The second form of the models is discriminative. An example of such as model is the HCRF. Much like the HMM, the HCRF is also capable of providing a compact representation of the underlying probability distribution. It is also capable of modeling temporal structure in the data. However, the HCRF circumvents some of the restrictions that have to be made on the HMM for the purposes of tractability. An example of such a restriction is that assumptions about independence of the data at each time point conditioned on the states should be made [Saa and Çetin, 2012]. The HCRF circumvents this by defining a joint distribution over the class label and hidden state labels conditioned on the observations, with dependencies between the hidden variables expressed by an undirected graph [Quattoni et al., 2007]. The HCRF uses the hidden variables to model the latent structure of the input domain [Quattoni et al., 2007].

In summary, the main difference is that the discriminative models provide the conditional probability distribution of the observed data, as opposed to the joint probability distribution which is provided by the generative models. Discriminative models tend to provide better classification performance than their generative counterparts [Saa and Çetin 2012, Quattoni et al. 2007, Morency et al. 2007]. In this work, we experiment with discriminative PGMs with hidden state variables.

4.3.3 Model Representation and Properties

As discussed above, PGMs make use of a graph structure in order to represent the probability distribution it encodes. The probability distribution is represented as a collection of nodes connected with edges. The nodes represent disjoint, but possibly related aspects of the complete probability distribution.

The interpretation of the nodes varies depending on the type of network model selected. For instance, if a Bayesian network is used, then the connected nodes represent probability distributions. In a Markov network, the connected nodes represent factors which are functions of the given variables. Both network models are depicted in Figure 2.19. The distinction being that the factors are not necessarily normalized, and thus they do not necessarily represent probability distributions. In the case of the Markov network, a normalization partition function Z is defined, as depicted in Figure 2.19. The partition function is further discussed in Section 3 of Chapter 5. In both kinds of networks, if there is a direct relationship between the nodes, then an edge connects the two nodes.

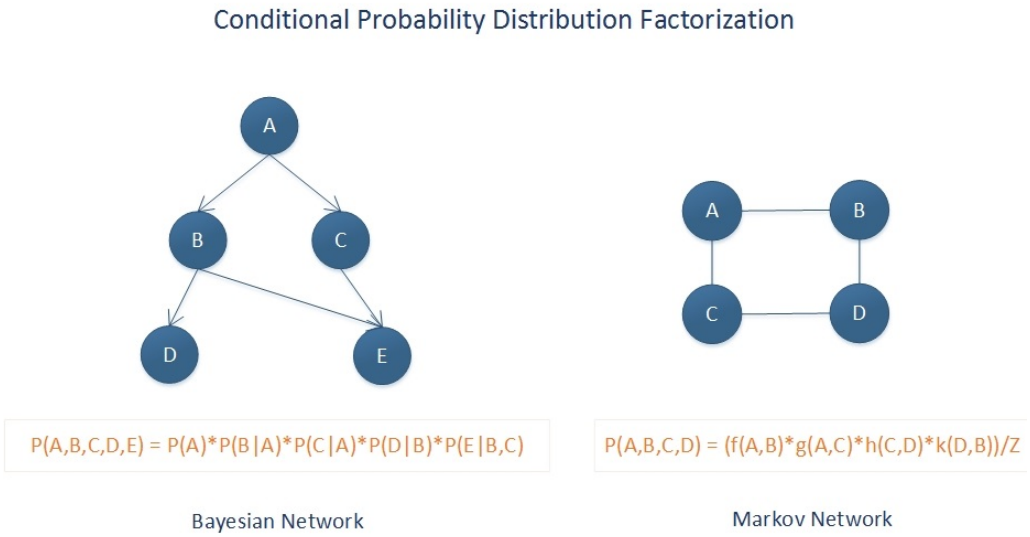


Figure 2.19: Bayesian and Markov probabilistic graphical models.

Using a graph provides various advantages, for example, the ability to readily represent the conditional independencies between variables, and the ability to represent the complete probability distribution in terms of statistically independent groupings of the variables. These advantages facilitate efficient querying of the model, and facilitate intuitive reading of the representation (see Figure 2.20).

4.3.4 Model Parameter Estimation

The graphical model has parameters that have to be optimized in order to improve the classification performance of the model on both the training and test data. The learning algorithm uses the training data in order to determine an appropriate set of model parameter values. In the context of PGMs, the model parameters can be the weights which are found in logistic regression problems for example. In logistic regression problems, the probability distribution is defined to be proportional to an exponential function such that $P \sim e^{\mathbf{w} \cdot \mathbf{x}}$.

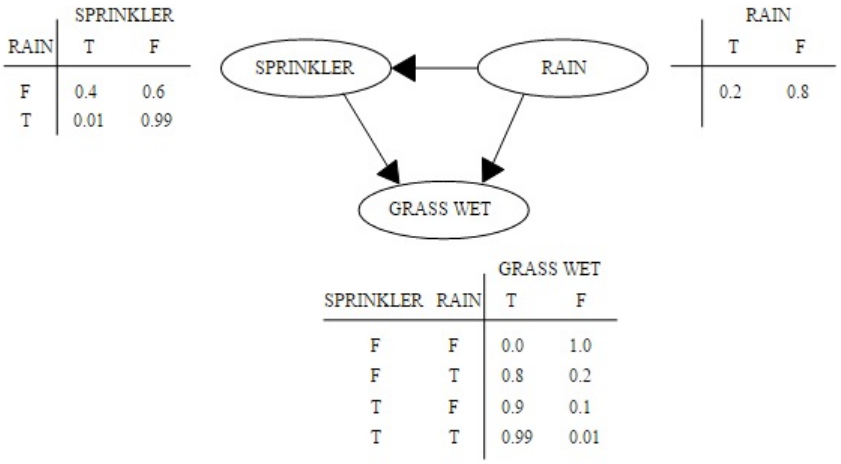


Figure 2.20: Probabilistic graphical model and associated disjoint conditional probability distributions.

Source: Bayesian network [Wikipedia Bayesian network 2016].

The parameter estimation process is then posed as a non-convex optimization problem, in which the objective function to be optimized is the likelihood function. There are two sets of parameters: the one set represents parameters within one node of the graph, and the other set represents parameters between two nodes. The former set ensures that the local factors or conditional probability functions have the appropriate distribution over the allowed values of the variables that determine the scope of the function.[18]

4.3.5 Model Querying and Inference

Querying the model entails obtaining the likelihood distribution over the set of output classes for a given data sample. The query represents a question about the likelihood of an event. The exact inference algorithm used to determine the likelihood and the class with the highest likelihood depends on the PGM used. In the context of the HCRF, the objective is to obtain the class label y that maximizes the likelihood, given the model parameters and the sequence of sample data. This is further discussed in Section 3.1 of Chapter 5. In the context of the LDCRF, the objective is to obtain the vector of class labels $\mathbf{y}$ that maximizes the likelihood, given the model parameters and the sequence of sample data. The vector $\mathbf{y}$ represents the sequence of labels, with each label corresponding to each sample in the sequence of sample data. The elements of the vector $\mathbf{y}$ are obtained by finding the value of the element $y_i \in \mathbf{y}$ such that the likelihood is maximized for the sequence element i in the sequence of sample data.

[18]For example, the variables that determine the scope of the conditional distribution function in the *grass wet* node of Figure 2.20 are sprinkler status and rain status.

5 Discussion and Conclusion

This chapter introduced the background theory used as the motivation for this work. The first section introduced the neuroscience theory. In particular, the computational neuroscience perspective was extensively discussed. Computational neuroscience allows us to think about the brain activity in terms of information processing. This perspective justifies our approach to creating a BCI based on cortical activity. The idea being that cortical activity, as measured via EEG, is related to the mental task that is performed by the user of the BCI. With this idea, our task is then reduced to identifying different, but repeatable activation patterns. The activation patterns can then be associated with specific sets of instructions that a computer has to execute.

The activation patterns exhibit complex spatio-temporal dynamics. In order to extract useful information from the EEG data, we transform the representation from the temporal domain into the frequency domain. This is achieved by employing the FFT algorithm from signal processing theory (see Section 2.13). The resulting frequency information also exhibits complex structure. In order to get around the complexity associated with attempting to mathematically model the dynamics of the signal, we recast the problem into a machine learning framework.

Machine learning provides the tools for automated association between certain input data and certain output data (see Section 4). We use machine learning to associate the frequency information with the most likely mental task. The output from the machine learning model can then be used to select the appropriate set of instructions that the computer should execute.

In this chapter, we discussed the probabilistic graphical models on which we report experiment results. A motivation for selecting these kinds of models was also outlined in Section 4.3.2 and Section 4.3.1. Three other topics pertaining to how the models are defined and used to obtain the classification results are developed, in which the following are discussed model representation and properties (see Section 4.3.3), model parameter estimation (see Section 4.3.4), and model querying and inference (see Section 4.3.5). Further discussion concerning the mathematical details of the models is presented in Section 3 of Chapter 5.

The next chapter provides an outline of the research from other researchers who have performed related work. The chapter discusses the problems that were addressed in the studies and the findings in those studies. Application domains in the related work primarily include: text input, robot control, and rehabilitation BCIs.

Chapter 3

Related Research

1 Introduction

There have been significant studies done in the BCI domain, most of which are geared towards creating assistive technology. Another application domain is gaming, which is a topic not discussed in this work. Research output continues to grow as research level BCI challenges such as the BCI Competition are introduced.[1] The BCI Competition facilitates EEG data sharing for researchers working in this field. Another factor that is driving interest in the field is the emergence of low-cost EEG technologies. These technologies stand a good chance of yielding normal day-to-day use of the technology, even amongst people of modest income. In this work, we make use of both BCI Competition data and data collected using the low-cost Emotiv EPOC EEG technology.

The reviewed studies in the field aim to address various challenges associated with creating BCI technology. The particular context of the problem considered in each of the studies determines the technology and the techniques used to address it. However, in the case of the BCI Competition, the EEG data sources and the problems are defined by the organizers of the competition. The researchers are required to submit solutions typically based on machine learning techniques which solve the required problem of the competition. The researchers with the best performing solution are deemed winners of the competition.

In this chapter, a broad outline of related research is presented. There are three main themes discussed, all of which are within the context of BCI technology. The first of these is EEG classification, in which a review of some of the work purely focused on classification of signals is presented (see Section 2). This is then followed by the text input application domain (see Section 3.1). A review of the rehabilitation and robot control application domain is presented last (Section 3.2). Most of the research is based on clinical EEG equipment, with an electrode montage which almost always includes the C3 and C4 channels. Pfurtscheller et al. [2005] reports that separability between left and right hand motor imagery is best performed with signals recorded from electrode positions C3 and C4. The exceptions are work done by Fok et al. [2011] and Roula et al. [2012], both of which make use of the EPOC, albeit with different objectives from ours.

[1] http://bbci.de/competition/

2 EEG Classification

One of the pertinent problems in BCI technology concerns the classification of EEG activity. EEG classification entails associating a set of samples of EEG measurements with some abstract label, such as a class label. The purpose of classifying EEG data is to create an association between the execution of a mental task, and the execution of a set of computer instructions. When a classification result has been determined, a set of predetermined instructions can be invoked. The execution of the instructions represents what the user intends to do. EEG classification is of paramount importance because it provides the foundation for BCI technology. There is a large amount of work that has been performed to reliably associate EEG activity with the appropriate abstract categories.

Two of the main techniques used for classification are described below. The synchronous protocol EEG classification is simple to work with, both from implementation and use-case perspectives. The main drawback is that control doesn't fully reside with the user of the BCI, since the BCI determines specific intervals in which control signals are processed. The asynchronous protocol EEG classification is considerably more complex [Pfurtscheller et al. 2004], and not much work with user feedback has been done with it. One of the major and long-standing problems in the classification context is the multi-class discrimination problem. In an ideal situation, the user would associate mental tasks with different EEG activity to as many of the controls supported by the application logic of the BCI as possible. Essentially this means that given for example a robotic prothesis with eight degrees of freedom, the user would ideally want to have direct control of each of those degrees of freedom. Such access would necessitate that eight mental tasks be used, and each should have different EEG activity from the remaining seven. The challenge reduces to creating a model that can reliably discriminate accordingly. This is not possible with current technology. Pfurtscheller et al. [2004] also discusse the importance and the difficulties associated with creating BCI technology that can support the discrimination of multiple classes of EEG activity.

A large number of studies in the literature make use of the synchronous protocol. Although to a lesser degree, there also is some work that has been done with an asynchronous protocol. In the first subsection below, a review of three papers from the literature is presented, based on the synchronous protocol. The subsequent subsection provides a review of one paper that is based on an asynchronous protocol. Our focus is on the classification techniques used and the resulting classification performance. In these studies, there was no control feedback communicated to the subject using the BCI. Thus the application domain is disregarded in this discussion since it is not defined.

2.1 Synchronous EEG Classification Protocol

EEG classification that is based on synchronous protocols requires that control signals be acted upon only within well-defined intervals in time. The intervals essentially represent recurring time windows in which the BCI is active. Outside of those intervals, the BCI does not act on the incoming EEG signals. Usually there are fixed time-intervals between and within the windows, although the two types of intervals may have different lengths.

Garret [2003] reports the classification results obtained from three models: one linear and two non-linear. They were applied to the classification of spontaneous EEG during five mental

tasks. Garret [2003] reports that the results show that the non-linear models produced only slightly better classification results when compared to their linear counterpart. The linear model is based on linear discriminant analysis (LDA). The non-linear models are based on artificial neural networks (ANN), and kernel based support vector machines (SVM).

Feature selection is performed by a genetic algorithm. The features are used to represent the state information which the model acts on. The features used in the classification are not described in the paper. All the data used was collected in one day, and the power spectra were computed on that data. Given that the genetic algorithm need not provide features that admit a neurophysiological explanation, it would be interesting to investigate the temporal evolution of the features, as this has consequences for inter-day operability of the BCI.

The LDA approach implements a clustering technique such that each mental task is associated with a cluster. The problem is reduced to determining the parameters: prior probabilities, mean vector, and covariance matrix. The model and its parameters describe the clusters taking into consideration the given training data. A data sample is then assigned to the cluster that is closest to it. The ANN is an abstraction of biological neural networks. At a conceptual level, the idea behind the ANN is to have connected nodes organized into layers. The nodes have an activation function which is applied to the output from nodes at a previous layer. The purpose of the activation function is to determine the response of the node. The response is then propagated to nodes in the next layer as input, or is used as the output of the network. In this model, each mental task is associated with a classification result. The problem is then reduced to determining the model parameters, which are scalar weights between each pair of connected nodes. The model and its parameters map the input space to the output space of classification results.

The kernel based SVM implements a classifier which uses a kernel function in a high-dimensional space which is different from the input space defined by the data. Conceptually, data from the input space is projected into a high-dimensional feature space in which the data is more likely to be linearly separable [Garret 2003]. The SVM is then applied in the feature space with the idea being that certain regions of the space correspond to certain mental tasks. The feature space is partitioned using hyper-planes equipped with support vectors. The problem solved by a SVM is the maximization of the margins between the hyper-planes and their support vectors. The constraint is that support vectors should partition the space in which the data is represented.

The performance of the classifiers was based on five mental tasks: resting, multiplying numbers, visualizing a sequence of numbers, rotating a three-dimensional figure, and composing a letter. The presentation of the results makes it unclear from which subject the results were obtained. For the LDA, the classification results varied from 47.3% to 51.1% across the mental tasks. Classification results from the ANN varied from 47.3% to 64.4% across the mental tasks. In the case of SVM, the classification results varied from 44.5% to 59.4% across the mental tasks. The largest differences between linear and nonlinear methods occurred for the resting and rotation tasks, suggesting that EEG measurements for these tasks are more difficult to distinguish than for other tasks [Garret 2003].

Pfurtscheller et al. [2005] report the results obtained from four models. The tests involved both single and multi-channel analysis of the data. The classification was performed on data with four classes, and the data was obtained from five individuals [Pfurtscheller et al. 2005]. There are some similarities with the work by Garret [2003] as far as the chosen

models is concerned. However, there are differences in the methodology followed. Pfurtscheller et al. [2005] model the EEG signal as an auto-regressive process of order three with adaptive parameters chosen using a Kalman filter. Five different classification approaches were applied and compared with each other [Pfurtscheller et al. 2005]. The classification approaches are: single-channel minimum distance analysis MDA, MDA based on the three best channels, LDA using 60 channels, SVM using 60 channels, and k nearest neighbours (kNN) using 60 channels.

In the single-channel analysis, the relative importance of each electrode position is evaluated. The dominance of electrodes in the single trial analyses that overlay the sensorimotor and pre-motor areas, confirms the modulation of sensorimotor rhythms during motor imagery [Pfurtscheller et al. 2005]. In the results presented by Pfurtscheller et al. [2005], the sets of important channels are different across the users. For one of the subjects, classification accuracy was better with the use of a single channel, as opposed to three channels. These discrepancies are not accounted for in the paper. It is also noted by Pfurtscheller et al. [2005] that in general, adding more channels does improve classification accuracy on some mental tasks.

The basic principle behind LDA is to find the best discriminating projection direction so that the distance between the classes is maximized, while the distance within a class is minimized [Pfurtscheller et al. 2005]. The minimum distance analysis (MDA) can be based on the Mahalanobis distance measure as done by Pfurtscheller et al. [2005]. It was assumed that each class can be represented by a Gaussian distribution with mean μ_c and covariance $\sum_c$. The mean μ_c and the covariance $\sum_c$ define the multivariate normal probability density function that corresponds to class c. Any point in the n-dimensional feature space can be associated with a certain distance to each class c. For each point x in the n-dimensional feature space, a distance to each class c is obtained. Each x is assigned to the class with the smallest distance to x. In the kNN classifier method, a sample is assigned to the class which is most frequently represented among the k nearest training samples. The nearest neighbors are determined by calculating the Euclidean distance function between those samples. The SVM has been described in the review of Garret [2003] above.

In the single-channel analysis case, the overall classification accuracy ranged from 46.5% to 56.9% across the subjects. The results obtained using the best three channels produced similar results, and the overall classification accuracy ranged from 38.5% to 66.6%. The last set of results was obtained using 60 channels on all the classification models. The SVM consistently produced the best results, which ranged from 52.4% to 77.2% overall classification accuracy across the subjects. The kNN produced the worst results, which ranged from 32.5% to 46.5% overall classification accuracy across the subjects. The MDA technique was used for obtaining the first two sets of results. The mental tasks in the study were the imagined movement of the left hand, right hand, foot, and tongue [Pfurtscheller et al. 2005].

Saa and Çetin [2012] report a BCI based on the classification of two imaginary motor tasks using the hidden conditional random fields (HCRF) model. HCRFs are discriminative graphical models that are attractive for this problem because they can model the temporal structure of EEG. They include latent variables that can be used to model different brain states in the signal. The HCRF also circumvents some of the limitations of generative models by learning the conditional distribution associated with the classification task. A detailed description of the HCRF is provided in Section 4.2 of Chapter 6.

Saa and Çetin [2012] take the view that changes in the power of the signals during execution of motor tasks reflect the underlying states in the brain, and that the sequence of states provides

useful information for discrimination of different imaginary motor tasks. This motivated the adoption of latent variables because the underlying sequence of states is not observed.

In the study, the common spatial pattern (CSP) technique and autoregressive modeling are applied on the EEG data. These techniques are applied after the EEG data had been processed to reduce the interference caused by electrooculographic artifacts. The reduction of the electrooculographic artifacts was facilitated by the use of three reference electrodes placed on the face such that there is an electrode on either side of each eye, and there is one electrode in-between the eyebrows. The purpose of the CSP technique is to separate a multivariate signal into additive subcomponents which have maximum variance between two observation windows. The autoregressive modeling of the CSP components produces the feature vector that is provided to the HCRF-based classifier. The feature vector is obtained by computing the power spectrum, and selecting the appropriate frequency bands using neurophysiological insights [Saa and Çetin 2012].

Saa and Çetin [2012] avoid the personalization methodology, in which frequency bands are selected on a subject-dependent basis. This contrasts the approach adopted by other researchers such as Pfurtscheller et al. [2003]. Saa and Çetin [2012] motivate this by pointing out that although personalization could lead to higher accuracy, the use of common frequency bands for all subjects makes their approach more general. Saa and Çetin [2012] justify this stance by further pointing out that the obtained performance shows the robustness of the method. However, a model trained with subject-specific data was used for each subject.

On the training data, Saa and Çetin [2012] applied the cross validation technique to determine how well the model learnt the structure in the data. The results range from 68% to 99% classification accuracy across nine subjects. On the test data, the model produced results ranging from 11% to 100% classification accuracy across nine subjects on one test data-set. Saa and Çetin [2012] also report 0% to 96% classification accuracy across nine subjects on another test data-set. In both the test data contexts, the average classification accuracy across all subjects is 66%.

Saa and Çetin [2012] conclude by asserting that the method used is based on modeling the temporal changes of the EEG signal, and the analysis of the state sequences could provide insights into the physical phenomena underlying the execution of the imaginary motor tasks. It is further indicated that this is part of future work.

2.2 Asynchronous EEG Classification Protocol

Asynchronous protocols have not been studied extensively, owing to their complexity. In this subsection, a review of a study that is based on an asynchronous protocol is presented. When asynchronous protocols are used, the BCI is active at all times. This essentially means that there is a single window that is as long as the usage session, and the BCI acts upon the control signals that are related to the continual EEG signal stream. The model used in the study serves both to label and segment the data. This means that the model is capable of representing both the characteristics which are intrinsic to a particular class and extrinsic to that class, thereby facilitating the representation of inter-class transition characteristics.

Saa and Çetin [2013] describe how the asynchronous BCI problem can be posed as a classification problem based on conditional random fields (CRF) or latent dynamic conditional random

fields (LDCRF).[2] This is done by defining appropriate random variables and their relationships. The LDCRF incorporates latent variables that permit modeling the intrinsic structure for each class and at the same time allows modeling extrinsic dynamics [Saa and Çetin 2013]. These characteristics allow the LDCRF model to be directly applied for labeling unsegmented sequences. The theoretical stance assumed by Saa and Çetin [2013] is the same as that assumed by Saa and Çetin [2012]. The main difference being in the modeling of the dynamical effects between the different classes. One of the fundamental problems in the asynchronous protocol domain is that the subjects execute different mental tasks without cues, and this means that it is unknown when the subjects start the execution of a specific task.

CRFs can model the extrinsic dynamics of the data, which in asynchronous BCI corresponds to dynamics across different tasks. However, they lack the ability to model intrinsic dynamics (i.e. the temporal evolution in the course of execution of a particular task [Saa and Çetin 2012]). HCRFs have been used for synchronous BCI by Saa and Çetin [2012]. The HCRF takes into consideration the dynamics of the signal during the execution of one task. However, it assigns a unique class label to an entire segment of EEG signals. The problem with such an approach is that the selection of a fixed and appropriately sized window for data segments that are presented to the model is necessitated. In our work, the window size was determined empirically. The LDCRF circumvents this calibration problem by directly working out the segment length without depending on a fixed-width window.

Feature selection was performed using the sequential floating forward selection algorithm [Saa and Çetin 2013]. The algorithm is applied on a set of features which are non-overlapping bands of the power spectrum. This algorithm determines the best set of features by systematic elimination of inferior features. The starting point is an initial given set of features. Features are eliminated if they either decrease or have no effect on the classification accuracy. The objective is to reduce the number of elements that are used to denote the current state because an unnecessarily large set of features entails computational cost [Saa and Çetin 2013].

There were three participants in the study, and the classification results reported are based on three classes: imagined left hand movements, imagined right hand movements, and the generation of words beginning with the same random letter. The classification accuracy on the training data ranges from 59.3% to 91.55% on the LDCRF and 59.73% to 89.34% on the CRF across the subjects. The classification accuracy on the test data ranges from 72.36% to 95.63% on the LDCRF and 61.81% to 92.95% on the CRF across the subjects. On a separate data-set, the classification accuracy results ranged from 62.76% to 81.31% and 54.85% to 80.75 using the CRF and LDCRF across the subjects, respectively. The mean classification accuracy across the subjects is reported as being 69.53% and 67.47% on the LDCRF and CRF, respectively.

3 Application Domains

The literature reviewed in this section features an application domain in which the performance of the BCI is assessed. Two application domains of interest are discussed. The first domain largely pertains to communication, in the context of text input. The second domain largely

[2]The CRF is a counterpart of the HCRF. The main difference between the two is that the latter can model intermediate structure that is not explicitly annotated in the data, while the former cannot.

pertains to mobility, in the context of controlling a wheelchair and regaining motor control through orthotic device assisted recovery.

3.1 Text Input

Text input is one of the fundamental building blocks of modern digital communication. The loss or the degradation of the neuromuscular connection presents itself as an impediment to using digital communication technology. There is a substantial amount of work that has been devoted to the development of text input systems for people with certain disabilities. In this section, two papers that report work done on BCI for text input are reviewed. The first paper is based on a synchronous protocol [Roula et al. 2012], and the second paper is based on an asynchronous protocol [Pfurtscheller et al. 2003]. In both studies, the users are provided with control feedback.

Roula et al. [2012] report the classification accuracy and the text input rate achieved using an EEG based BCI. The BCI implements a hybrid architecture using motor-imagery and event related P300 signals.[3] In the study, the Emotiv EPOC headset is used to measure both the P300 signals and motor-imagery linked de-synchronization in the EEG data. In order obtain control information (such as selecting a particular key) corresponding to P300 signals, on-screen control elements flash for a specific time duration. If the user is looking at the flashing control element, then the corresponding control is executed. Control information corresponding to motor-imagery is obtained by measuring the changes in certain frequency bands in response to the user imagining a particular motor-task. The BCI reported was used for text input by using both signals in order to take advantage of the two modalities.

Unlike motor-imagery signals, the main limitation with using only P300 signals is that the resulting speed of typing averages 22 seconds for each character [Roula et al. 2012], which is clearly undesirable. However, there is a trade off between speed and accuracy. Unlike P300 signals, the main limitation with using motor-imagery signals is that they produce less accurate results, especially in character selection tasks [Roula et al. 2012].

In the study, the wavelet transform was used for feature extraction. The computed wavelet coefficients were used as feature vector components. It is reported that only the C3 and C4 channels have been used to provide data [Roula et al. 2012]. This is an unusual setup since the fixed montage of the Emotiv EPOC does not feature these channels [Emotiv 2014]. Three classes were used for the classification step, for which the Mahalanobis distance was used to associate feature vectors to clusters formed by mean vectors [Roula et al. 2012]. The three motor-imagery tasks with which the model is trained are imagined left arm movement, imagined right arm movement, and a rest state in which no movement is imagined.

As a performance benchmark, sentences of the following character length were used: 3, 8, 21, 25, and 29. The P300 system is reported to produce higher accuracy across all sentence lengths (85% at 25 characters), followed by the hybrid (70% at 25 characters), and finally by

[3] P300 signals are EEG measurable responses to a certain external stimulus such as a flashing light. They appear in EEG measurements approximately 300 milliseconds after the stimulus is presented to the subject. Typically the stimulus is a flashing light and the response is measured over the occipital region (see Figure 2.1) of the brain.

the motor-imagery system (70% at 25 characters) [Roula et al. 2012]. All systems produced 100% classification accuracy on 3 characters. The time durations required for the described set of classification accuracies were: 595 seconds for the P300 system, 306 seconds for the hybrid system, 1 427 seconds for the motor-imagery system. Although the task of text input is successfully performed, the reported latency makes using the system a slow process.

Pfurtscheller et al. [2003] report the text input rate measured such that errors are also corrected for by the user, the result being error free output. The classification problem was addressed using an asynchronous protocol. Amongst a set of mental tasks, the two most discriminable were selected for each subject individually. The motor imagery tasks were: imagined left hand movement, imagined right hand movement, imagined movement of both feet, and imagined movement of the tongue. One purely cognitive task was an arithmetic task [Pfurtscheller et al. 2003]. Amongst the aforementioned mental tasks, only two were selected for each user and the resulting classification problem had only two classes. The control model is based on selecting one of two possible control options that appear on the screen. Text input is achieved by successively selecting control options that are presented in manner analogous to tree traversal.[4]

The features used for classification were the band-power estimates in subject specific EEG bands. The classification model used is the hidden Markov model (HMM). The classification was arranged such that two separate HMMs were used and with each HMM trained to learn the characteristics of the EEG activity of one mental task. The final classification result was obtained by taking the maximum of the two likelihoods returned by the HMMs. The classification result is attributed to the HMM with the highest likelihood value [Pfurtscheller et al. 2003].

Three subjects took part in the study [Pfurtscheller et al. 2003]. It is interesting to note that for each subject in the study, an appropriate set of channels, mental tasks, and EEG bands had to be determined. This is a good step towards personalization, but is also non-trivial in the context of BCI problems. Pfurtscheller et al. [2003] reports an error-free text input rate which ranges between 0.5 to 0.85 letters per minute across the subjects. This spelling rate results in a slow rate of text input, although when viewed in light of current result, .

In a follow-up study, Pfurtscheller et al. [2004] noted an improved text input rate by redesigning the input method. Pfurtscheller et al. [2004] presents an asynchronously controlled three-class brain-computer interface-based spelling device virtual keyboard that is operated by spontaneous electroencephalogram and modulated by motor imagery. Concerning the results of three able-bodied subjects operating the virtual keyboard, two were successful and which showed an improvement of the spelling rate. The number of correctly spelled letters per minute obtained went up to 3.38, with a mean of 1.99. However, it is further noted that one of the participants could not correctly spell a single word. Additionally, no consistent spelling performance could be achieved [Pfurtscheller et al. 2004].

Pfurtscheller et al. [2004] accounts for the observed problems by pointing out that the visual input in the form of control feedback has a strong impact on motor cortex activity, and can lead to a deterioration or changing of the motor imagery related EEG patterns. Pfurtscheller

[4]Using graph theory concepts, the specification of a sequence of actions can be achieved by organizing it like a tree data structure.

et al. [2004] states that more research on this effect is necessary. Another problem that is not typically discussed in the other studies is the effect of providing control feedback to the user. Pfurtscheller et al. [2004] states that multiple classes and asynchronous control can limit the usability of the system, and that users do require more training and the cognitive load is higher.

3.2 Rehabilitation and Robot Control

BCI technology has application potential in rehabilitation, especially if applied in a physiotherapy context in which BCI-based control of an orthotic device can be used to strengthen neuro-muscular connections and increase muscle tone. Such technology could be a significant improvement in early-stage rehabilitation. In a study conducted by Fok et al. [2011], an orthotic hand coupled with a BCI based on the Emotiv EPOC headset is described. In another study by Iturrate et al. [2009], an EEG brain-actuated wheelchair with automated navigation is presented. Research devoted to this kind of wheelchair technology is becoming wide-spread. However, we limit our discussion to the work described in Iturrate et al. [2009] as it is representative of this work.

Fok et al. [2011] report an orthotic hand that is coupled with a BCI. The purpose of the reported BCI is to assist a user in opening and closing the user's hand. The main concept in the study is the use of motor-imagery activity that emanates from an ipsilateral section of the brain with respect to the would-be affected side of the body. This is important since it is the contralateral side of the brain that is often affected in stroke survivors, thus making it difficult to obtain reliable EEG signals from that side of the brain. Fok et al. [2011] proposes the adoption of the presented technique as an alternative scheme for rehabilitation and restoration of hand control for stroke survivors. This is important when one considers that the potential for recovery is unhampered by the severity of neural pathway injury since we circumvent the entire injured pathway [Fok et al. 2011]. The hope is that with the help of BCIs the remaining contra-lesional areas of the brain can be trained to take over motor control of the impaired hand [Fok et al. 2011].

It is not clear which algorithm was used for signal processing. However, the authors mention that the BCI 2000 framework was used.[5] This is a vague statement since the BCI 2000 framework does not perform a unique function for signal processing purposes. What is clear however is that the power spectrum is computed and used as the input signal. In addition, there is no application of a transparent classification scheme. The transformed signal obtained from EEG data is applied to dynamically setting the gain parameters using the least means square adaptive filter. The gain parameters are then used by some unspecified classifier to produce output which sets the position of the linear actuator that drives the orthotic hand.

Fok et al. [2011] report that the classifier produced an 81.3% success rate for this task. It is also noted that the metric used for performance measurement is not specified. Fok et al. [2011] does however provide a plot indicating how well the system could track the target position (see Figure 3.1). Through visual inspection, it is clear that the system does perform reasonably well. What should also be kept in mind is that the actuator produces a linear and not a step-wise response. So the target and actual curves must not be expected to coincide at

[5] http://www.schalklab.org/research/bci2000

all times. Additionally, the actual curve must not be expected to have regions in which the gradient approaches infinity. That is to say, set-points cannot be attained instantaneously using their mechanism. Therefore, there will be unavoidable discrepancy between the curves.

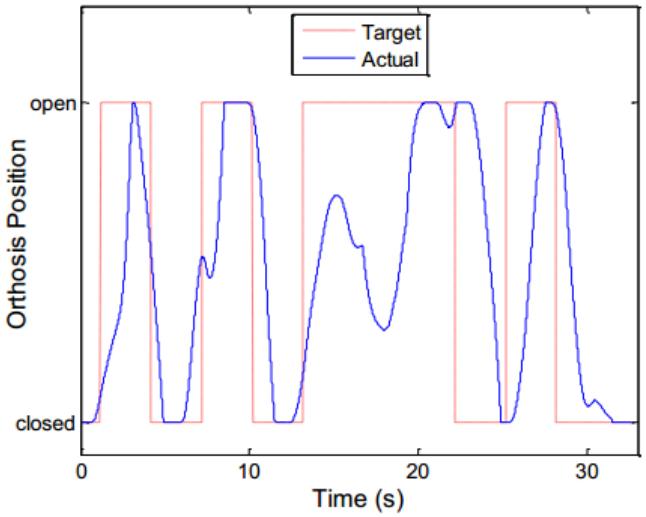


Figure 3.1: Orthosis position actuation.

Source: An EEG-based Brain Computer Interface for Rehabilitation [Fok et al. 2011].

Fok et al. [2011] concludes by recognizing some of the challenges that remain to be addressed. Future development pertaining to expanding the system's ability to adapt to spatially non-stationary signals is required. Implementing adaptive spatial filters or an adaptive classifier would determine the strongest correlated channel automatically and continuously would improve robustness for a long-term outpatient orthosis [Fok et al. 2011]. More importantly, spatial and temporal filters that remove artifacts are essential to the device performance outside of a research setting. The artifact sources considered are induced by eye blinks, EMG, and breathing.

Iturrate et al. [2009] describe a new non-invasive brain-actuated wheelchair that relies on a P300 signal protocol and automated navigation. When the system is in use, the user faces a screen with a real-time virtual reconstruction of the scenario and concentrates on the area of the space to reach [Iturrate et al. 2009]. A visual stimulation process elicits the P300 response and the EEG signal processing detects the target area. The identified target area represents a location that is given to the autonomous navigation system. The autonomous navigation system drives the wheelchair to the desired place while avoiding collisions with the obstacles detected by the laser scanner. The classification accuracy of the brain-computer interface is above 94%. In addition, the flexibility of the sensor-based motion system allows for navigation in non-prepared and populated scenarios. The study involved five healthy participants, and all the participants were able to successfully use the device. The participants are reported to have used the system with relative ease [Iturrate et al. 2009].

Iturrate et al. [2009] made use of step-wise linear discriminant analysis (SWLDA) for classification. The SWLDA has been extensively studied in P300 classification problems, especially in problems involving online communication using visual stimulation [Iturrate et al. 2009]. SWLDA is reported to have produced classification accuracy higher than 90% in less than an hour

of training. Two trajectories were used to test the performance of the system. Of particular interest for us is the variability analysis provided in the paper. It has been shown in numerous studies that there is significant variability in the choice of the parameters used, as can be seen in Pfurtscheller et al. [2003]. The parameters pertain to channel selection, EEG band selection, and the selection of the most discriminable mental tasks. Iturrate et al. [2009], however reports very little variability.

Iturrate et al. [2009] analyze two types of variability observed in the experimental sessions. First is intra-subject variability, which measures the variability of a subject among trials of the same task. Second is inter-subject variability, which measures the variability of execution among subjects during the execution of the same task [Iturrate et al. 2009]. The main objective of the analysis was to determine what Iturrate et al. [2009] describes as the homogeneity of the system. In more general terms, it is to test whether a homogeneous group of participants offers similar results under similar experimental conditions.

Iturrate et al. [2009] apply the Pearson correlation coefficient test on various metrics. It is reported with some exception that the intra-subject variability is not substantial (0.94 Pearson correlation coefficient). Iturrate et al. [2009] asserts that the low intra-subject variability indicates that the subject tried to perform the task in a similar way in both executions. Furthermore, low inter-subject variability (0.92 Pearson correlation coefficient) is also reported, and the low variability is justified as indicating that the users executed the task in a similar and analogous way [Iturrate et al. 2009]. It is not clear what is meant by performing the task in a similar way. However, Iturrate et al. [2009] asserts that the results and the intra-variability allow us to infer that under the same experimental conditions, the group performs similar actions and gives the system a high degree of homogeneity and invariability against these situations. The reported level of homogeneity is not found in the other studies reviewed in this work, and it raises interesting questions about the variations that other researchers report. It is possible that the homogeneity is an effect of the neurophysiological invariance which characterizes P300 signals.

4 Discussion and Conclusion

This chapter discussed the related research. The discussion entailed a broad outline concerning the current state of the field. It was noted that some of the factors facilitating research in the field are the emergence of low-cost EEG technology and the introduction of research-level competitions based on BCI problems. In subsequent sections, the general activities in the field were discussed and it was shown that the application domains range from text input to brain-actuated wheelchair operation. It is also made clear that there are still significant challenges that remain to be addressed systematically.

In Section 2, the general problem concerning the classification of EEG data was outlined. At present, the problem is being addressed by adopting the synchronous or asynchronous protocol. The former is more commonly used owing to its relative simplicity. Concerning the synchronous protocol, three studies which made use various techniques to address the problems were reviewed. Some success was reported by the researchers. Saa and Çetin [2012] reported the best classification results. Concerning the asynchronous protocol, only one paper was reviewed. Saa and Çetin [2013] apply the LDCRF to perform data segmentation and labeling of unsegmented EEG data. The technique adopted is an asynchronous BCI protocol. The performance of the LDCRF is noted as being not substantially different from that of the simpler CRF, which

is somewhat counter-intuitive given the dynamic nature of the data. It could however indicate that most of the important information is related to transition dynamics between different mental tasks.

In Section 3.1, two studies involving text input as application domains were discussed, Roula et al. [2013] and Pfurtscheller et al. [2003] respectively. The first study was based on the Emotiv EPOC headset and a synchronous protocol. The BCI made use of both motor and P300 signals for control. Although the system is functional, the text input rate is very low. The average typing rate is reported as being 22 seconds per character. The second study was based on clinical-grade EEG technology and an asynchronous protocol. The BCI used motor imagery for control. The performance of the BCI could not be consistently assessed, and one of the three participants could not even type a single word using the BCI.

In Section 3.2, a discussion concerning the application of BCI technology to rehabilitation [Fok et al. 2011] and robot control [Iturrate et al. 2009] is presented. First, Fok et al. [2011] present their work as an alternative treatment programme technology for stroke survivors. The Emotiv EPOC measured EEG signals that were used to actuate an orthotic hand. Second, a brain-actuated wheelchair with autonomous navigation capability is presented by Iturrate et al. [2009]. All the participants being able to control the wheelchair to complete the challenge of tracing out the test trajectories.

In conclusion, it appears that there is still a significant amount of work that still has to be done in order to formalize the approach to BCI problems. Text input application domains are still problematic. We find that either the BCI functions properly, but is very slow, or that it doesn't function properly at all. Neither case is desirable. There is uncertainty over the procedure to follow in determining the most discriminable mental tasks, channels, and EEG bands. Despite the views and findings of Saa and Çetin [2012] and Iturrate et al. [2009], it is clear that in order to improve the performance of BCI technology, there should be a great deal of personalization in the operation of BCI technology. The personalization techniques that were reviewed are all empirically based. Perhaps it is a matter of speculation whether a unified framework for determining the parameters could be created. However, a more structured and simpler approach is required to make BCI technology useful outside research laboratories. Scope for further development of the field was discussed in most of the papers. Topics that have not received much attention from researchers excluding Pfurtscheller et al. [2004], are the effects of control feedback and cognitive load when multiple classes are to be discriminated.

In the next chapter, the research method used in this work is outlined. The chapter provides a detailed discussion concerning the framework used to perform this work. The discussion entails stating the research hypothesis and providing motivation of why it is expected to hold true. The discussion proceeds to the software components that were required to test the hypothesis. The software components were designed and implemented with a modular architecture in mind. A modular architecture is versatile and facilitates relatively simple component reuse or replacement. The chapter ends with a review of the low-cost experimental equipment used in the research.

Chapter 4

Research Framework Overview

1 Introduction

This chapter discusses the experiment framework under which this research was conducted. In Section 2, the motivation for initiating the work is presented and is summarized as follows. Although the search for a solution that would alleviate motor disabilities is interesting from both scientific and engineering perspectives, the impact of these disabilities are both substantial and negative for society at large and especially so for afflicted individuals. The work we present is a step towards alleviating motor disability related problems in a manner that is accessible and affordable to the general population.

Section 3 presents our research hypothesis, as well as a brief motivation concerning the feasibility of achieving the objective put forward in the hypothesis. Contrast is also drawn between what this study aims to achieve and what other researchers with similar objectives have achieved. Of significance to our approach is the use of graphical models in conjunction with low-cost EEG sensor technology, which constitutes the main contribution of this work.

In Section 4, the method followed in order to carry out the research is presented. This begins with a discussion of the design of the system (see Section 4.1.1) and the software components which have been implemented (see Section 4.1.2). The software components provide complementary functionality. Some of the components are critical to the functioning of the BCI, , whilst others are used for assessing the performance of the system and for determining appropriate parameter values in the system components. The subsequent subsections describe the data collection methodology (see Section 4.2) and how the performance of the system is evaluated (see Section 4.3). Section 5 presents a discussion of the experimental equipment used. Various other topics which include a basic description of the sensor technology are also discussed. In particular, a critical review of the chosen low-cost sensor technology is provided (see Section 5.1).

2 Research Motivation

At the fundamental level, the motivation is to address one of the basic human needs. Our focus is on the ability to communicate and be part of society after motor control has been lost. There are two basic ways in which people lose motor control. The first is through degeneration of

tissue owing to the manifestation of some neuro-degenerative ailment. This may result from the transmission of some genetic condition from parents to offspring [Purves et al., 2004]. The second is through injury, which is a society-wide problem since it transcends both genetic and behavioural parameters. In this sense injury recognizes neither genes, age, nor lack of recklessness.

Collectively, these conditions affect a large number of people. In 2013 there were approximately 2 861 028 people with disabilities above the age of four years old in South Africa [Statssa 2013]. The associated socio-economic impact includes loss of income for individual families, and a substantial government expenditure on disability grants (R17 450 371 800 per annum in South Africa as measured in August 2014 [Sassa 2014]). If some of the problems could be alleviated, the government expenditure could be channeled towards capital expenditure for further advancement of society. As discussed in Section 2 of Chapter 1, we propose an asynchronous BCI based on low-cost EEG technology. Our approach is based on the use of machine learning techniques for extracting user intention from the EEG data. In the next section we present the research hypothesis we aim to test.

3 Research Hypothesis

Formally stated, our hypothesis is that: *At least 3 independent mental tasks can be asynchronously and reliably classified using the Emotiv* EPOC *headset if feature vectors are extracted using frequency decomposition techniques, such as FFT, and classified using probabilistic graphical models, such as HCRF and LDCRF.* In the subsection below, a brief motivation of the feasibility of our hypothesis is provided.[1]

Various studies have indicated that, using certain techniques from both signal processing theory and machine learning, it is possible to reliably classify 3 different classes of mental activity using EEG signals. Notable examples include:

1. **Pfurtscheller et al. [2004]:** a three class asynchronous BCI based on clinical-grade EEG technology. The mean spelling rate of 1.99 letters per minute was obtained from two participants, although there were three participants in the study. One of the participants could not spell a single word using the system. Linear discriminant analysis was used for classification.

2. **Roula et al. [2012]:** a three class synchronous BCI based on the Emotiv EPOC EEG headset. However, substantial command execution latency is reported in this work. Three characters are spelled with 100% accuracy in 321 seconds. The performance was assessed on two participants in the study. A Mahalanobis distance based technique was used for classification.

3. **Saa and Çetin [2012]:** a two class synchronous BCI based on clinical-grade EEG technology. The classification performance is reasonable, with mean classification accuracy of 66-67% on nine participants. The classification technique was the HCRF.

It is noted that above listed researchers have reported reasonable performance of their BCIs,

[1]The HCRF and LDCRF are acronyms for hidden conditional random fields and latent dynamic conditional random fields, respectively. These models are further discussed in Section 3 of Chapter 5.

and this motivates our work further. However, the transformation between our work and their work is not straightforward.

First, our EEG technology is low-cost and that entails some difficulties. Some of the difficulties that were anticipated prior to the study are discussed in Section 5.1. Second, the probabilistic graphical models that we use are suitable to learning temporal sequence structure in data and they provide a compact representation of the probability distribution which facilitates efficient querying. Additionally, they are members of the discriminative family of probabilistic graphical models (PGMs), and that means that they do not require independence assumptions to be made about the variables used to describe the data. A detailed discussion about the machine learning techniques is presented in Section 3 of Chapter 5. Thus the main differences between our work and that of other researchers are along the following factors:

- **EEG sensing technology** Most of the work in the literature makes use of clinical-grade EEG sensing technology. The main advantage of using clinical-grade technology is that the quality of the data can be very high. Consequently, the theory can corroborate the findings from neuroscience and this makes assessing the effectiveness of the data processing and machine learning algorithms easier.[2]

  The main disadvantage of using clinical-grade technology is the associated cost. As previously stated, in this work we aim to make BCI technology accessible to people of average financial means. The cost factor rules out the use of clinical-grade technology for the application domain that is considered in this work. The studies which have made use of low-cost technology did not make use of models that are capable of learning temporal dynamics in the data, with examples including Fok et al. [2011] and Roula et al. [2012]. Amongst these studies, it is with regards to the machine learning techniques and consequently classification methodology that we draw a distinction to their work.

- **Machine learning techniques** There are various studies in which PGMs have been used for BCI applications. Some examples include hidden Markov models, hidden conditional random fields (HCRF), and latent dynamic conditional random fields (LDCRF). Amongst others the list includes work by Hasan and Gan [2011], Saa and Çetin [2011], Saa and Çetin [2012], Saa and Çetin [2013]. The PGMs are reported to be an appropriate fit for the classification task on EEG based BCIs. This follows from the observations that they provide compact representations of the probability distributions that they model, and that they are capable of capturing temporal structure in the data. The capabilities of the HCRF are discussed in Section 3.1 of Chapter 6. In the reviewed studies, clinical-grade EEG technology and a synchronous protocol are primarily used. Amongst these studies, it is with regards to the EEG technology and consequently data quality that we draw a distinction with their work.

4 Research Methodology

In this section, a detailed discussion of the various methodological matters is presented. The ultimate goal of our proposed methodology is to perform an empirical analysis of the performance

[2]This kind of assessment is also performed in this work, using data obtained from researchers who used clinical-grade technology, as discussed in Chapter 5. This topic is developed in detail in Section 4.1 of Chapter 6.

of the system. However, the experimental framework tools had to be designed and implemented before this goal could be achieved. The experimental framework tools form part of the deliverables of this work. The three major milestones in the research are as follows: system design and development (see Section 4.1), data collection (see Section 4.2), and system performance evaluation (see Section 4.3).

4.1 System Design and Development

One of the major problems with using an EEG BCI is the contiguous signal stream that must be processed is such that the bulk of the data corresponds to being in an idle state. Consequently, one of the major tasks is identifying this idle state. In order to facilitate the identification of the idle state, BCI protocols are separated into synchronous and asynchronous modalities.

Synchronous protocols require a perceptible cue to be generated by the system before a user can provide control signals. This means that a certain time-interval is designated as the control time-interval. By definition, all signals that fall outside of the control time frame are discarded since they are treated as noise. In this mode of operation, the BCI dictates when the user is allowed to interact with the BCI. This mode of operation reduces the appeal of usage. Although they are effective and simpler to implement, synchronous protocols do not provide a natural mode of interaction with a BCI.

An asynchronous protocol is implemented in this work. An asynchronous protocol provides a natural mode of interaction with a BCI, since the control commands are provided at a rate that is dynamically determined by the user. This mode of operation is more appealing because control resides with the user. As a result, there are no cue signals that would potentially interrupt the user when interaction with the BCI is not required. This protocol is however more complex to implement, and Pfurtscheller et al. [2004] raises two concerns. First, the visual input has a strong impact on motor cortex activity and can lead to a deterioration, or changing of the motor imagery related EEG patterns. Second, multiple classes and asynchronous control can limit the usability of the system [Pfurtscheller et al. 2004]. Taking into account that the control window is always open, more noise has to be recognized since the signal being sought may appear at any segment within a stream of noise. In the discussion below, the design and development details pertaining to the system are presented.

4.1.1 System Design Considerations

There are various software components that were developed in order to facilitate experimentation in this work. The components include visualization tools, data processing tools, and plug-in applications. The key issue is that the algorithms, their implementation, and even the specification of the auxiliary entities such as test beds may change over time. In order to minimize the negative impact of these possible changes, the following software engineering principles are employed.

The architecture that is adopted for the system is modular. Modular designs ensure that creating different implementations of a particular module does not require re-working of other modules. The connections between most of the modules are mediated by web-socket technology. This modularity follows from the fact that modules can be implemented as stand-alone entities which can even be independent of the particular details of the problem being addressed. With

a modular design, test-beds such as text input or wheelchair control can be readily augmented to form part of the system.

The schematics of the designs are provided in Figure 4.1 and Figure 4.2 and indicate that the system can be operated in two modes. The first mode is on-line, in which the system is used in real-time to generate classification results. An example application domain for this mode could be to manipulate the controls in the application logic graphical user interface (GUI). The second mode is off-line, in which the researcher uses the system to assess the performance of the command inference and consequently that of the model. In the off-line mode, the parameters of the system are also tuned in order to provide better on-line performance.

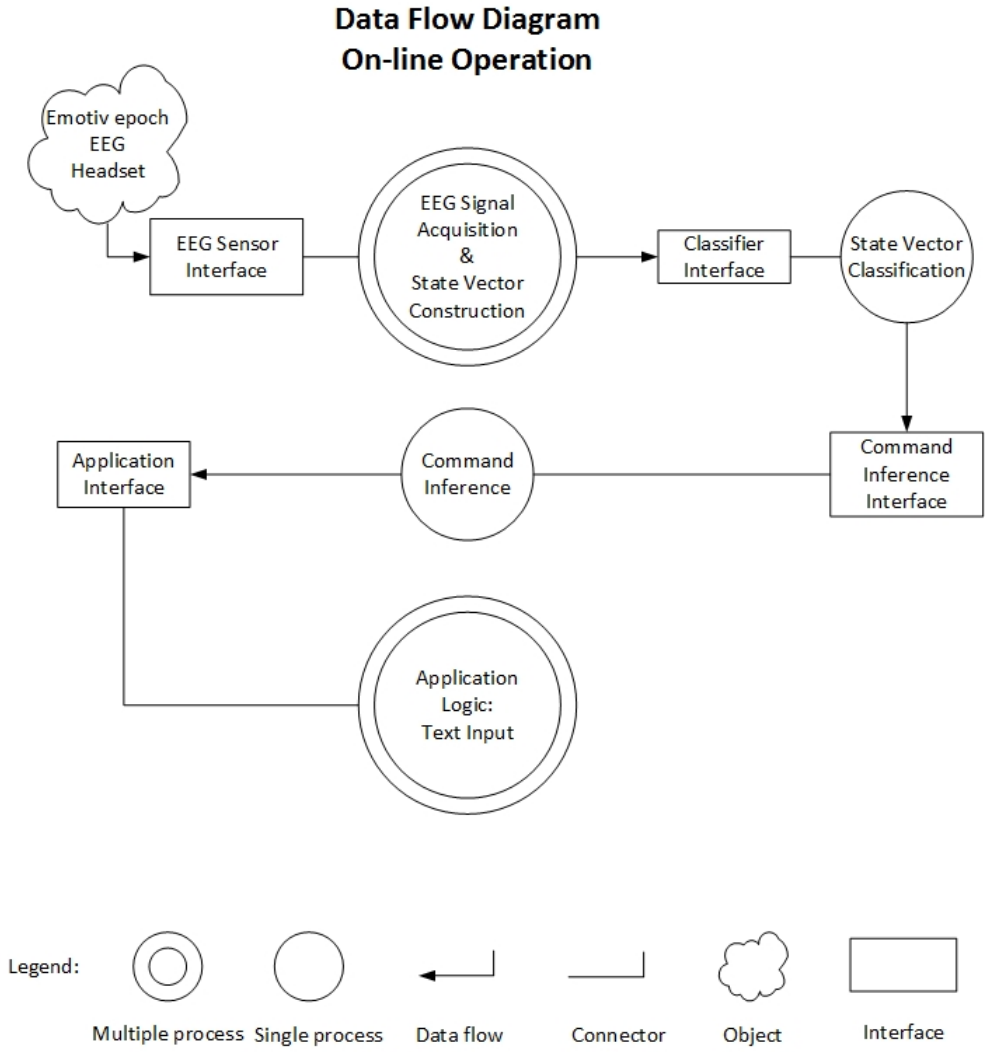


Figure 4.1: System design schematic with signal acquisition for on-line mode data-flow.

4.1.2 System Components

Deliverables in this work include functional components of the system described above. In addition to the primary components depicted in the design schematics, there are secondary components which were also developed. The secondary components complement the system,

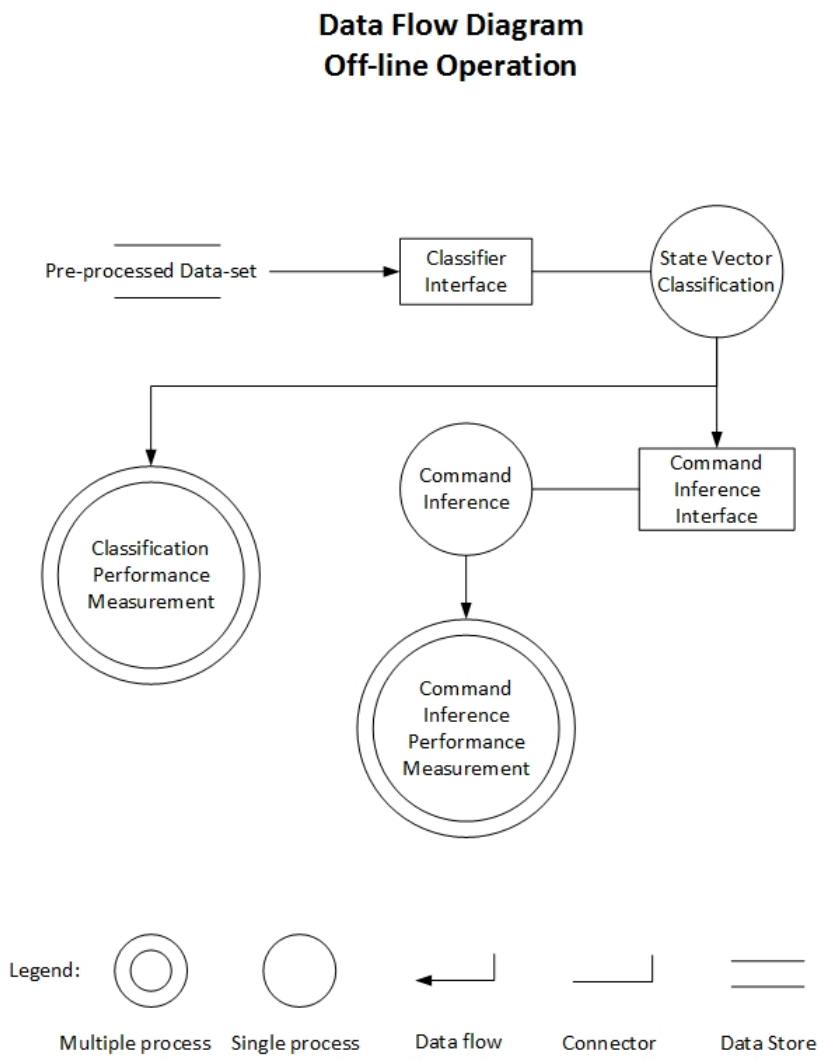


Figure 4.2: System design schematic for off-line mode data-flow.

because they provide information which can be used by the researcher to assess whether the primary components are functioning properly or not. These secondary components are in the form of visualization tools. A discussion of the primary and secondary components that have been developed is presented below.

EEG Sensor State Monitoring Module This module communicates with the EEG sensor via an EEG data stream service in order to obtain the signal and device status data. The GUI of this module as depicted in Figure 4.3 allows the user and the researcher to monitor the state of the sensor.

The circles within the boundary of the head figure represent electrodes. The meaning of the colour-code appears above the head figure. The ideal situation is when all circles have a green colour. The exception being the two white circles with a dotted area, because they represent the reference electrodes which are used for calibration purposes. When a circle has a black colour, then the signal corresponding to the electrode is corrupted. The connection referenced in the figure is the WIFI connection between the headset and the USB dongle. The sampling rate

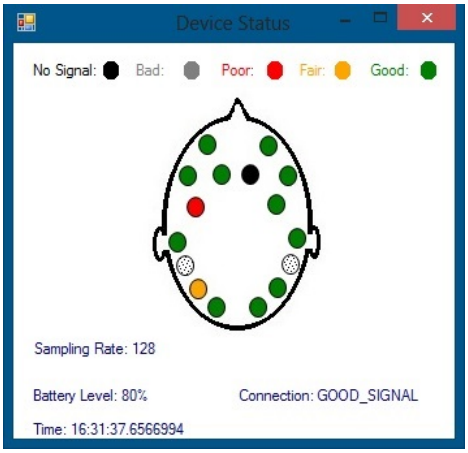


Figure 4.3: Device status monitor GUI depicting device status information.

indicates the current sampling rate of the device. For reasons unknown to us, the rate varies sporadically from time-to-time. This module corresponds to the *EEG sensor interface*, and the *EEG signal acquisition* and *state vector construction* multiple process symbols in Figure 4.1.

Frequency Decomposition and State Vector Construction Module This module transforms the input signal from the temporal to the frequency domain. The exact transform is the computation of the power spectrum of the EEG measurements. The power spectrum represents the squared magnitude of the coefficients obtained from the FFT algorithm which is described in Section 3.2.2 of Chapter 2. After the power spectrum has been computed, a specific subset of the entire spectrum is selected, as outlined in Section 2 of Chapter 5. The selected elements represent features and form the components of the state vector. This module corresponds to the *EEG signal acquisition* and *state vector construction* multiple process symbol in Figure 4.1.

State Vector Classification Module This module provides the state vector classification functionality. The input state vectors are associated with a set of abstract classes. The abstract classes correspond to labels of certain mental tasks, and can be associated with a specific set of instructions that the computer has to execute. Figure 4.4 depicts the GUI of this module. The GUI has controls that allow the researcher to perform the following tasks:

- **Set the classifier activity state** This determines whether the model is active or not. When the model is off-line, then the BCI is inactive since instruction sets to be executed cannot be invoked.[3] When the model is on-line, then the BCI is active and the classification results can be used to invoke the execution of certain instructions.
- **Train a model** This pertains to applying a learning algorithm to the underlying model until certain convergence criteria are satisfied. The learning algorithm uses the labeled training data that is provided by the researcher. The learning algorithms for the HCRF and the LDCRF are different. A short description of the algorithms is provided in Section 3.1 and Section 3.2 of Chapter 5, respectively.

[3]Note that "off-line" here has no bearing on the "off-line" referenced in Figure 4.2. A similar argument applies to "on-line".

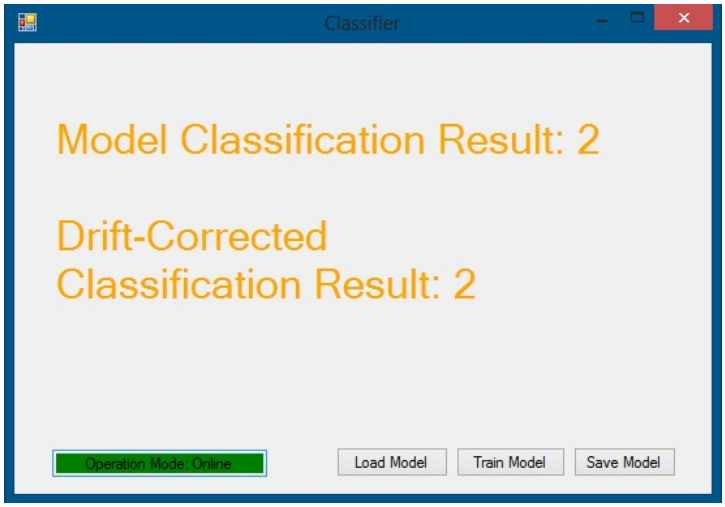


Figure 4.4: Classifier GUI depicting model classification and command inference results.

- **Load a model** This pertains to selecting and loading a model into primary memory. It is assumed that the model has been saved on secondary memory from a preceding experiment session. After the model has been loaded, it can be used for classification purposes, or it can be trained once again with potentially new data.
- **Save a model** This pertains to writing a binary file into secondary memory. The file contains all the information about the model. The model can be either new or it can be a model which has been loaded from secondary memory in the first instance. When changes are made to the model, such as perhaps when the model is retrained, then the binary file can be overwritten with the newer version of the model.

This module corresponds to the classifier interface and the state vector classification process symbols in Figure 4.1 and Figure 4.2.

Command Inference Engine This module determines which class the user is most likely trying to select. Following that, it can invoke the set of instructions that should be executed when a particular command has been determined. It is important to note that at this stage the invoked set of instructions represents invoked operating system events, and the application logic module then determines how the application responds to the events. This engine receives a contiguous stream of classification results from the state vector classification module. The purpose of this engine is to determine the classification result which best represents the intent of the user, given a set of possibly conflicting classification elements. This topic is developed further in Section 4.3.1.

This module also implements the logic for wrapping around the controls on the GUI of the text input module, which in this case represents the application logic module. The wrap around functionality is used to reduce the number of distinct control signals required to manipulate the GUI controls from five to three. This is critical when distinguishing between more than 4 mental tasks proves to be difficult. In Figure 4.8, the arrows associated with the different mental tasks allow the user to scroll along the controls. By implementing wrap around, two of the four arrows can be rendered inactive, in which case only two *scroll* (one horizontal and one vertical) and one *select* commands are required. This engine corresponds to the command inference interface and the command inference process symbols in Figure 4.1 and Figure 4.2.

Application Logic Module This module acts on the symbols emitted by the command inference engine. The logic in this module determines which instructions must be executed by the computer. Commands correspond to scrolling to or selecting a particular on-screen button in order to specify the text to be input (see Figure 4.8).

In order to increase the typing-rate, the text prediction engine called Presage is used for generating text suggestions. Presage is a text prediction platform that assists in reducing the number of keystrokes required to type text [Presage 2015]. It does this by offering suggestions that are based on statistical properties of the English language and the usage patterns of a user. The required system events are generated with Autoit. Autoit is a scripting language designed for automating the Windows GUI and general scripting. Amongst other capabilities, Autoit can generate operating systems events such as *key-press* events [Autoit 2014]. Figure 4.5 provides an image of the GUI of the text input application. The GUI has the following controls on it:

- **Text box** the text box maintains the string of characters that have to be displayed and which correspond to the input from the user.
- **Alphabet keys** these keys append the indicated character at the end of the text in the text box.
- **Delete key** this key removes the last character or the last prediction to have been appended in the text box.
- **Space key** this key appends a blank character at the end of the text in the text box.
- **Enter key** this key appends a *new-line* character at the end of the text in the text box.
- **List box** the list box provides the predictions which are generated by the Presage engine. When the prediction engine is invoked, focus is transferred to the first element in the list box. The first element in the list box, which is depicted as *ESC* performs an escape function. When this option is selected, focus is transferred to the alphabet keys so as to resume text input.
- **Labels** the labels on the GUI provide the state information. The red label indicates the last key to have been pressed. The blue label indicates the current control in focus. When the *select* command is executed, then it is the control in focus that is selected.

This module corresponds to the application logic interface and multiple process symbols in Figure 4.1.

Data Visualization Tools allow the researcher to visualize input data streams, frequency decompositions, and classification performance results. Three of the visualization tools and the graphical information produced by each of them are described below. These tools are associated with the classification and command inference performance measurement multiple process symbols in Figure 4.2.

Temporal classification performance plots depict the temporal sequence of mental task classification and ground-truth labels. The purpose of this tool is to illustrate how well the model can infer the correct mental state given the input data. The tool provides visual information pertaining to the actual sequence of mental tasks which were performed by the user. In addition to the actual sequence, the tool also provides an overlay of the inferred

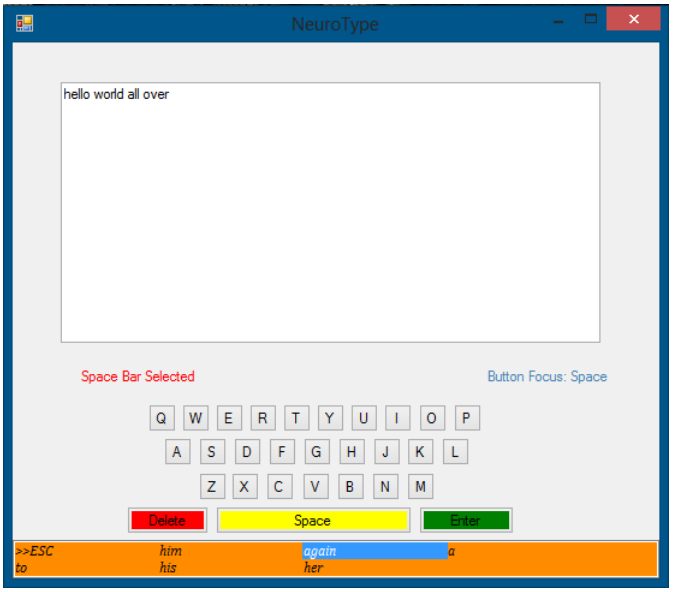


Figure 4.5: Text input GUI with predictive text functionality.

sequence of mental tasks. The inferred sequence is produced by a model in response to the data it receives as input. Figure 4.6 provides an example output produced by the tool. A further discussion on this tool is provided in Section 2.3 of Chapter 6.

Temporally-Sequenced Spatio-Spectral Topographic Maps (TeSSTMaps) depict the underlying cortical activity by using EEG as a correlate signal. TeSSTMaps information is important because it provides two fundamental aspects of the nature of the underlying cortical activity.

First, it provides information concerning the temporal evolution of the underlying cortical activity. Cortical activity patterns evolve differently in time depending on the mental task being executed. Second, it provides information concerning the spatial distribution of the power spectrum of the EEG signal. Within a single spatial region, EEG activity is distributed unevenly with respect to frequency components of the signal (see Figure 4.7). Thus it is important to gain further resolution into the spectral distribution of the EEG data, because certain activation patterns manifest only within a narrow frequency band.

Figure 4.7 provides a snap-shot view of the TeSSTMaps visualization tool. In the figure, the rows represent the mental task performed, and the columns represent the frequency component at which the spatial distribution of the activity is presented. The heads are to be interpreted as the view from above the head of the user of the BCI, with the nose pointing in the direction of the arrow. The implementation of this visualization tool was achieved by extending the EEGLab toolbox.[4]

[4]The EEGLab toolbox is developed and maintained by the Swartz Center for Computational Neuroscience, and may be obtained from http://sccn.ucsd.edu/eeglab/.

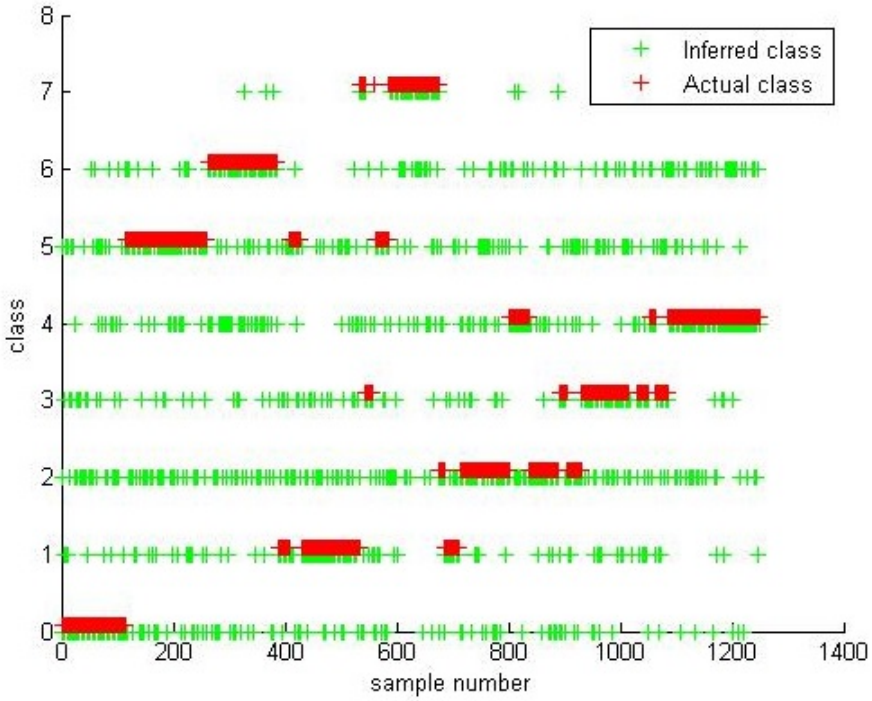


Figure 4.6: Temporal classification performance visualization tool.

4.2 Data Collection

Training data with ground-truth labels has to be collected from different individuals and at different times while performing certain mental tasks. The data have to be labeled because they provide a ground-truth reference, and because supervised machine learning techniques are employed in this work. The data is collected from different individuals at different times so that the model does not learn spurious patterns. EEG data exhibits non-stationarity between recording sessions, and this non-stationarity leads to a BCI performing better under specific contexts [Delorme et al. 2010]. Consequently, collecting samples under different conditions is required in order to better assess the performance of the system. Other data-sets have been obtained from other researchers working on BCI technology, and who use clinical-grade EEG technology. The exact details pertaining to how the experiments were performed are outlined in Chapter 5.

4.3 System Performance Evaluation

System performance evaluation pertains to assessing whether the system provides operational capabilities that satisfy reliability constraints. In this work, the performance of the system is evaluated using two metrics. The metrics provide a measure of how well the reliability constraints are satisfied. The reliability constraints are associated primarily with the classification performance of the model and the command inference engine. The first metric measures the classification accuracy of the model. The second metric measures the inference accuracy of the command inference engine. In the interest of ensuring readability, an outline of command inference is provided prior to discussing the metrics. Command inference is developed first in order to give meaning to the second metric.

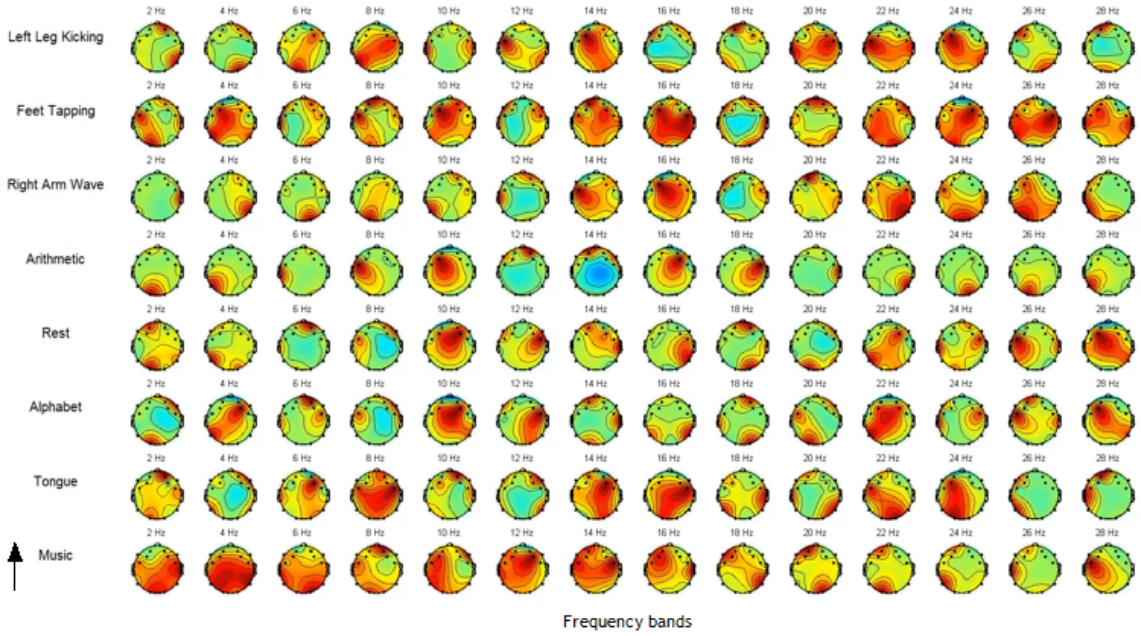


Figure 4.7: Snap-shot of temporally-sequenced spatio-spectral topographic maps. The row labels in the image represent mental tasks, and the column labels represent frequency bins that are spaced in 2Hz intervals from 2Hz to 26Hz.

4.3.1 Command Inference

Command inference is the process by which an intended command is determined from a stream of potentially conflicting commands. This is essential if a stream of symbols contains noise in the form of unintended commands elements. An example of this phenomenon can be seen in Figure 4.6. In the figure, the segment of correct classification results is depicted in red and corresponds to 0. However, the inferred classification results are distributed across all classes except class 7. This can occur if there is a drift effect in the process that produces the symbols. In the context of BCI applications, the drift occurs when a person fails to focus on a single mental task, or when the signal is corrupted by unaccounted artifacts. This leads to different classification results when there should be a single consistent class label. Consequently, conflicting instructions may be executed if steps aren't taken to mitigate the problem. In the context of this work, the aforementioned problem results in the scenario described below.

Using Figure 4.8 as a reference, suppose a user intends to scroll rightwards until the last prediction element gains focus, in which case the element would be highlighted. A transient interruption in the execution of *mental task 4* would result in an undesirable classification result which would not correspond to *mental task 4*. If command inference were not applied, there would be excessive scrolling or selection of inappropriate controls on the GUI. The spurious activity would be demotivating to the user of the BCI. In this situation, there must be a mechanism that evaluates a segment of the stream of symbols so that the most likely command is extracted. In this work, command inference is part of the post-classification step, and its impact is discussed in Section 3.2 of Chapter 6.

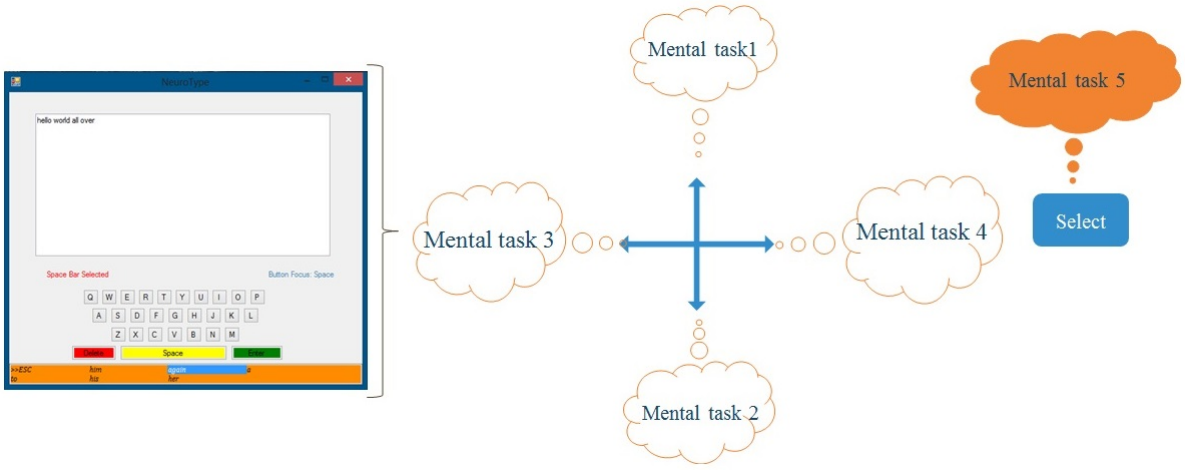


Figure 4.8: Control model as defined on the text input interface.

4.3.2 Classification Performance Metric

The classification performance metric measures how well the model learns the structure of the training data, and how well it can generalize when it is presented with previously unseen test data. The metric is defined as the classification accuracy. Classification accuracy is basically the correspondence between the inferred classification sequence and the ground-truth classification sequence, and this represents the true positive classification rate. Other complementary metrics are combined with the aforementioned metric. The complementary metrics include classification results pertaining to the true negative rate, false positive rate, and false negative rate. The mathematical formulation of the metrics is presented below:

$$true\,positive\,rate = \frac{true\,positive}{true\,positive + false\,positive} \tag{4.1a}$$

$$true\,negative\,rate = \frac{true\,negative}{true\,negative + false\,negative} \tag{4.1b}$$

$$false\,positive\,rate = \frac{false\,positive}{true\,positive + false\,positive} \tag{4.1c}$$

$$false\,negative\,rate = \frac{false\,negative}{true\,negative + false\,negative}. \tag{4.1d}$$

The complementary metrics provide insight pertaining to how spurious the classification results are. The intuition behind the metrics is as follows. Consider two cases in which the true positive rate is low. In one case, the classification results are spread out evenly across all classes. In the other case, the classification results are concentrated in one class. From a machine learning perspective, the former case is not too problematic since it indicates that the model is not learning the disambiguation between the classes. In a sense, the model exhibits low levels of confidence in the classification and this is acceptable when the representational capabilities of the model are taken into consideration. The latter case is problematic because the model is incorrect with a high level of confidence. The confusion matrix discussed in Section 2.5 of Chapter 6 provides a coherent framework for presenting the metric results.

This assessment is primarily performed using the off-line data flow model (see Figure 4.2). The temporal classification performance tool is one of the visualizations used in this context (see Figure 4.6). The tool reports how well the classifier performs without the use of command inference at inferring the mental task being performed by the user. If the classification accuracy is zero, then the red and green symbols in the tool are completely misaligned along the sample number axis for all samples. If the classification accuracy is one, then the red and green symbols in the tool perfectly align along the sample number axis for all samples.

4.3.3 Control Performance Metric

The classification performance metric provides information concerning how well the mapping has been learnt. However, it does not fully report how well the whole system responds to the intentions of the user. The primary concern pertains to how well the system can be controlled to perform the task intended by the user. The control performance metric provides this information.

The definition of the control performance metric is similar to that of classification performance. The difference is that command inference is applied to the classification results before the correspondence between the ground-truth and inferred sequences is assessed (see Section 4.3.1). This essentially means that some of the errors that manifest as the results of excessive variation in the classification results are discounted. The discounting is appropriate because at the level of control, the excessive variation in the classification results has no influence owing to the variation being averaged out by the application of techniques such as categorical binning.

The basic idea is that the bin with the highest number of symbols at some point in time is selected. The symbol it contains is thus assumed to be the intended symbol (i.e. the command that the user intended to issue). Mathematically, this corresponds to the following. Assume the number of elements that can be stored in a bucket is k and that all r buckets have the same capacity. At any time t, the collection of buckets $\mathbf{x}(t) = \{x_1(t), x_1(t), .., x_r(t)\}$ contains the number of elements in each of the buckets. The task is to select the bucket corresponding to class $y^*(t)$ such that:

$$y^*(t) = \arg\max_y x_y(t) \tag{4.2}$$

All the buckets are periodically emptied after a time period that is equivalent to obtaining k samples. The emptying is performed in order to avoid the situation in which one particular bucket accumulates more elements than the other buckets, the consequence would be getting stuck on a single result.

Another feature that is incorporated into the control performance metric is the discounting based on low likelihood values associated with the classification. If for instance the classification result is less than 10% above the probability of being randomly chosen, then the result is substituted with a classification result that corresponds to *no action*, in which case the BCI would not execute an incorrect control action. Thus the overall control performance would improve under the aforementioned circumstances. This form of discounting has to be applied to the classification results before categorical binning is applied. Mathematically, the discounting is formulated as:

$$y = \begin{cases} y_{\{infered\ by\ model\}}, P(y_{\{infered\ by\ model\}}|\theta, \mathbf{x}) \geq \frac{1}{r} \\ y_{\{no\ action\}}, P(y_{\{infered\ by\ model\}}|\theta, \mathbf{x}) < \frac{1}{r} \end{cases} \tag{4.3}$$

where θ represents the model parameters and $\mathbf{x}$ represents the data sample provided to the model.

More sophisticated approaches would employ machine learning techniques such as the hidden Markov model (HMM). An HMM would be particularly useful for predicting what the user would likely intend to do in the next time-step. This could be achieved by determining the next most likely sequence of instructions given the current instruction. The same visualization tool used in Section 4.3.2 is used in this assessment as well. However in this context, the focus is primarily on the metric, and not so much on the visualization.[5]

5 Experimental Equipment

Obtaining a reliable and high quality EEG signal presents itself as an engineering problem. The problem is complicated by constraints that must be satisfied in order to make the technology practical for everyday use. Some of the major constraints placed on the device are low production cost, high portability, and ease of use. The objective of this work is to attempt to build a low-cost BCI for people of average income means, so as to improve prospects of applicability in everyday life. As such, the chosen EEG sensor is a commercial device that is not suitable for clinical purposes [Emotiv 2014]. The benefits associated with using this sensor are that the constraints alluded to above are satisfied.

Emotiv Systems produces an EEG sensor that satisfies the requirements that would make the technology usable in everyday situations (see Figure 4.9). The Emotiv EPOC headset has the following 14 wet-contact electrodes that correspond to the cortical areas as mandated in the international 10-20 system: AF3, F7, F3, FC5, T7, P7, O1, O2, P8, T8, FC6, F4, F8, AF4 (see Figure 2.11). The sensor communicates with its USB dongle via a WI-FI connection, and can provide 12 hours of continual use [Emotiv 2014]. The sensor is light-weight and is easy to use in comparison to conventional clinical EEG devices, in that the preparation step only requires that a saline solution is dripped onto the channel contacts. The sensor has an SDK that simplifies the application development task. In this work, the .NET dynamic link library was used to extract EEG samples and the status of the device as depicted in Figure 4.3.

5.1 Experimental Equipment Limitations

The Emotiv EPOC is not free of problems. First, the manufacturer states that the sensor is not suitable for clinical purposes [Emotiv 2014]. This follows from the fact that the sensor does not have the capabilities of its clinical counterparts. For example, sampling rates usually have a minimum of 256 Hz, and channel numbers are usually above 64 in clinical sensors. In contrast,

[5] It is important to note that in terms of the visualization, the subsequences that have no correspondence will be unchanged so as to reflect error. The metric merely discounts the lack of correspondence in the interest of maintaining the stability in the inferred command.



Figure 4.9: Emotiv EPOC EEG headset.

Source: Emotiv EPOC specifications [Emotiv 2014].

the Emotiv EPOC has a sampling rate of 128 Hz and 14 channels. These shortcomings affect both the temporal and spatial resolution power of the sensor. Consequently, the granularity in both the temporal and spatial distributions of the EEG data is limited. However, the limitations may not be too restrictive if the researcher does not seek to unveil neuroscience principles with the sensor.

Second, there is usually a compromise on the signal-to-noise ratio (SNR) when low-cost electronics are used. This follows from the fact that sophisticated electronics circuitry that would limit the SNR to acceptable values is costly. As a consequence of reducing production cost, high SNR does not constitute a major production priority. In light of experimental results, this topic is taken up again in Section 3.1 of Chapter 6 where experiment results are discussed.

Third, saline based wet-electrode technology such as the EPOC and which is not compatible with conductive paste, or gel, provides suboptimal conductivity. With clinical technology, the conductivity is improved by the use of conductive paste or gel. Good conductivity is essential for obtaining good quality data. When conductive paste is used, the electrodes stick to the scalp location to which they are attached. While in the case of conductive gel, the tension on the mounting cap ensures that tight physical contact is maintained at all times.

The Emotiv EPOC does show reliability problems in the context of conductivity. The problems manifest as contact quality problems in the electrodes. For an illustration, see the black circle in Figure 4.3. One of the reasons why contact problems arise is that mechanical pressure on the scalp is determined by the fixed structure of the EPOC, and the shape of the head of the user. The Emotiv EPOC is simply mounted on the head. Beyond that, not much can be done to maintain good conductivity. These problems were almost uncontrollable during experimentation. Even with the intervention of cleaning the scalp with alcohol and padding of the electrode contacts with saline solution-soaked pieces of cotton wool, the problems persisted. Lastly, a residue that builds up on the surface of the metal caps in the electrodes seems to imply an occurrence of electrolysis. There is a recurring build-up of a solid substance on the surface of the metal caps which is reminiscent of copper sulphate. The effect of the substance has not been established in this work.

6 Discussion and Conclusion

In this chapter, the framework within which the experiments were performed was discussed. The motivation for initiating this work was provided, in which the social and economic cost of the loss of the neuro-muscular connection was discussed (see Section 2). Section 3 described the research hypothesis that this works aims to test, and a brief motivation concerning the feasibility of achieving the result was presented in the same section.

The subsequent section described the methodological aspects of the work. This entailed a discussion of the system design and development issues (see Section 4.1). Given that this work is our first attempt at the subject matter and that extensibility to other settings is a design objective, substantial modifications of the software components are anticipated in future work. Thus a modular architecture was identified as being the most appropriate for this work (see Section 4.1.1). The software components required for experimentation and their implementation details were also discussed (see Section 4.1.2). Snap-shots of some of the GUI components were also exhibited. Subsequent topics tackled include the data collection process (see Section 4.2), and the system performance evaluation strategies along with their associated metrics (see Section 4.3). A description of the methodology concluded with a critical review of the low-cost experimental equipment used in this study (see Section 5). The review included a discussion of the limitations of this equipment, in which electrode contact quality and sampling rate concerns were highlighted (see Section 5.1).

In the next chapter, the experiment procedure is outlined. The classification performance assessments of the probabilistic graphical model based techniques were performed using the experiment procedure outlined in that chapter. The experiments were performed using EEG data obtained using the Emotiv EPOC headset, and data obtained from other researchers who used clinical-grade EEG data.

Chapter 5

Investigation Procedure

1 Introduction

This chapter outlines the experimental procedure followed in this work. In particular, two experimental issues are addressed. The first pertains to the data collection methodologies, in which the two groups of data-sets are discussed. The first data-set was obtained using low-cost EEG technology and the experiments were carried out as part of this work. The second data-set was obtained from the BCI Competition data-sets under BCI Competition III, and the particular data-set V was provided by the Idiap Research Institute. A discussion pertaining to this is provided in Section 2.

The second experimental issue pertains to the machine learning models that were used for the classification of mental task related EEG data (see Section 3). In that section, two different probabilistic graphical models are reviewed. The first model is the hidden conditional random fields (HCRF). The HCRF is suited to modeling intrinsic structure within a given mental task. The second model is the latent dynamic conditional random fields (LDCRF). The LDCRF is suited to modeling both the intrinsic structure within a given mental task, and the extrinsic structure pertaining to transitions between mental tasks. The latter is crucial to the detection of mental task switching. A detailed account concerning the structure and the functioning of the models is provided in Section 3.1 and Section 3.2, respectively.

2 Data Collection Methodologies

The purpose of this section is to outline the structure of the experimental trials and how the data was obtained. The discussion begins with the experimental procedure that was followed using the Emotiv EPOC EEG headset, and ends with a description of the experimental procedure that was followed in the BCI Competition III, in data-set V by Chaippa [2004].

2.1 Emotiv Headset Collected Data

As discussed in Section 2 of Chapter 1, our objective is to create a BCI based on low-cost technology. To achieve this, we conducted our own experiments with 5 participants in our study using the Emotiv EPOC headset. The experimental paradigm performed with each subject is

outlined below. The experiments were carried out over three phases. In each phase, there were specific objectives to be met and tests to be performed as described below.

2.1.1 Phase 1: Training without control feedback

The primary purpose of phase 1 was to assist the subjects with improving their ability to focus on a particular mental task. This is essential for obtaining data that exhibits some degree of consistency. The secondary purpose was to familiarize the subjects with the technology, and to make them aware of the various sensitivities that affect the measurement process. Examples of such sensitivities include excessive blinking, body movements, and changing lighting conditions. The subjects were required to perform various mental tasks, including those presented in Section 2.1.2. The data that was obtained was used to assess how well the mental tasks could be disambiguated.

This phase extended over a period of one month. There were two sessions per week, and each session lasted for 20 minutes. Each mental task was performed three times over three trials, and each trial lasted for 30 seconds. The mental tasks were performed in a random order in each session. All the subjects went through phase 1.

Frequency Band Selection The frequency bands that were assessed in phase 1 ranged from 3 Hz to 28 Hz, and the objective was to see if the expected cortical activation responses could be observed using the Emotiv headset. The visualization tool used to perform this is the TeSSTMaps depicted in Figure 4.7. It is also in this phase that candidate mental tasks were identified. The requirement for these was that the temporal evolution of the data looks different between the candidate mental tasks. The candidate mental tasks that were selected in phase 1 were used in the subsequent two phases.

2.1.2 Phase 2: Discontinuous task execution without control feedback

The primary purpose of phase 2 was to determine if the intrinsic structure of each of the candidate mental tasks could be identified using the HCRF and the Emotiv headset. In order to determine this, each subject was asked to perform the following mental tasks and then the trained HCRF was used to verify if the structure had been learnt:[1]

1. Imagined repetitive left leg kicking.
2. Imagined repetitive right leg kicking.
3. Imagined repetitive left arm waving.
4. Imagined repetitive right arm waving.
5. Imagined repetitive tongue movements.
6. Performing continual arithmetic operations.
7. Continual alphabet recitation in reverse order.

[1] A discussion about how the HCRF was trained is presented in Section 3.1

Table 5.1: Emotiv Data: Frequency bands and channels used for HCRF and LDCRF state representation.

Frequency Bands (Hz)	Channels
10-12	AF3, F7, F3, FC5, T7, P7, O1, O2, P8, T8, FC6, F4, F8, AF4
16-18	AF3, F7, F3, FC5, T7, P7, O1, O2, P8, T8, FC6, F4, F8, AF4
20-22	AF3, F7, F3, FC5, T7, P7, O1, O2, P8, T8, FC6, F4, F8, AF4

8. Maintaining a state of rest without focusing on a particular task.

These tasks were selected such that different cortical regions could be activated. This would increase the chances of having differentiated data from the associated mental tasks. Identifying differentiated mental tasks is important because each of those mental tasks can then be mapped to an abstract class. That abstract class can then be associated with the execution of a certain set of instructions on the computer. Each subject performed each mental task twice in a single session. Each mental task was performed for about one minute at a time, and this represented a trial. There was a 10 second break between the trials. Apart from maintaining a state of rest, the mental tasks were performed in a random order in each session and across all subjects. In addition, all sessions began with the rest state.

Frequency Band Selection In phase 2, we used the frequency components used by Fok et al. [2011] and Roula et al. [2012]. The frequency components are presented in Table 5.1. There are two reasons for this selection. First, in both studies the Emotiv headset was used thus forming a reference point for our work. Second, the frequency bands are supported by neurophysiological arguments and experimental findings [Fok et al. 2011].

2.1.3 Phase 3: Continuous task execution without control feedback

In phase 3, only two subjects participated in the study. The purpose of phase 3 was to determine if the intrinsic structure of each of the mental tasks, and extrinsic structure between different mental tasks, could be captured using the LDCRF and the Emotiv headset. In order to determine this, the subjects were required to perform mental tasks similar to those in data-set V of the BCI Competition III. The mental tasks are outlined below. The trained LDCRF was used to verify if the structure had been learnt:[2]

1. Imagination of repetitive left hand movements.
2. Imagination of repetitive right hand movements.
3. Generation of words beginning with the same random letter.

The above mental tasks were selected in order to facilitate comparison between the performance of the LDCRF when the Emotiv data is used, as opposed to when the BCI Competition data is used. The comparison does not permit direct correspondence, because the montage used in data-set V of the BCI Competion III is different from the montage provided by the Emotiv

[2]A discussion about how the LDCRF was trained is presented in Section 3.2

headset.[3] We follow an experimental paradigm similar to that outlined in Section 2.2. The commonality lies along the following specifications: "sessions of a given subject were acquired on the same day, each lasting 4 minutes with 5-10 minutes breaks in between them. The subject performed a given task for about 15 seconds and then switched randomly to another task at the operator's request. EEG data is not splitted in trials since the subjects are continuously performing any of the mental tasks" [Chaippa 2004].

Frequency Band Selection The table detailing the frequency bands and the channels used is provided in Table 5.1. The frequency bands have been selected using neuroscience insights, and they are similar to those used by Fok et al. [2011] and Roula et al. [2012].

2.2 BCI Competition Sourced Data

This data was provided in BCI Competition III, as data-set V.[4] The data was provided by Silvia Chiappa and José del R. Millán of the Idiap Research Institute. EEG signals were recorded with a Biosemi system using a cap with 32 integrated electrodes located at standard positions of the International 10-20 system. The sampling rate was 512 Hz. Signals were acquired at full DC, and no artifact rejection or correction was employed [Chaippa and Millán 2004].

This data-set contains mental imagery data that has been recorded continuously so as to allow the testing of the performance of a BCI with an asynchronous protocol. This is the same data that was used by Saa and Çetin [2013]. Chapter 6 presents a comparative analysis between our findings obtained using this data and the findings reported by Saa and Çetin [2013]. The experimental paradigm presented by Chaippa and Millán [2004] is outlined below.

2.2.1 Continuous task execution without control feedback

This data-set contains data from 3 normal subjects during 4 non-feedback sessions. The subjects sat in a normal chair, with relaxed arms resting on their legs. There are 3 tasks:[5]

1. Imagination of repetitive self-paced left hand movements (class 2).
2. Imagination of repetitive self-paced right hand movements (class 3).
3. Generation of words beginning with the same random letter (class 7).

All 4 sessions of a given subject were acquired on the same day, each lasting 4 minutes with 5-10 minutes breaks in between them. The subject performed a given task for about 15 seconds and then switched randomly to another task at the operator's request. EEG data is not splitted in trials since the subjects are continuously performing any of the mental tasks. Our classification algorithm is thus required to provide an output every 0.5 seconds using the

[3] In order to see the differences between the montages, see the description in Section 2.2.1 and Table 5.1.

[4] bbci.de/competition/iii/desc_V.html

[5] In this work, these class labels were later redefined, respectively, as: 0, 1, and 2. This was done to make the results consistent with the implementation of the machine learning models. The implementation does not treat the results as abstract labels. The numerical ordering of the labels starting from 0, and with an increment of 1 unit, is required without any "integer gaps" in the set of labels.

Table 5.2: BCI Competition Data: Frequency bands and channels used for HCRF and LDCRF state representation.

Frequency Bands (Hz)	Channels
10-12	C3, Cz, C4, CP1, CP2, P3, Pz, P4
16-18	C3, Cz, C4, CP1, CP2, P3, Pz, P4
20-22	C3, Cz, C4, CP1, CP2, P3, Pz, P4

last second of data (see clarification in the paragraph 'Requirements and Evaluation'). Data are formatted in two ways:

1. Raw EEG signals. The sampling rate was 512 Hz.
2. Pre-computed features. The raw EEG potentials were first spatially filtered by means of a surface Laplacian. Then, every 62.5 ms –i.e., 16 times per second– the power spectral density (PSD) in the band 8-30 Hz was estimated over the last second of data with a frequency resolution of 2 Hz for the 8 centro-parietal channels: C3, Cz, C4, CP1, CP2, P3, Pz, and P4. As a result, an EEG sample is a 96-dimensional vector (8 channels times 12 frequency components).

Frequency Band Selection The frequency bands used in this phase are the same as those presented in Table 5.2. The channels used are those which are provided in the data-set, as described above. It is important to note that the locations of the electrodes determines what is measured in EEG data. Consequently, the results from this data are not directly comparable to those obtained using the Emotiv headset because of the differences in the electrode placement. The data does however provide a reliable benchmark for assessing the performance of the LDCRF, and that is precisely the purpose of this validation.

3 Machine Learning Models

In this work, the performance of two probabilistic graphical models was assessed. The two models are the hidden conditional random fields (HCRF), and latent dynamic conditional random fields (LDCRF). These models were chosen for three reasons. First, they provide compact representations of the probability distributions that they model. Second, they are capable of capturing temporal structure in the data. Lastly, the family of discriminative probabilistic graphical models does not mandate that the state vector that is used have uncorrelated elements. This makes selecting the elements of the state vector significantly less complex. This flexibility is in contrast to generative probabilistic graphical models such as the hidden Markov model which require this statistical independence assumption.

Both the HCRF and LDCRF models have been applied in various other domains, and have also been used previously in the BCI domain [Saa and Çetin 2012, Saa and Çetin 2013, Quattoni et al. 2007, Morency et al. 2007]. These two models are depicted with blue labels in Figure 5.1. The main difference between the two models is in what each model is intended to represent, and this difference is outlined in Sections 3.1 and 3.2, respectively. In the next subsection, we present the structure that is common to both the HCRF and the LDCRF.

The probabilistic graphical models that are used in this work are provided by the Accord.NET framework and the HCRF library. The Accord.NET Framework is a .NET machine

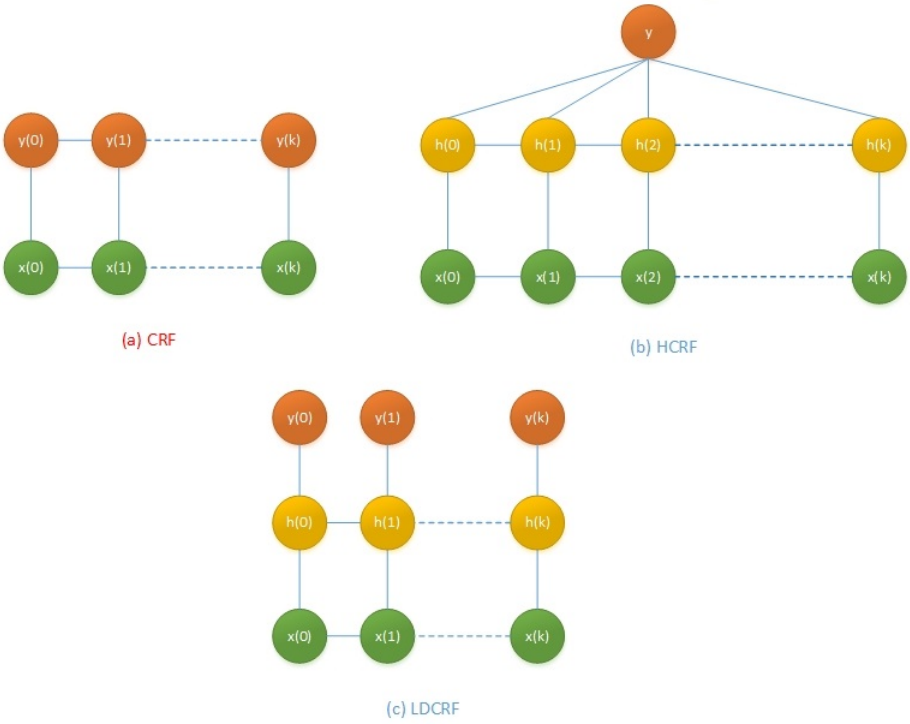


Figure 5.1: The evaluated discriminative probabilistic graphical models have blue labels. The CRF is shown for comparison purposes, and both the HCRF and LDCRF have been used in the experiments.

learning framework combined with audio and image processing libraries completely written in the C# language [Accord 2015]. It is a complete framework for building production-grade computer vision, computer audition, signal processing, and statistics applications [Accord 2015]. In this work, the signal processing and machine learning libraries were used for power spectrum estimation and classification functionality, respectively. Amongst other implementations, the HCRF library provides implementations of both the HCRF and LDCRF [HCRF-lib 2015]. The Matlab toolbox of the HCRF library is used to evaluate the performance of the LDCRF model.

3.1 Hidden Conditional Random Fields

The HCRF is suited to problems in which the modeling of the intrinsic structure of each class is required. These kinds of problems arise when assessment concerning internal structure corresponding to each class is sought. In the context of this work, the classes are abstract numbers which represent different mental activities. An example mental task is imagining waving the left hand. The HCRF is trained on data-sets that have pre-segmented sequences, and it provides a single classification result for each segment presented to it. In concrete terms, the segment length corresponds to the value of the k that appears in Figure 5.1. The length should be selected such that the classification performance of the model is maximal. This requirement of pre-segmentation is problematic for various reasons.

It is difficult to determine the appropriate size of the window of samples that are supposed to be given the same label, as there is no clear theoretical basis for setting that window size. Thus, the HCRF cannot adequately model the transition dynamics between different mental

tasks. Given the asynchronous protocol problem being addressed in this work, this limitation is substantial and motivates the use of the LDCRF described in Section 3.2.

The main objective when using the HCRF is to obtain the conditional probability distribution. The conditional probability distribution can subsequently be used for classification. This conditional probability distribution is defined as follows. Assume a prediction task in which a label y is associated with each m-dimensional input vector $\mathbf{x} = \{x_1, x_2, .., x_m\}$, and that y is selected from a discrete set of labels Y. Thus, the training data is of the form $(\mathbf{x}_i, y_i)$ with $y_i \in Y$ and $\mathbf{x}_i = \{x_{i,1}, x_{i,2}, .., x_{i,m}\}$. In this work, the set Y represents the set of labels corresponding to each of the mental tasks used in the experiments. The vectors $\mathbf{x}_i$ represent the input data which are the selected components of the power spectra of the EEG data, and are observed every half a second during runtime.[6]

For any $\mathbf{x}$, a vector of unobserved latent variables $\mathbf{h} = \{h_1, h_2, .., h_m\}$ is assumed. Each of the h_j are selected from a discrete set of labels H and represent the hidden states which assume positive integer values. The purpose of these latent variables is to capture intermediate structure in the input data. A point of significance here is that the labels defining this intermediate structure are not specified in the training data, and may represent structure that may not be readily interpreted using neuroscience insight. The joint probability of the labels and the hidden states given the data is defined as:

$$P(y, \mathbf{h}|\mathbf{x}, \theta) = \frac{e^{\Psi(y, \mathbf{h}, \mathbf{x}; \theta)}}{\sum_{y', \mathbf{h}} e^{\Psi(y', \mathbf{h}, \mathbf{x}; \theta)}}, \tag{5.1}$$

where θ represents the parameters of the model, and $\Psi(y', \mathbf{h}, \mathbf{x}; \theta)$ is a potential function parameterized by θ [Quattoni et al. 2007]. In order to determine the conditional probability of class y given the input data and the model parameters, we marginalize out the hidden state vector. This conditional probability is then given by:

$$P(y|\mathbf{x}, \theta) = \sum_{\mathbf{h}} P(y, \mathbf{h}|\mathbf{x}, \theta) = \frac{\sum_{\mathbf{h}} e^{\Psi(y, \mathbf{h}, \mathbf{x}; \theta)}}{\sum_{y', \mathbf{h}} e^{\Psi(y', \mathbf{h}, \mathbf{x}; \theta)}}. \tag{5.2}$$

HCRFs use undirected graphical structures, with the graph defined by $G = (V, E)$ where V denotes the vertices in the graph and E denotes the edges [Saa and Çetin 2012]. Given this, the potential function is defined as:

$$\Psi(y, \mathbf{h}, \mathbf{x}; \theta) = \sum_{j=1}^{m} \sum_{l \in L_1} f_{1,l}(j, y, h_j, \mathbf{x}) \theta_{1,l} + \sum_{(j,k) \in E} \sum_{l \in L_2} f_{2,l}(j, k, y, h_k, \mathbf{x}) \theta_{2,l}, \tag{5.3}$$

where $f_{1,l}$ and $f_{2,l}$ are feature functions of the HCRF related to the nodes and edges of the graph, respectively [Saa and Çetin 2012]. The feature functions respect the structure of the graph, in that no feature depends on more than two hidden variables h_j , h_k. If a feature does depend on variables h_j and h_k there must be an edge (j, k) in the graph G [Quattoni et al. 2007].

[6] See Table 5.1 and Table 5.2 for reference.

The $\theta_{1,l}$ and $\theta_{2,l}$ represent the model parameters corresponding to the nodes and edges. The parameters have to be learnt using the training data. The model parameter learning is treated as an optimization problem, in which gradient ascent algorithms such as the quasi-Newton gradient ascent algorithm can be used. The optimization problem is non-convex, and the function being optimized on θ is the log-likelihood function defined as:

$$L(\theta) = \sum_i log(P(y_i|\mathbf{x_i}, \theta)) - \frac{1}{2\sigma^2}\|\theta\|^2, \tag{5.4}$$

where the second term is a regularization term which facilitates the prevention of over-fitting [Saa and Çetin 2012, Quattoni et al. 2007]. This term is the logarithm of a Gaussian prior with variance σ^2 [Quattoni et al. 2007]. Gradient ascent is then used to find local optimal parameter values, $\theta^* = argmax_\theta L(\theta)$. The graphical model contains information directly related to the decomposition of the potential function $\Psi(y, \mathbf{h}, \mathbf{x}; \theta)$ [Saa and Çetin 2012]. In the case in which the edges in E form a tree and that $\Psi(y, \mathbf{h}, \mathbf{x}; \theta)$ assumes the form given above, then exact methods such as belief propagation can be used for inference and parameter estimation in the model [Quattoni et al. 2007]. If E contains cycles, then approximate methods such as loopy belief-propagation may be necessary for inference and parameter estimation [Quattoni et al. 2007]. Given a new test example $\mathbf{x}$ and parameter values θ^* induced from the training set, the label for the example is taken to be $\arg\max_{y \in Y} P(y|\mathbf{x}, \theta^*)$ [Saa and Çetin 2012].

3.2 Latent Dynamic Conditional Random Fields

The LDCRF model is meant to overcome the limitation concerning sequence segmentation and labeling of the HCRF. For the purposes of asynchronous BCI protocols, the limitation of the HCRF is not negligible. Similarly, the CRF lacks the ability to represent internal states for each class which can be used to increase the differentiability between classes [Saa and Çetin 2013]. Morency et al. [2007] presents the LDCRF which incorporates hidden state variables with an explicit partition function for normalization, and for which inference can be efficiently computed during both training and testing. Thus the LDCRF model can be considered as an extension of both the CRF and the HCRF in the sense that the strengths of both models are utilized, while the limitations are removed. The LDCRF addresses both of the aforementioned problems of the CRF and HCRF by modeling both the intrinsic and extrinsic dynamics of the mental tasks [Morency et al. 2007]. These two attributes allow this model to be used for labeling data sequences that have not been segmented.

The main task when using the LDCRF is to learn the mapping between a sequence of observations $\mathbf{x} = \{x_1, x_2, .., x_m\}$ and a sequence of labels $\mathbf{y} = \{y_1, y_2, .., y_m\}$ [Morency et al. 2007]. Each y_j is a class label for the j^{th} sequence element of $\mathbf{x}$, and is a member of the set Y of possible class labels. In this work, the set Y represents the set of labels corresponding to each of the mental tasks used in the experiments. The vectors $\mathbf{x}_i$ represent the input data which are the selected components of the power spectra of the EEG data, and are observed every half a second.[7] Each observation x_j is represented by a feature vector $\phi(x_j) \in \mathbf{R}^d$ [Morency

[7] See Table 5.1 and Table 5.2 for reference.

et al. 2007]. For each sequence $\mathbf{x}$, a vector of hidden variables $\mathbf{h} = \{h_1, h_2, .., h_m\}$ is assumed. The vector of hidden variables is not provided in the training data. Its purpose is to capture unspecified structure in the data, as in the case of the HCRF. Similar to the HCRF, the objective is also to obtain the conditional probability distribution. In this case however, the conditional probability distribution is used to label unsegmented sequences. Morency et al. [2007] defines this conditional probability distribution as:

$$P(\mathbf{y}|\mathbf{x}, \theta) = \sum_{\mathbf{h}} P(\mathbf{y}|\mathbf{h}, \mathbf{x}, \theta) P(\mathbf{h}|\mathbf{x}, \theta), \tag{5.5}$$

where the θ are parameters of the model. In order to keep training and inference tractable, Morency et al. [2007] proposes that the model should have disjoint sets of hidden states associated with each class label. This means that each h_j is a member of a set H_{y_j} of possible hidden states for the class label y_j. The set H which represents the set of all possible hidden states, is defined to be the union of all the H_{y_j} sets [Morency et al. 2007]. As a consequence of the above requirements, sequences which have any $h_j \notin H_{y_j}$ will by definition have $P(\mathbf{y}|\mathbf{h}, \mathbf{x}, \theta) = 0$ and thus give the more succinct:

$$P(\mathbf{y}|\mathbf{h}, \mathbf{x}, \theta) = \sum_{\mathbf{h}:\forall h_j \in H_{y_j}} P(\mathbf{h}|\mathbf{x}, \theta). \tag{5.6}$$

Morency et al. [2007] defines $P(\mathbf{h}|\mathbf{x}, \theta)$ using the CRF formulation:

$$P(\mathbf{h}|\mathbf{x}, \theta) = \frac{1}{Z(\mathbf{x}, \theta)} e^{(\sum_k \theta_k \cdot \mathbf{F}_k(\mathbf{h}, \mathbf{x}))}, \tag{5.7}$$

where the partition function Z is defined as:

$$Z(\mathbf{x}, \theta) = \sum_{\mathbf{h}} e^{(\sum_k \theta_k \cdot \mathbf{F}_k(\mathbf{h}, \mathbf{x}))}, \tag{5.8}$$

and $\mathbf{F}_k$ is defined as:

$$\mathbf{F}_k(\mathbf{h}, \mathbf{x}) = \sum_{j}^{m} f_k(h_{j-1}, h_j, \mathbf{x}, j), \tag{5.9}$$

where each feature function $f_k(h_{j-1}, h_j, \mathbf{x}, j)$ is either a state function $s_k(h_j, \mathbf{x}, j)$ or a transition function $t_k(h_{j-1}, h_j, \mathbf{x}, j)$ [Quattoni et al. 2007]. The state functions s_k depend on a single hidden variable in the model while the transition functions t_k can depend on pairs of hidden variables. Parameters associated with a transition function for hidden states that are in the same subset H_{y_j} model the substructure patterns. The parameters associated with the transition functions for hidden states from different subsets model the external dynamics between different mental tasks. The model parameters associated with the transition functions model both the intrinsic and extrinsic dynamics. Given the definition of the model, the next task is to determine the optimal model parameter values θ^*, given training data-set of n labeled sequences $(\mathbf{x_i}, \mathbf{y_i})$ for $i = 1..n$. As in the case of the HCRF, determining these values is reduced to an optimization problem in which the log-likelihood is maximized. Morency et al. [2007] defines the log-likelihood function as:

$$L(\theta) = \sum_{i=1}^{n} log(P(\mathbf{y}_i|\mathbf{x_i}, \theta)) - \frac{1}{2\sigma^2}\|\theta\|^2, \tag{5.10}$$

where the second term is a regularization term, similar to that of the HCRF. This term is the logarithm of a Gaussian prior with variance σ^2. Gradient ascent is then used to find local optimal parameter values such that $\theta^* = argmax_\theta L(\theta)$ [Morency et al. 2007]. Once the model has been trained, a new test sequence $\mathbf{x}$ is presented and the objective is to determine the most probable label sequence $\mathbf{y}^*$ that maximizes the conditional probability:

$$\mathbf{y}^* = \arg\max_{\mathbf{y}} P(\mathbf{y}|\mathbf{x}, \theta^*), \tag{5.11}$$

Continuing with the assumption that each class label is associated with a disjoint set of hidden states, the previous equation can be rewritten as:

$$\mathbf{y}^* = \arg\max_{\mathbf{y}} \sum_{\mathbf{h}:\forall h_i \in H_{y_i}} P(\mathbf{h}|\mathbf{x}, \theta^*). \tag{5.12}$$

In order to estimate the label y_j^* of sequence element x_j, the marginal probabilities $P(h_j = a|\mathbf{x}, \theta^*)$ are computed for all possible hidden states $a \in H$ [Quattoni et al. 2007]. Then the marginal probabilities are summed according to the disjoint sets of hidden states H_{y_j} and the label associated with the optimal set is chosen. The marginal probabilities can efficiently be computed using belief propagation[8], while another option would be to compute the Viterbi path.[9]

4 Conclusion

In this chapter, the experiment procedure was outlined. Various discussions concerning the experimental paradigms and the machine learning models used were presented. There are two groups of data-sets that have been used in this work.

The first corresponds to the data we collected using the Emotiv headset. This is the data on which we investigate the performance of a low-cost BCI. There are various phases that were discussed. One of the phases included verifying whether or not the structure of the data within a single mental task can be learnt using the HCRF (see Section 3.1). This forms the basis for further investigations which ultimately lead to sequence segmentation and labeling, in which the LDCRF model is used (see Section 3.2). The LDCRF is used for this task since it provides a systematic way to model both the intrinsic and extrinsic dynamics of the mental tasks. The latter is crucial to the detection of mental task switching. The second data-set corresponds to the data obtained from BCI Competition III, data-set V. This data-set was provided by the Idiap Research Institute, and was collected using clinical-grade EEG technology. This data-set is used to validate the functional capabilities of the LDCRF model on sequence segmentation

[8] http://people.csail.mit.edu/billf/publications/Understanding_Belief_Propogation.pdf
[9] http://www.cim.mcgill.ca/~latorres/Viterbi/va_alg.htm

and labeling. Frequency components for testing both the HCRF and LDCRF on the two groups of data-sets were identified using neurophysiological insights, and frequency components used in related studies such as those of Fok et al. [2007], Roula et al. [2011], and Saa and Çetin [2012, 2013].

In the next chapter, the experiment results obtained from the various experimental paradigms are presented. There are various tools that are used to depict the results, therefore the presentation begins with a brief description of each of the tools (see Section 2). This is followed by a presentation of the actual experiment results, in which some of the observed irregularities are reported. The chapter then concludes with an analysis of the obtained results, in which the model performance is discussed.

Chapter 6

Results and Analysis

1 Introduction

In this chapter, we present the experimental results and the corresponding analysis of those results. As motivated in Section 2.2.2 of Chapter 2, our analysis is empirical and it is assumed that the underlying neurophysiological model which generates the EEG data is not accessible. The primary task in this work is to model the structure contained in the EEG data. This effectively means that the model need not be correct in neurophysiological terms. Thus the empirical analysis is not intended to provide epistemological insights concerning the neurophysiology of brain. Rather, it provides insight on the effectiveness of the models adopted to model the structure in the EEG data.

The first section of this chapter provides a discussion of the various visualization tools used in this work. Section 3 presents the results that were obtained from the experiment sessions. The results include some irregularities that were observed. Following that, results obtained from both the hidden conditional random fields (HCRF) and the latent dynamic conditional random fields (LDCRF) models are presented. Section 4 provides an analysis of the experiments results, and the significance of the research findings. The correspondence between the data obtained using the low-cost Emotiv headset and neuroscience is a topic of particular interest, and the correspondence is discussed in Section 4.1. The chapter concludes with a performance analysis of both the HCRF and LDCRF models.

2 Visualization Tools

The data visualization tools that were used in this work are discussed in the subsections below. Each of the tools provides insight into the various data-sets. The data-sets range from the EEG data to the model classification performance results. Each of the tools is discussed, and the interpretation of its representation is outlined.

2.1 EEG Time Series

The time series visualization tool depicts how the EEG signal varies over time (see Figure 6.1). This tool was used in the experiment sessions involving the Emotiv headset. The tool was used

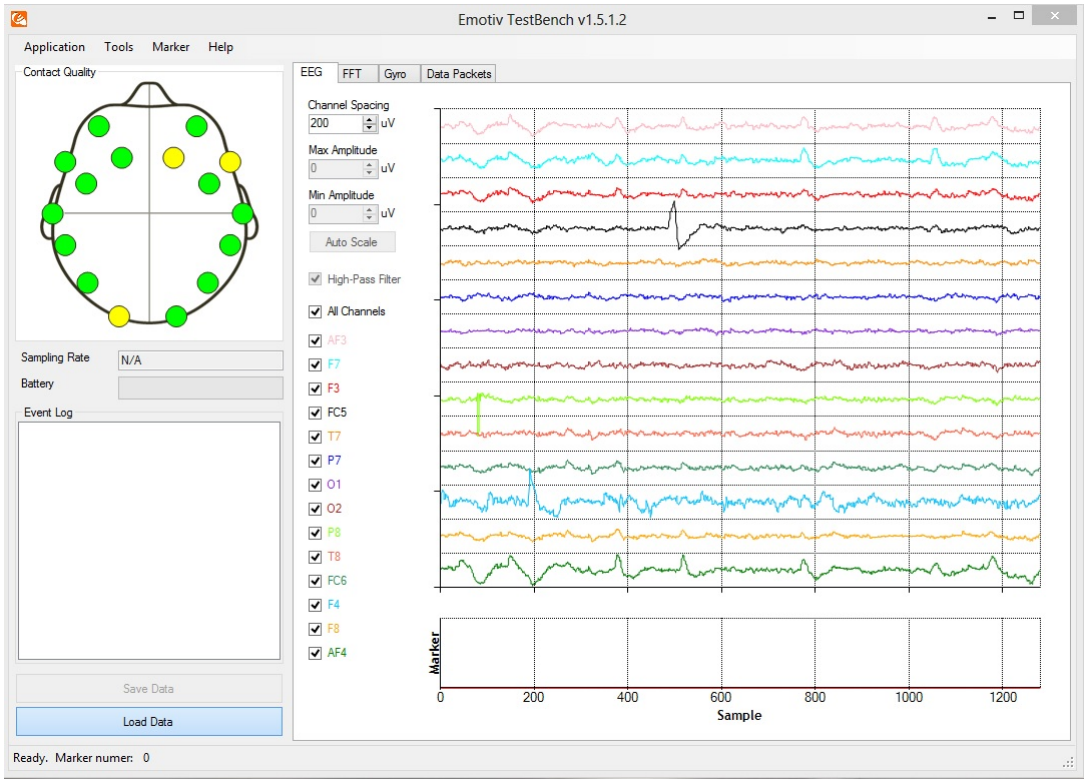


Figure 6.1: Emotiv EEG time series visualization program snap-shot.

to monitor the quality of the signal. Large variations in the amplitude usually indicates that there is muscle activity that is measured in the form of electromyographic (EMG) signals. These EMG signals cannot be ignored because they have a large influence on the EEG data. The time series visualization tool also provides insight corresponding to the effect of the variation of both the sampling rate of the EEG sensor and the contact quality of an electrode. This topic is expanded further in Section 3.1.

2.2 Temporally Sequenced Spatio-Spectral Topographical Maps

Temporally sequenced spatio-spectral topographical maps (TeSSTMaps) provide information pertaining to the temporal evolution of the topographically represented spatial distribution of the spectral decomposition of the EEG signal. A detailed discussion TeSSTMaps is provided in Section 4.1.2 of Chapter 4, and a snap-shot of the tool is depicted in Figure 4.7. Briefly, the TeSSTMaps indicate how the cortical activity is spatially distributed as time progresses. Essentially, this corresponds to a video showing how the cortical activity is varying in time. This tool was used to identify candidate mental tasks, as discussed in Section 2.1.1 of Chapter 5.

2.3 Temporal Classification Performance Plot

The temporal classification performance plots reveal information about how well the model infers the correct class when presented with state information. A description of the tool is given in Section 2 of Chapter 4. This tool was used to provide both a visual representation that depicts how well the ground-truth labels are tracked by the model, and a quantitative measure

of the true positive classification rate (see Figure 4.6). One of the main benefits of using this tool is that the temporal characteristics of the true positive classification rate are not obscured, in contrast to the confusion matrix plots. In other words, this tool makes it apparent when during mental task execution the misclassification occurs. This kind of resolution is important because it may happen that misclassification mostly happens around mental task switching, in which case we could infer that the model doesn't represent the transition dynamics well. In the context of intra-mental task classification, the objective is to assess how well the HCRF performs in modeling the characteristics of a particular mental activity.

2.4 Segmentation and Labeling Performance Plot

The segmentation and label performance plot is the same tool as the temporal classification performance tool. The difference is determined only by the data being provided by the tool. Thus, the interpretation of the plot is context dependent: in the context of sequence segmentation and labeling, the tool depicts how well the model can track the ground-truth labels. In this case, both within a particular mental task and between mental task transitions. In our work, the LDCRF was used for segmentation and labeling of EEG signals.

2.5 Confusion Matrix

The confusion matrix implements the classification performance metric described in Section 4.3.2 of Chapter 4. The tool is important because it provides a systematic way of determining which mental tasks can be easily disambiguated using the models adopted.

The basic principle behind the confusion matrix is that of counting the number of observations that fall into a category which coincides with correct or incorrect classification. Ground-truth labels are required in order to facilitate this categorization and counting. Given N classes, the confusion matrix is set up as an NxN grid and the grid elements are related to the aforementioned categories. The elements that lie along the main diagonal of the grid represent the true positive classification results, and the off-diagonal elements represent misclassification results.

In an ideal situation, the model perfectly reproduces the correct classification results and the confusion matrix has non-zero diagonal elements and zero off-diagonal elements. It is also possible that the values are distributed evenly amongst the row elements of the matrix. This would represent the situation in which the model is producing what can be considered as random classification results. The worst situation that can occur is when the diagonal elements are zero-valued. Such a situation corresponds to the lowest possible accuracy rate, and is indicative of the scenario described in Section 4.3.2 of Chapter 4.[1]

One of the limitations of the tool is that it does not preserve the temporal characteristics of the data. What is meant by this is that from the perspective of the representation, two temporally different classification patterns can map into the same set of statistical properties. The analysis that we present in Section 4 is based on the training data confusion matrix and the test data confusion matrix. In the analysis, the mental tasks used were selected on the basis that their diagonal elements of the confusion matrices had comparably high values, and they

[1] In such a scenario, the model produces anything except that which it has been trained to produce.

had comparably low values in the off-diagonal elements in the confusion matrix. That choice is in-line with optimal model performance, as it maximizes the true positive classification rate and at the same time minimizes the misclassification rates, such as the false positive classification rate.

We present the confusion matrix in a graphic format in order to simplify the analysis of the results (see Figure 6.2). The visualization represents the information contained in the confusion matrix. The number of histograms corresponds to the number of classes. Each of the histograms represents the classification results corresponding to that particular class. If the classification performance were perfectly correct, then each of the histograms would have a single bar that has 100% selection frequency with an appropriate class label. That is, for "histogram 1" the first bar would reach 100%. Continuing with the logic, "histogram 2" would have the middle bar at 100%, a similar argument applies to "histogram 3". The observation that in non-zero elements appear in incorrect histograms indicates misclassification.

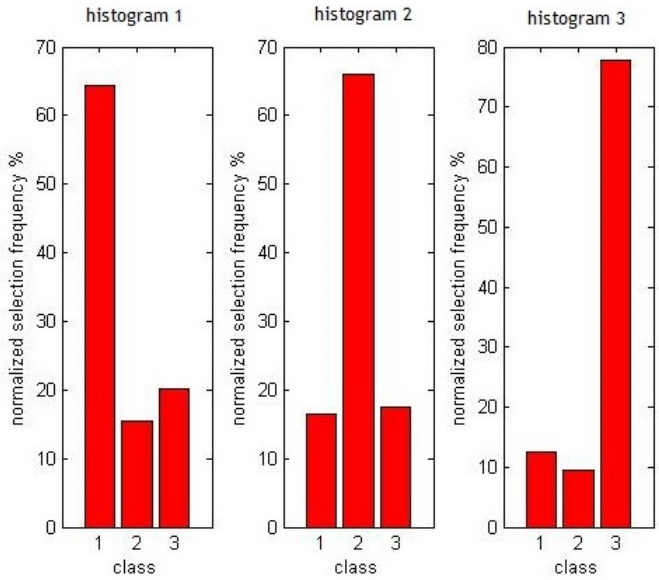


Figure 6.2: Confusion matrix results for sub2 corresponding to row two in Table 6.1.

3 Experiment Outcomes

In this section we discuss the experiment outcomes. The discussion starts with a presentation of irregularities in the data. These irregularities are reported only for the data obtained using the Emotiv headset (see Section 3.1). Thereafter, the model performance results for both the HCRF and LDCRF are presented in Section 3.2 and Section 3.3, respectively.

3.1 Data Irregularities

The significance of the fidelity of the data cannot be overstated, because the analysis of the performance is dependent on the quality of the data. EEG measurements can be corrupted by various factors. Some of these cannot be readily controlled, such as environmental conditions and sampling rate variations. One of the observations from the experiment sessions is that

the sampling rate variations heavily influence the amplitude measurement of the EEG signal and this ultimately reduces classification performance. It was expected that the choice of EEG technology would negatively influence the signal quality. However the degree of the variation depicted in Figure 6.3 was not expected.

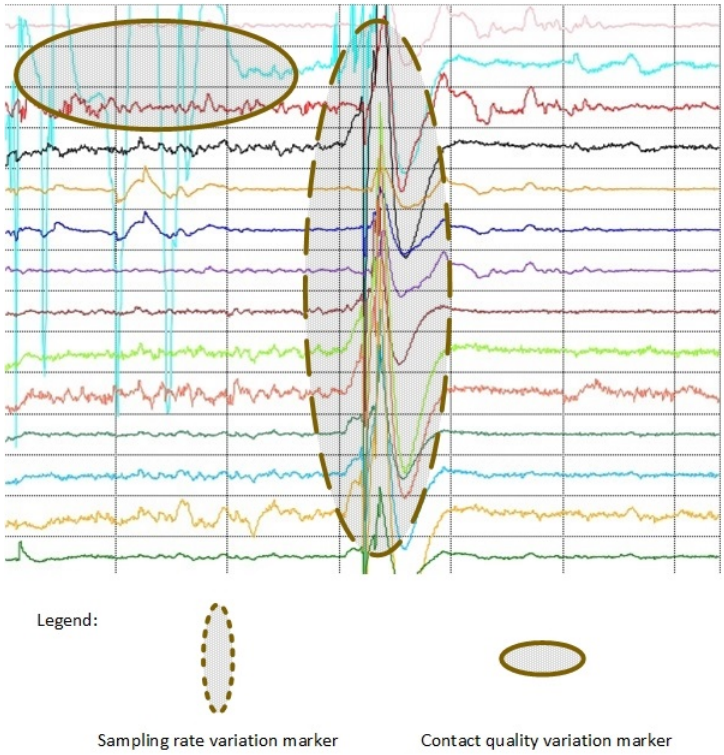


Figure 6.3: Effect of sampling rate and contact quality variations on the EEG measurements. Variations of this kind can be three orders of magnitude higher than the normal EEG activity, and negatively influence the classification performance of the models.

3.2 Classification Results

In this subsection we present the classification results obtained from the HCRF model. The HCRF was applied only to the EEG data obtained using the Emotiv headset. As discussed in Section 2.1.2 of Chapter 5, each mental task was performed separately from the other mental tasks. Thus the data-sets are separated with respect to the mental task. Given the structure of the data, determining the class label for the entire sequence was straightforward. However as outlined below, treating the entire sequence proved to be problematic. This problem required that an appropriate window size of the vector encoding state information be empirically determined. We found that the classification performance is greatest when the window size is chosen between 1 to 3 samples, and in our case this corresponds to 0.5 to 1.5 seconds worth of data.

The classification scheme provided by the HCRF is based on providing a single label for a sequence of inputs. The task that we are then faced with is determining the appropriate size of the window. Determining this is problematic two reasons. First, we expect that the dynamics associated with a particular mental task are repetitive over a timescale in which the user can perform the task multiple times. Taking this observation into consideration, it is clear that

Table 6.1: HCRF classification using Emotiv data. Each row represents an indepedent trial.

Subject	Test data classification accuracy(%)	Test data control accuracy(%)
sub1 (3 classes)	41.82	52.95
sub2 (3 classes)	71.29	81.61
sub3 (5 classes)	39.60	55.20
sub3 (5 classes)	31.79	51.67
sub4 (3 classes)	87.09	91.94
sub4 (3 classes)	34.81	65.19

the full 30 second interval shouldn't be used as a single training sequence such that the HCRF provides only one classification result over that entire sequence.

Second, there are transitions between mental tasks when online testing is taking place, and they do not necessarily occur at time intervals that match the length of the training sequence blocks. That is to say, at any given time during data collection the user may switch to a different mental task, either intentionally or not. In such an event, the HCRF provides poor classification performance because there is a mixture of incompatible classification labels in the sequence, in that some sequence elements belong to a different class and yet they are all classified to a single class. In both cases it is clear that the sequences have to be broken up into smaller segments. However, the HCRF doesn't provide a structured procedure to deal with the problem. This is one of the problems that motivate the use of the LDCRF. The command inference scheme described in Section 4.3.1 of Chapter 4 was applied to the classification results obtained from the HCRF and the results are shown in Table 6.1.[2] Applying the command inference in conjunction with the control accuracy metric discussed in Section 4.3.3 of Chapter 4 produced substantially better control signal classification accuracy. This can be justified using the arguments presented in the aforementioned sections.

3.3 Sequence Segmentation and Labeling Results

This subsection presents the classification performance results obtained from the LDCRF. The LDCRF was applied to both the BCI Competition and Emotiv headset data-sets. The results from both data-sets are presented below. In both cases we used a window of length three which is equivalent to 1.5 seconds of mental activity data. The length of the window was selected on the basis that the user would prefer to provide control signals as quickly as possible, and given constraints of the technology used. Additionally, the cognitive load that is associated with rapid control signal issuing is taken into consideration.

Inferring the control signal using the command inference technique did not produce considerable improvement in the accuracy of the LDCRF. This follows from the observation that the LDCRF does not produce the excessive variation in the classification labels, that is to say there is a higher degree of clustering in the classification results. It should also be noted that the discounting procedure is not applicable in this context, which follows from the observation that the chosen three mental tasks do not include a *rest* mental task which can be associated with

[2]This entailed using categorical binning and discounting low likelihood classification results.

"no control signal".[3] As depicted in Figure 6.5, the model has a tendency to stick to a particular task and not transition between them. Thus we see clustering in the symbols corresponding to the inferred class label. This finding does not however suggest that command inference is not useful in this context. We suspect that a more sophisticated form of command inference would be more appropriate for this context, as discussed further in Section 2.6 of Chapter 7.

Results From Emotiv Collected Data

The results obtained from the LDCRF using the Emotiv headset are presented in Table 6.2. The classification accuracy for both subjects is of the order of 40%. This performance is not satisfactory. However, it is better than the baseline rate of 33.34% which corresponds to a random classification strategy over three classes. There are various factors that contribute to this relatively low classification accuracy, as discussed further in Section 4. The foremost factors are the electrode placement provided by the sensor, and the quality of the signal acquisition as discussed in Section 3.1. Figure 6.4 depicts the confusion matrix results, and a visual illustration of the classification results as they appear in time is presented in Figure 6.5.

Table 6.2: LDCRF classification using Emotiv data.

Subject	Test data classification accuracy(%)
sub1 (3 classes)	39.42
sub2 (3 classes)	40.10

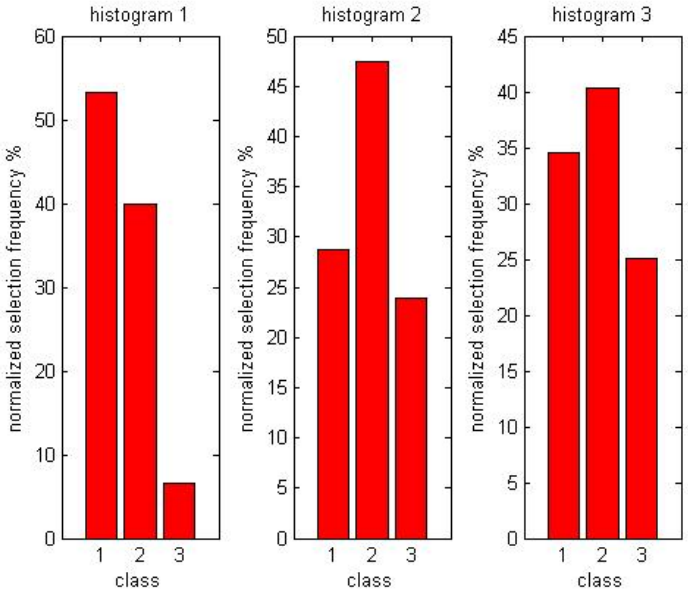


Figure 6.4: Confusion matrix results obtained using sub1 test data. The LDCRF doesn't perform well on the word generation mental task. The motor-imagery tasks appear to be discriminable, and the structure they encode has been learnt by the LDCRF.

[3]This discounting procedure is part of the control accuracy metric.

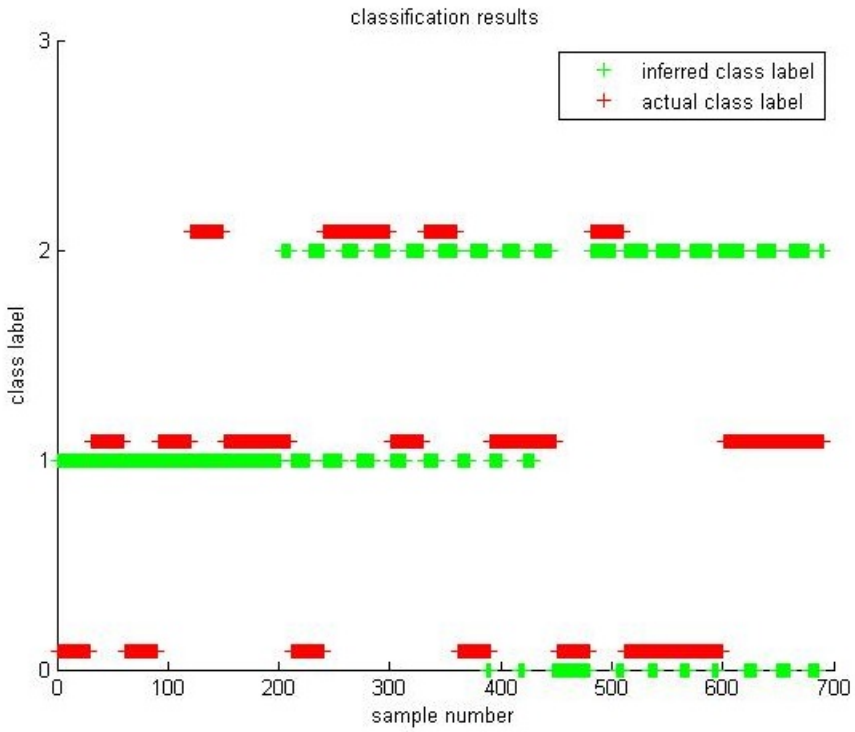


Figure 6.5: Temporal classification performance on Emotiv data for sub1. The plot represents the results obtained from the test data. The results obtained from sub2 are similar in form. The misalignment in the inferred and actual class labels indicated misclassification.

Results From BCI Competition Sourced Data

The results obtained from the LDCRF using the BCI Competition data are provided in Table 6.3. The classification accuracy for all subjects is in the range of 63.53% to 100%, with a mean classification accuracy of 82.31% as measured across all subjects. In contrast to the results using the Emotiv headset, the performance obtained using the BCI Competition data is satisfactory, which indicates that the general methodology adopted in this work is appropriate for the task.

Figure 6.6 and Figure 6.8 depict the confusion matrix results for subj2 and subj3, results corresponding to subj1 are not analyzed further since the results are perfectly correct. A visual illustration of the classification results as they appear in time is presented in Figure 6.7 and Figure 6.9. In these figures, it is clearly visible that there is significantly better agreement in the inferred and actual classification labels when compared to the results presented in Section 3.3. There is also a significant degree of stability in the classification results, in the sense that when a mental task is performed, the model does not produce oscillations between classes.

Table 6.3: LDCRF classification using BCI Competition data.

Subject	Test data classification accuracy(%)
subj1 (3 classes)	100
subj2 (3 classes)	83.41
subj3 (3 classes)	63.53

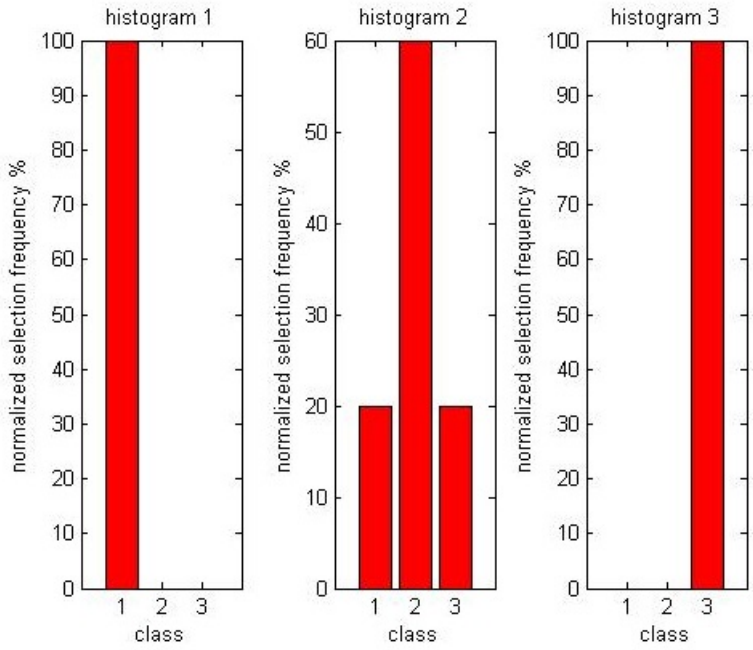


Figure 6.6: Confusion matrix results obtained using subj2 test data. The results indicate that both the first and third mental tasks are inferred with 100% accuracy when the actual class label coincides with any of those classes. Using Figure 6.7, it is clear that the classification error observed with the second mental task results is due only to the first two clusters in the data, beyond that the LDCRF produces error-free classification performance.

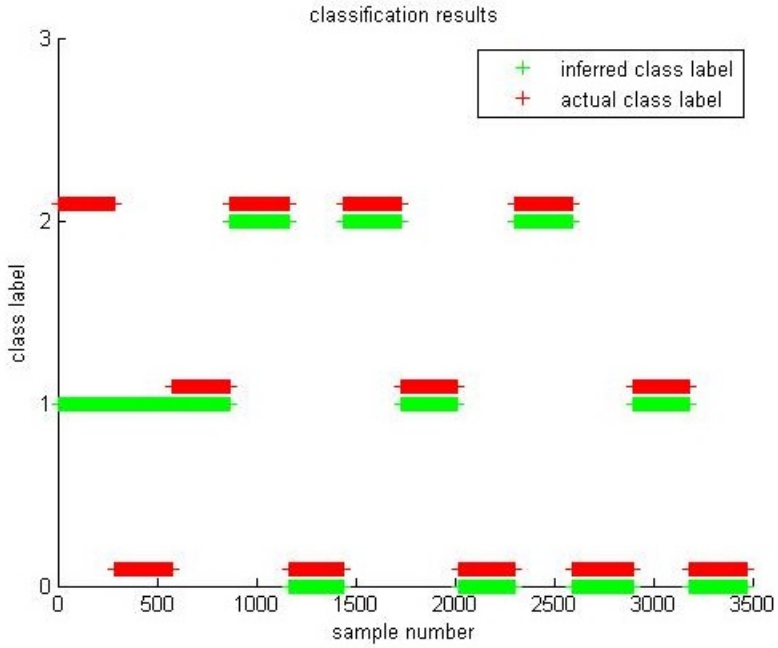


Figure 6.7: Temporal classification performance on BCI Competition data for subj2. The plot represents the results obtained from the test data. The misalignment in the inferred and actual class labels indicated misclassification

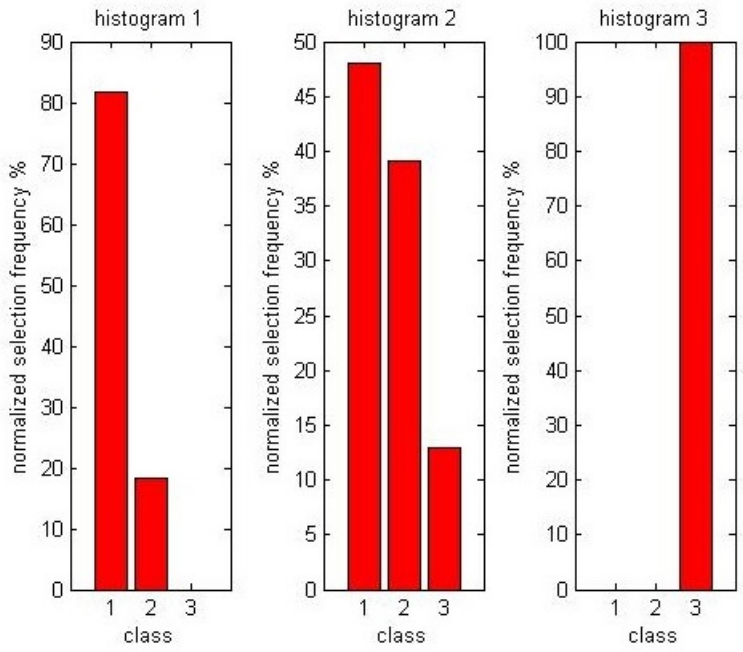


Figure 6.8: Confusion matrix results obtained using subj3 test data. The results indicate that the third mental task is inferred with 100% accuracy when the actual class label coincides with that class. A large proportion of the misclassification is attributable to the second mental task, in which the structure appears to not being learnt well by the LDCRF.

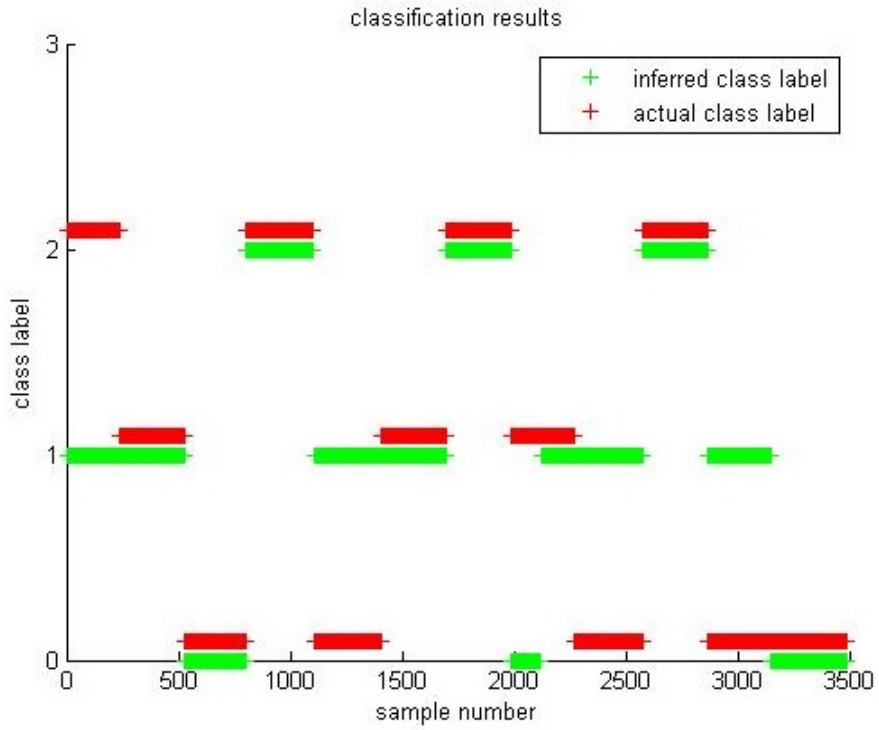


Figure 6.9: Temporal classification performance on BCI Competition data for subj3. The plot represents the results obtained from the test data. The misalignment in the inferred and actual class labels indicated misclassification

4 Analysis

In this section we present the analysis of the results that were obtained from the experiments. The primary objective of the analysis is to reveal insights concerning the performance of the probabilistic graphical models used in this study (Section 4.2). Another objective of the analysis is to assess the correspondence between our experimental results and standard notions from current neuroscience theory (Section 4.1).

4.1 Neuroscience Correspondence

In this subsection, we assess the correspondence between neuroscience theory and what we obtained using the Emotiv EEG headset. The correspondence is defined as the agreement between the expected observations and what was observed from the measurements obtained using the Emotiv headset. A complete correspondence would not only carry scientific import, it would also simplify the analysis of the data. Essentially, it would indicate that the sensor is operating as expected. Complete correspondence would entail the reinforcement of the epistemological foundation of neuroscience. However, it is not required in this work. Concerning the analysis task that is required to draw conclusions, we aimed to establish whether the Emotiv headset could be used to discriminate between the data corresponding to different mental tasks. Thus, it was important that we distinguish between the two cases presented below:

1. Complete correspondence and the model fails to produce the required results.

2. Minimal correspondence and the model fails to produce the required results.

The cases described above should be distinguishable in order to facilitate the drawing of conclusions given the results. The technique that was adopted in performing the disambiguation is the visual inspection of the TeSSTMaps visualizations. Given that the classification accuracy is low, we can then infer which of the two cases we are observing based on the correspondence of the cortical activation patterns that are expected and the ones observed from the data. If the first case is observed, then the failure can then be attributed to the techniques used to extract and model information from the EEG data. In such a case, a different set of techniques would have to be adopted. If the second case is observed, then there are two scenarios that would have to be considered, as presented below:

1. The low-cost EEG technology is failing to measure signals of significance, but the techniques are appropriate for the task at hand.
2. The low-cost EEG technology is failing to measure signals of significance, and the techniques are not appropriate for the task at hand.

The above described cases can be disambiguated by testing the performance of the models using EEG data obtained from clinical-grade EEG technology. This test was performed in this work, and the results are presented in Section 3.3. The test provides meaningful results since the clinical-grade EEG technology does measure signals of significance. Thus, if the performance of the models is unsatisfactory, then it can be inferred that the models are not appropriate for the task at hand and this would motivate the use of a different set of techniques. The results from all the case tests described above would facilitate the drawing of conclusions concerning whether the research hypothesis is true or false.[4]

In this work, minimal correspondence was observed between the expected cortical activation patterns and the cortical activation patterns observed from the measured data. However, the performance of the models is acceptable. Improvement in the performance of the models possibly can be attained. However, as discussed in Section 2 of Chapter 7, it need not be attained through discarding the models. In the discussion that follows, we present some of the factors that presented challenges in establishing the correspondence.

4.1.1 Hardware Limitations

Most of the limitations of the hardware became clear when the experiments were performed. A detailed description of the hardware limitations is presented in Section 4.1.1 of Chapter 4. Some of the problems were readily observable and permitted remedial action: one example of this is the electrode contact quality. Other problems did not lend themselves to remedial action, and thus their effects were much more challenging to deal with. An example is the variation in the signal sampling rate on the headset (see Figure 6.3). The observations suggested that there is still considerable scope for improvement regarding the design and manufacturing of low-cost EEG technology. However, such activities fall outside of the scope of our work.

[4]The cases in which the models succeed at producing the required results are not considered, because the success would imply that the technology works. Thus the research hypothesis would be true, but possibly for non-trivial reasons.

4.1.2 Inter-day Variations

The EEG data obtained using the Emotiv EPOC headset exhibited significant variation over timescales of a day for each of the subjects. The variation was investigated using the TeSSTMaps visualization tool and the classification performance of the HCRF model. It was noted that the classification performance of the model decreased when the data collected from a preceding session was used for training purposes and the data from the present session was used for testing purposes. The decrease in classification performance was significant, and typically resulted in classification performance that is worse than random. This problem was observed across all the subjects in the study. In this work, this variation problem was dealt with by collecting both training and testing data-sets in one day. This provided considerably better performance results, and these are the results reported in this work with regards to the Emotiv data-set. An alternative scheme, which can be investigated in the future, to deal with this problem is outlined in Section 2.7 of Chapter 7. With the given data it is not possible to determine to what extent the variation is due to the Emotiv EPOCH headset, since the BCI Competition data was also collected in one day. For reasons pertaining to lack of experimental control, further analysis is deferred.

4.1.3 Inter-subject Variations

There was a great deal of variation in the EEG data patterns obtained from the different subjects. The degree of variation is undesirable because treating subject-dependent matters is non-trivial. However, the inter-subject variations should be anticipated if neuroscience is used as background insight. An anticipated source of this variation is attributable to cortical folding, as discussed in Section 2.1 of Chapter 2. The main technique used to establish this inter-subject variation was performance validation of the machine learning models on the subjects on which they were not trained. The main observation was that the classification performance was drastically reduced on both the Emotiv EPOC and BCI Competition data-sets.

The results presented in Figure 6.10 and Figure 6.11 illustrate the changes in the performance of the LDCRF on the BCI Competition data. The assumption is that the clinical-grade EEG sensor has minimal effect on the variation of performance. Consequently, the variation is assumed to be indicative of factors that are beyond the sensing technology. The cross-subject validation was performed by using the LDCRF model that was trained using the EEG data from subj1, this was the same model that provided error free performance when applied to the test data from subj1 (see Table 6.3). The test data used in the cross-subject validation was the test data from subj2. It is clear that the performance observed in this cross-subject validation is worse than random. At first glance, Figure 6.11 seems to suggest that the true positive classification accuracy for mental task 3 is 100%, as read from 'histogram 3". However, this is misleading because the model appears to be consistently producing the same result for almost all the data samples, regardless of the actual class label (see Figure 6.10).

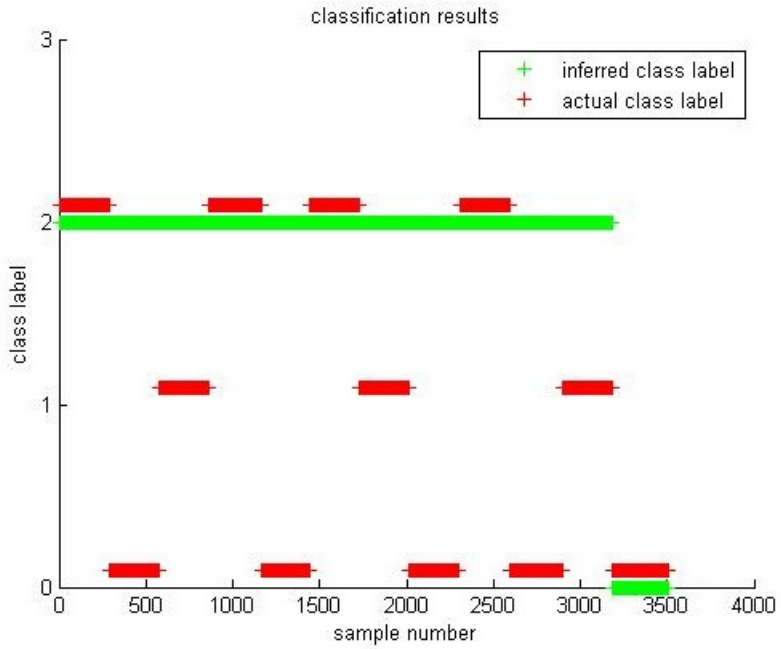


Figure 6.10: Temporal layout of classification results. The figure indicates that the model consistently infers the same class label for almost all data samples. It should also be noted that only two of the total three classes are inferred throughout the whole process.

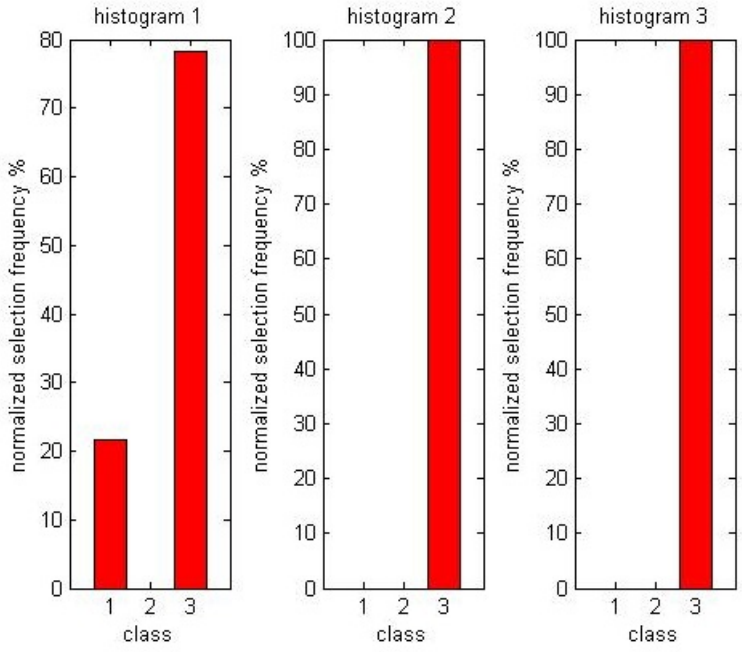


Figure 6.11: Confusion matrix results obtained from the cross-subject validation. The results indicate that the model doesn't infer one of the classes at all, either correctly or incorrectly. The 100% true positive accuracy noted in one of the classes could be spurious.

A few factors that were noted to have a reinforcing influence on the inter-subject variations

on the Emotiv EPOC data are described below.

Hair characteristics The density of the subject's hair negatively influenced the contact quality in the electrodes. Although it was possible to move the hair away from the electrode positions on some of the subjects, with the rest it was not possible to do so. A partial solution was to have the hair shaved off on some subjects, but it was not possible to have all the subjects shave off their hair.

Head shape The curvature of the subject's head negatively influenced the contact quality in the electrodes. This resulted from the fixed structure for electrode placement provided by the Emotiv EPOC headset. In order to mitigate the impact of this problem, the space that appeared between the electrodes and the head was filled with a piece of cotton-wool that was soaked with the saline solution used on the electrode contact pads in order to facilitate conductivity.

4.1.4 Control Feedback Effects

Control feedback appeared to have a destabilizing effect on how well a subject could focus on a given mental task. This problem is not unique to our investigations. Pfurtscheller et al. [2004] briefly discusses the effects of the higher cognitive load on the user of an EEG based BCI with an asynchronous protocol. However, further discussions pertaining to the problem do not feature in the literature. It is also noted that there is a much larger volume of work directed towards the synchronous protocol, in which the higher cognitive load would not have a substantial effect on the results. This follows from the observation that with a synchronous protocol, the users have a recurring time-frame in which they can rest or perform any other mental task without influencing the current state of the information processed by the control logic in the BCI. This stands in stark contrast to our work, in which such a filtering window does not exist.

In this work, the subjects reported having difficulties with focusing on a particular mental task and keeping track of the button which has been assigned focus at that moment (i.e. the highlighted button on the text input interface). All the experiments for which we report results were thus performed without control feedback. The results obtained from the BCI Competition data are also of this nature, as discussed in Section 2 of Chapter 5.

4.2 Model Performance Analysis

In this subsection, we present the analysis of the performance of the HCRF and the LDCRF, respectively. Regarding the former, only Emotiv EPOCH data was used to assess the performance of the model.[5] Regarding the latter, both Emotiv and BCI Competition data were used to assess the performance of the model. The BCI Competition data was used to benchmark the performance of the LDCRF, because it was collected using clinical-grade EEG technology. Thus it constitutes our validation scheme for our research methodology. The BCI Competition data

[5]It was sufficient for us to validate that the HCRF could represent some aspects of the structure of the EEG signals, even if it meant the transition dynamics are not considered. Non-random performance on Emotiv EPOCH data is sufficient to infer that non-random performance would be obtained if clinical-grade EEG technology were used.

also provides a reference for the results obtained using the Emotiv headset, so as to provide insight pertaining to the performance of the models used in this work.

Hidden Conditional Random Fields

The HCRF provided acceptable performance results. As can be seen in Figure 6.2, the HCRF was capable of learning the internal structure of the data associated with each mental task. Although Figure 6.2 is not representative of all the classification performance results, it did however indicate that it is appropriate to proceed to the next phase in which the LDCRF was used.

There is a great deal of variation in the classification results obtained from the HCRF. It is not clear at this stage what the source of the observed variations is. Observations made with the assistance of the TeSSTMaps tool indicate that there are instances in which there is a match in the representation of the activity of different mental tasks (see Figure 6.12). It is under such circumstances that the HCRF is liable to producing poor classification performance. An assessment of the probability distribution also indicated that the HCRF produced almost equivocal probabilities across different classes when misclassification was observed.

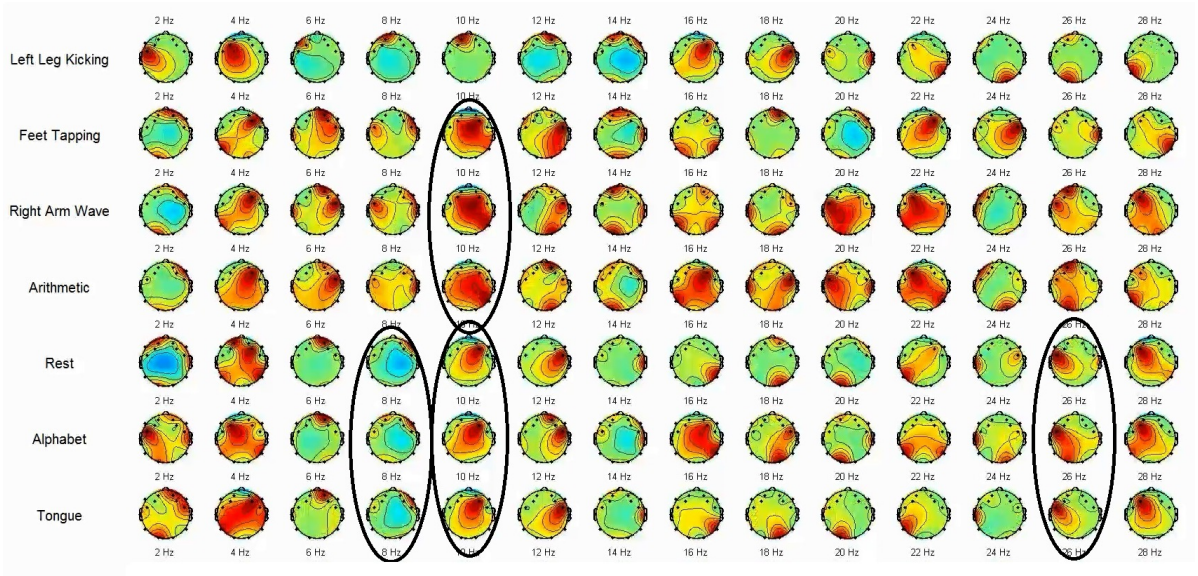


Figure 6.12: In some samples there were similarities in the measurements, as highlighted with ovals in the image. The samples were collected in one session and from one subject. Each mental task was performed independently of other mental tasks, consequently they do not share the same timestamp. It is a matter of coincidence that they appear in the same frame on the TeSSTMaps visualization tool. Thus the degree of similarity depicted here is a gross underestimate of the true amount of similarity between the sequences. All the tasks were mental, and that means that the motor tasks were imagined.

Latent Dynamic Conditional Random Fields

The LDCRF provided very good performance on the BCI Competition data (see Section 3.3), and acceptable performance on the Emotiv EPOC data (see Section 3.3). Overall, the model proved to be an appropriate tool to be used for sequence segmenting and labeling. Our findings

suggest that our research hypothesis is true. There are significant discrepancies that were observed in the performance of the LDCRF between the above mentioned data-sets. In the discussion that follows below, we provide some of the factors we suspect are responsible for this discrepancy. A note of significance is that our validation procedure that made use of the BCI Competition data provided results that corroborate the results presented by Saa and Çetin [2013]. This essentially means that our investigation is methodologically sound, in particular by assuming that EEG signals encode measurable structure and that machine learning would facilitate the extraction and representation of that structure.

First, there are significant differences in the nature of the data collected using the Emotiv EPOC headset and that obtained from the BCI Competition. In particular we highlight that the montages are different. Different montages produce results that are not directly comparable because the EEG measurements are collected in different regions and with different electrode couplings. One particularly important aspect concerning observing the differential cortical activations corresponding to the left or right hand side motor imagery tasks relates to the C3 and C4 electrodes. These two electrodes were used in the BCI Competition experiments, but were not used in the Emotiv headset experiments because they are simply not available on the hardware. This topic is further developed in Section 2.4 of Chapter 7.

Second, the data quality itself is noticeably different. Data quality issues are discussed in Section 3.1. It is clear in the discussion that the observed effects should negatively influence the performance of the LDCRF. Our findings confirmed the validity of this expectation. Lastly, Saa and Çetin [2013] reports the use of the sequential floating forward selection algorithm in order to select the optimal set of features to be used for state vector construction, and this is presented in Section 2.2 of Chapter 3. We have not made use of such an algorithm. Our approach was motivated by the frequency values found in the neuroscience literature, as described in Section 2.1.3 of Chapter 5.

5 Discussion and Conclusion

In this chapter, the experiment results and the corresponding analysis of the results was presented. We highlight key topics developed in the chapter data visualization, experiment outcomes, and analysis. With regards to data visualization, we presented the primary visualization tools that we used in this work (see Section 2). The tools include: a visualization of the temporal evolution of the EEG signal as presented in time series format (see Figure 6.1), the TeSSTMaps visualization tool for assessing the temporal structure of the classification results (see Figure 4.7), and the confusion matrix for determining the mental tasks which have resulted in clustering in the classification results (see Figure 6.2).

Pertaining to the experiment outcomes and analysis, we first showed that there were data irregularities that manifested during experimentation (see Section 3 and Section 4). The irregularities were primarily related to variations in the sampling rate on the sensor and the variation in the contact quality of an electrode. Thereafter, we presented the experiment results obtained from both the HCRF and the LDCRF models, respectively. Regarding the HCRF, we performed experiments only on data collected using the Emotiv headset. Although there was a great deal of variation in the classification performance, it was nonetheless shown that the model was capable of learning the internal dynamics of each of the data corresponding to different mental tasks from the selected set of mental tasks. Consequently, the classification performance was better than that which would be equivocally random across the classes. Furthermore, we

presented sequence labeling and segmentation results obtained using the LDCRF. The LDCRF was applied on data collected using the Emotiv EPOC headset and data obtained from the BCI Competition, respectively. Regarding the former, we showed that LDCRF provides performance which is better than that which would be obtained using an equivocal random labeling strategy across the classes. The latter data was used for methodological validation, and the outcome indicated that our methodology is sound. We concluded that the research hypothesis is true.

The analysis also indicated that there was minimal correspondence between the data that was collected and what neuroscience theory suggests we should expect (see Section 4.1). Various factors are presented describing why this is the case, and how we can still draw meaningful inferences related to the research hypothesis given this problem. The discussion includes the inter-day variations that were observed in the data. There is a good reason to assume that this is a problem that is particularly accentuated by the EEG technology used in this work. We also discuss inter-subject variations that were observed. Lastly, we discuss the effect of control feedback. In particular, we highlight that subjects encounter difficulties when they have to handle the higher cognitive load. This topic is developed further in Section 2.3 of Chapter 7.

The next chapter represents the last chapter of this document. The chapter represents the research conclusion, which provides an overall summary and the future prospects of the work. Some of the key findings from the research are highlighted. This includes the insights gathered from performing the research and the unexpected results encountered. The chapter ends with a broad outline of some of the work that still has to be done in order to make the technology viable. Most of the discussed problems were directly encountered in this work, while others are identified as important from considerations that are beyond the scope of our current work.

Chapter 7

Research Conclusion

Brain computer interfaces (BCIs) are a promising new technology for human-computer interaction. People who have lost the ability to perform tasks that require neuro-muscular connections stand to benefit the most from technologies of this kind. Much like the wheelchair facilitated social reintegration for its users, BCI technology may prove to do the same, and perhaps in a much more fundamental way. At the family unit level, BCI technology could change the typical situation that arises when people suddenly becomes incapable of taking care of themselves. Although a care-giver may still be required, the function of the care-giver may change from the traditional notion as the responsibilities move from the care-giver to the disabled person. With a broader outlook it is clear that the success of BCI technology could fundamentally change how government policy is structured, both in how financial resources are handled and how employment equity policy is structured. With the development of BCI technology, we hope to improve the quality of life for disabled people by providing them with the opportunity to live productive and fulfilling lives. Perhaps more importantly, facilitating full reintegration into society should be a priority.

In this work we attempted to create a low-cost EEG based BCI that is based on an asynchronous protocol. To make this possible, we extracted key insights from neuroscience concerning the structure and physiological functions of the human brain. Our main interest in this work was the temporal dynamics of the cortical activity. Our observation scheme depended on the related EEG activity that manifests when certain mental tasks are performed. We also made use of signal processing techniques in order to process data that has complex spatio-temporal structure. To achieve this, we employed time-frequency analysis in order to extract the frequency dependent structure in the EEG data, thus revealing more elaborate spatio-spectral dynamics that evolve in characteristic ways as time progresses. Our attention then shifted to machine learning, in which we discussed concepts that allow a computer program to determine what the user intends to do. This is achieved by automatically learning the association between input data and the desired output, which our work are power spectra elements and classes, respectively. In particular we made use of the latent dynamic conditional random fields model in order to perform segmentation and labeling of EEG data.

Our research also required that we discern what other researchers in the field have achieved with regards to solving problems similar to those that our work aimed to address. In the literature, we found studies related to what we set out to achieve in our work. However, there were differences in how they defined the problem, solution, design principles to be followed, and the implementation of the solution. Some of the related studies indicate success in the BCIs

they created, given the success metrics that they defined themselves.

In the subsequent sections of this chapter, the overall research findings are discussed and proposed future work is outlined. Regarding the former, key findings are discussed first and this includes the insights gathered from doing the work, and unexpected results found during experimentation. Subsequent to that, a discussion of the performance of the BCI and a brief discussion concerning the research hypothesis is provided. The chapter concludes with an outline of future work that still has to performed in order to make this technology usable outside of the laboratory. Although there is considerable work that still has to be done on developing better hardware using low-cost EEG technology, most of the problems that are discussed were also encountered by other researchers in the field who make use of clinical-grade EEG technology. Primarily, the problems that can be addressed through the development of more sophisticated software are discussed below.

1 Research Findings Discussion

In this section, we present the discussion of the research findings obtained in this work. In particular, we discuss the key research findings (Section 1.1), and the general conclusions reached through analysis (Section 1.3).

1.1 Key Research Findings

The key findings pertaining to the software components of this work are outlined below:

- The application of probabilistic graphical models, and in particular the latent dynamic conditional random fields produced acceptable segmentation and labeling performance. However, the performance on Emotiv data is not of high quality.
- The current implementation of the cortical activation visualization tool (i.e. TeSSTMaps) is not suitable for detecting problems timeously. At present, certain irregularities are noticed during the analysis stage, and by that time the subject is not available to assist in trial repetition. A candidate solution to this problem is the creation of a realtime visualization tool as discussed in Section 2.1.
- The overall software framework that was created to facilitate experimentation performed well, and will provide a good platform for future work. In particular, the signal acquisition and processing was performed sufficiently fast enough such that a delay during runtime was not noticeable. The modular architecture of the system also facilitated the removal and addition of certain modules without having to make modifications to other modules.

The key findings pertaining to the hardware components of this work are outlined below:

- The fixed electrode placement layout of the device imposes considerable limitations on experimentation. We found that almost all of the work reviewed in the literature made use of the C3 and C4 positions, both of which are very effective in disambiguating between motor activity associated with either the left or right hemisphere of the brain. A more flexible electrode placement layout is required. Not only would it provide better classification results, but it would also facilitate a more direct comparison between results obtained from other researchers and results that we obtain.

- The device provided good performance with regards to the battery life, but the signal acquisition produced very noisy data. The former factor is essential for extended operation outside of the laboratory environment. The latter is required for creating a BCI that functions reliably, and there is considerable scope for improvement on the signal quality.
- There are some durability problems associated with the hardware. In particular, we experienced problems associated with the electrode components of the device, as outlined below. There is considerable room for improvement in this regard.

1.2 Unanticipated Outcomes

There were quite a number of unexpected results that appeared as the work was being conducted. Almost all these results are related to the hardware used in the work. A list of the most important of these results is presented below, including one notable result obtained from the latent dynamic conditional random fields (LDCRF):

- The Emotiv EPOC headset used in this work exhibited low durability, and this is problematic when everyday usage of the device is required. The locking mechanism of the electrode easily wears-out due to the design and the materials used in the mechanism. Not much could be done to avoid this problem. One of the reasons is that the manufacturer recommends that the electrodes should be removed after use, as part of the electrode metal disk maintenance plan. This necessitates that the mechanism be opened and closed multiple times, thus contributing to the wearing-out of the parts.
- The EEG measurements obtained using the Emotiv headset were of a lower quality than expected. It is suspected that the main driver underlying this phenomenon is the degrading quality of the electrodes. There are two factors in this regard. First, the metal caps appeared to be undergoing an electrolysis process, and this resulted in the formation of a hard blue-green substance which we suspect is copper sulphate. Second, the locking mechanism as described in the preceding point presented its own challenges. One of the problems that was observed is the power spectra of the EEG data contained unexplained spikes which went up as high as three orders of magnitude when compared to the average value.
- The results obtained from the LDCRF exhibited very stable temporal structure, which stands in stark contrast to the results obtained from the hidden conditional random fields (HCRF). Command inference applied to the LDCRF output had negligible impact, which follows from the minimal variation that was observed in the LDCRF output.

1.3 Conclusion Concerning Performance

The overall findings in the work suggest that the theoretical insights that been selected to formulate and solve the problem are appropriate. Regarding the hardware platform, the results also indicated that it is possible to use low-cost EEG technology for asynchronous BCI applications. Although the performance of the overall system is currently not satisfactory, we conclude that our research hypothesis is true. For future work, we propose software-oriented interventions aimed at improving the overall performance of the system as discussed in Section 2. Hardware-oriented interventions are also required as there is a strong requirement towards making the hardware more durable and reliable.

2 Proposed Future Work

In this section, we provide an outline of future work that has to be done in order to make BCI technology a viable and commonly used technology. Our focus is limited to non-hardware challenges. Most of the challenges were directly related to our work, but some are more general in scope and have been discussed by other researchers working in the field. Some of the latter challenges have been featured as standard problems in the BCI Competition. Our intention is to develop algorithms that will be applied to BCI problems in order to handle data processing and machine learning problems. The list of problems below is by no means exhaustive. However, the provision of the corresponding solutions would be useful to further development of the field.

2.1 Realtime Cortical Activation Visualization

Data visualization is critical for the purposes of gaining further insight into the problem studied in this work. For example, it is significantly easier to take note of structural dynamics in EEG data when the spectral domain is used, as opposed to the temporal domain. Pfurtscheller et al. [1976] discusses structural dynamics of event-related cortical desynchronization. Desynchronization information is easily discernable in the frequency domain, as opposed to in the temporal domain. One of the problems in BCI technology is data visualization, in particular, realtime data visualization that is easy to interpret. As a candidate solution, we propose a realtime cortical activation visualization tool with an intuitive interface for the user.

One of the problems that would have to be addressed is determining an appropriate representation of the temporal, spatial, and spectral properties of the data. In Section 4.1.2 of Chapter 4, we presented our attempt aimed at addressing the three aspects outlined (see Figure 4.7). TeSSTMaps is a cortical activation visualization tool. The representation does convey the information pertaining to the temporal, spatial, and spectral properties. The interpretation is also fairly straight-forward, as there is a direct mapping into how we perceive the progression of time, the spatial layout of the cortical activity, and the spectral distribution of the cortical activity. However, it lacks the required realtime visualization capability, and that is problematic.

One of the main benefits of having a tool of the proposed kind is that the data can then be assessed by the researcher in realtime. This would allow the researcher to correct for problems as and when they arise. This is essential for reducing the amount of time wasted during experiment sessions. In a lot of cases, this time wastage only becomes apparent during the post-experimentation analysis session. The finding is usually that there were problems with data collection, or that the expected neurophysiological response was not observed.

2.2 Automated Mental Task Selection

The functionality of BCI technology is fundamentally based on the distinguishability of the input signals. It follows then that selecting the most distinguishable mental tasks is of high importance. At present, insight on this is provided by neurophysiological arguments. The idea is that mental tasks which are most appropriate are those that produce different cortical activation patterns. This is a reasonable first attempt, but further considerations raise numerous concerns. Some of the concerns include defining an appropriate measure of similarity, dynamically determining the most distinguishable mental tasks, and dealing with informational degeneracy. In

this discussion, we focus on informational degeneracy.[1] At present, the degeneracy cannot be eliminated as it appears to be an intrinsic property of the phenomenon being studied. The degeneracy arises from various factors, but amongst the critical ones we find the factors discussed below:

Signal aggregation The electrical activity that results in the manifestation of EEG signals undergoes aggregation according to the superposition principle. This means that different cortical regions can be activated by different mental tasks, however the resulting EEG activity may appear to be the same. In order to partially resolve the degeneracy, source localization techniques may be used. Unfortunately, the resulting problem is not trivial. The main argument here is that mental tasks should not just be motivated by neurophysiological insights. Experimental EEG data also has to be taken into consideration. Thus the pre-specification of distinguishable mental tasks is difficult to implement.

Cortical folding The effects of cortical folding cannot be neglected. Current knowledge suggests that cortical folding is unique across individuals, and that affects functional localization in the human brain [Brett et al. 2002]. The main argument here is that it is difficult to specify a standard set of mental tasks that will consistently work across all individuals. In order to mitigate the impact of the above problems, we propose an automated mental task selection scheme. An automated scheme of this sort would assist in mitigating the degree of intervention from technically trained personnel, and thus facilitate wide-spread use of the technology outside the laboratory. As a first attempt to the solution, a statistical variation measure on the information conveyed by the TeSSTMaps tool could be a suitable candidate.

2.3 Control Feedback Effect Assessment

In the literature, there is a considerable amount of work in which the classification of EEG activity is performed without providing control feedback to the user. This is especially the case when mental imagery is used as the basis for control. This stands in contrast to the amount of work directed towards problems in which the classification of EEG activity is performed with control feedback.

The studies that we have reviewed containing control feedback report BCIs that have considerable latencies. Examples of such work are those of Pfurtscheller et al. [2003] and Roula et al. [2012]. Pfurtscheller et al. [2003] provides a short discussion of the problem. In particular, Pfurtscheller et al. [2003] highlights that BCIs that use asynchronous protocols induce a cognitive load that is higher than that of BCIs that make use of synchronous protocols. Upon a high-level reading of the situation, it appears as though the effects of control feedback have not been studied extensively. We suspect that there is currently an insufficient understanding of the effects of higher cognitive loads in the BCI context.

We propose preliminary research that would be aimed at providing some insight into the problem. As a first step, the assessment could be directed towards studying the classification

[1] Informational degeneracy pertains to information that is different in a fundamental sense, but the difference cannot be resolved due to certain constraints. An example of this can be found in statistical mechanics, in which energy levels have degeneracy in the sense that different ensemble configurations map into the same energy level distributions.

accuracy when there is no control feedback. The second step would involve studying the classification accuracy when the user is required to actively track the movement of an on-screen object, while performing particular mental tasks. A comparative analysis could then be performed on the obtained results.

2.4 Minimal Montage and Channel Selection

An interesting problem that deserves further investigation concerns selecting a montage with the minimal number of electrodes. This selection necessarily entails the selection of channels, and affects the choice of mental tasks that ought to be used. The minimization also has clear economical ramifications. Another problem would be that of automating the process. Pfurtscheller et al. [2005] reports findings on the former problem, and some of the research findings are discussed below.

The single-channel analysis uncovers the relative importance of each electrode position [Pfurtscheller et al. 2005]. Pfurtscheller et al. [2005] reports that the dominance of electrodes in the single trial analyses that overlay the sensorimotor and pre-motor areas, and especially the hand representation area, confirms the modulation of sensorimotor rhythms during motor imagery. Pfurtscheller et al. [2005] also discusses other findings, in which the best separability between left and right hand motor imagery is based on signals recorded from electrode positions C3 and C4. There is currently not much flexibility in determining an appropriate montage in the context of low-cost technology. Consequently, this constrains the choice of mental tasks that can be reliably discriminated. The Emotiv EPOC headset is an example of a technology that has a fixed montage. Unfortunately, it excludes the C3 and C4 electrodes that have been extensively used for motor-imagery problems. The C3 and C4 electrodes are well-suited to motor-imagery problems because those electrode positions correspond to the motor cortex. Variations induced by motor-imagery are easier to discern using these electrode positions [Pfurtscheller et al. 2005].

There are two aspects to this work. The first pertains to the hardware specifications. The second concerns the specification of an algorithm that can be used to select an appropriate montage, given a set of mental tasks. Concerning the former, the basis could be clinical-grade hardware, as it usually has a large number of electrodes and good coverage on the scalp. Concerning the latter, the basis could be a combinatorics inspired machine learning algorithm that would test the performance of the model, given different sets of channels.

2.5 Unsupervised Feature Learning

One of the problems when working with EEG data is that the data is represented in a space with a relatively high dimensionality. The high dimensionality is a result of the number of channels and the corresponding frequency decompositions that define the space. High-resolution analysis of the data leads to a very high-dimensional feature space [Garrett et al. 2003]. The main problem that arises is the extraction of insights from the data, since visualization becomes a non-trivial problem. There are various techniques that are currently used to reduce the dimensionality of the space. Effectively, these techniques implement truncation of potential unknown information.

First, neurophysiological arguments are used to specify what ought to be considered. However, this has limitations as exemplified by Pfurtscheller et al. [2003] in the subject-dependent choice of EEG bands, electrode positions, and mental tasks. Second, certain aspects of the

data are projected into low-dimensional spaces in order to facilitate visualization. For example, consider the TeSSTMaps tool that we presented in this work. The problem with the two aforementioned approaches is that the complex structure that becomes apparent after non-trivial transformations remains obscured. Consequently, the efficacy of the feature selection process is limited. To address this problem, we propose the use of unsupervised machine learning techniques in order to discover non-trivial features. The efficacy of those features would then be tested using the supervised machine learning techniques discussed in this work. Another important assessment would be the temporal stability of those features. This assessment would provide insight on whether the features were picked by chance and they are spurious correlations, or if there is some fundamental significance to those features.

2.6 Intra-subject Hierarchical Command Inference

The topic discussed here is an extension of the command inference that was discussed in Section 4.3.1. An extension to the command inference essentially corresponds to adding successive layers of intelligence to the classification results generated by the model. These successive layers are essential for enforcing stability in the command inference process. The general idea is that they should reduce the amount of variation in the classification results. Control accuracy is largely determined by the underlying classification accuracy. However, in the context of BCI technology, control accuracy is primal. These layers become even more important when the quality of the EEG sensor is not very high, such as when low-cost EEG sensors are used. Using economic arguments, it is clear that universally accessible BCI technology is more likely to arise from low-cost EEG sensors than from clinical-grade EEG sensors. Hence, techniques that would facilitate the extraction of maximal value from the former would depend on data processing techniques that go beyond classification.

We propose that the benefits of training machine learning models on the classification results that are produced by the model operating on EEG data be investigated. There are various candidate models that can be used in the investigation, including some of the probabilistic graphical models in this work. In particular, it would be interesting to assess the performance of generative models such as the HMM. Generative models would provide a rigorous approach to early-stage prediction of multi-action-dependent user intention by sampling from the model. The predictive capability would be used to suggest sequences of commands.

2.7 Inter-session Intra-Subject Transfer Learning

Physiological data is known to manifest non-stationary dynamics which make it difficult to ascertain the behaviour of the system as time progresses [Mayer-Kress 1994]. One of the problems with the non-stationarity is that data collected in the past may become irrelevant for future purposes. Physiological data is influenced by a multitude of factors which include emotional state, fatigue, and diet. EEG data also exhibits variations when the same mental activity is performed at different times. The variations can be observed even within several hours. The differences appear to become more pronounced when the duration between the sessions is increased. This is particularly problematic in the context of BCI technology as it necessitates frequent retraining of the models.

There is some work that is directed at solving this problem. Some of the studies and the

corresponding results are part of the BCI III Competition, under the Data set IVc challenge.[2] However, current efforts appear to resort to recalibration. It is doubtful if this problem can be completely alleviated. Depending on the philosophical stance that is assumed, it may be unlikely that the brain returns to a state similar enough to one of the past states. However, there are various techniques that could be examined. One technique would be to attempt to find features that exhibit slow variation in time. These features could be possibly identified with the assistance of unsupervised machine learning techniques. There is no guarantee that the techniques would uncover something useful, but the search is meaningful. When standard neurophysiological insights are taken into consideration, there are features that are expected to be stable.[3] Examining the temporal stability of the features would be straightforward.

2.8 Inter-subject Transfer Learning

Inter-subject transfer of some learnt attributes is interesting in the BCI domain. There are various reasons why this transfer capability is important. For example, suppose we have a large data-set that we can use to train a model and that the data was collected from different individuals. A problem that arises pertains to finding a transformation that would allow the data to be used to bootstrap the training process for a new user. On a first reading of neuroscience theory, it appears as though the parameters (e.g. EEG bands) to be used for BCI purposes have been determined. However, various studies including Pfurtscheller et al. [2003] show some deviations from the theoretical parameter values.

We posit that there should be transformations that would provide the required inter-subject transfer. There have been some attempts in this regard. Wang et al. [2004] and other researchers in the BCI Competition III provided techniques that produced positive inter-subject transfer results. In particular, Wang et al. [2004] provided a solution based on ensemble classifier and bootstrap aggregation, and the solution provided an accuracy of 94.17%. The outcome based on chance would have produced an accuracy of 50%. Integrating this capability in future work could provide significant improvement in reducing training time.

3 Conclusion

In this research we endeavored to create a brain computer interface based on the asynchronous protocol. Our primary objective is to make the technology accessible to people of average income means, and we propose to achieve this by adopting low-cost EEG technology. Low-cost EEG technology can be an order of magnitude cheaper than clinical-grade EEG technology, but there are some trade-offs between the cost and the quality of the technology. The research hypothesis we aimed to test pertains to investigating the feasibility of reliably classifying three mental tasks using the Emotiv headset. We make use of the fast Fourier transform algorithm for signal processing, and the latent dynamic conditional random fields model to perform the segmentation and labeling of the EEG data. Our research findings suggest that the test of the

[2] http://bbci.de/competition/iii/results/index.html#berlin3

[3] A prime example are event-related potentials in EEG measurements.

research hypothesis is true.[4]

In order to achieve the above stated objective, we made use of insights from neuroscience, signal processing, and machine learning. The exact technical preliminaries from these fields required in this work are presented in Chapter 2. First, the chapter discusses the neuroscience insights. In particular, the differential cortical activation depending on mental task being performed, and the EEG activity that manifests when that mental task is performed. Second, the chapter discusses the signal processing techniques required to extract the dynamical spatio-spectral structure in the EEG data. In particular, we discuss the FFT algorithm and the power spectrum which are used to construct an abstract mathematical object that represents the state of the brain. This mathematical object is an one-dimensional vector that has elements which are specific bands of the power spectra of each electrode. Lastly, the chapter discusses probabilistic graphical models which are used to model the structure that is represented by the state information. The models are used to provide an abstract classification result that can be used to invoke a set of instructions that the computer has to execute.

In the subsequent chapters we presented some of the related research, in which we discuss some of the work that appears in the brain computer interface field (see Chapter 3). In Chapter 4 we discussed the framework required to test our research hypothesis, and this entailed designing and implementing various software components of the overall system. In Chapter 5 we discussed the exact experiment procedure that we followed when performing the experiments. The chapter also includes a description of our experiment paradigm and the experiment paradigm used to provide the BCI Competition data we made use of for validation purposes. In Chapter 6 we present the experiment results obtained from both the data we collected using the Emotiv headset and the data obtained from the BCI Competition. The chapter also presents the analysis of the results, and we argue that the general methodology that we used in the work is appropriate for the task at hand. The last chapter discusses the general research findings. In particular we discuss some of the problems that were encountered when the experiments were performed and what was done to address those problems. The chapter also provides some direction on the future work that still has to be performed in order to improve the performance of the system, and to facilitate the transfer of this technology from the laboratory environment to the end-user environment.

[4]Hypothesis statement: *At least 3 independent mental tasks can be asynchronously and reliably classified using the Emotiv epoch headset if feature vectors are extracted using frequency decomposition techniques, such as FFT, and classified using probabilistic graphical models, such as HCRF and LDCRF.*

Chapter 8

References

[Accord 2015] Machine learning framework.
http://accord-framework.net/. Last accessed 11 March 2015.

[Autoit 2014] Automation and scripting language.
https://www.autoitscript.com/site/autoit/. Last accessed 11 March 2015.

[Adeli 2010] Hojjat Adeli and Samanwoy Ghosh-Dastidar. Automated EEG-Based Diagnosis of Neurological Disorders. Taylor and Francis Group, LLC 2010.

[Brett et al .2002] Matthew Brett, Ingrid S. Johnsrude, and Adrian M. Owen. The problem of functional localization in the human brain. Nature Reviews Neuroscience 3, 243-249 (March 2002).

[Chiappa and Millán 2004] Silvia Chiappa and José del R. Millán. BCI Competition III dataset V, 2004 competition. http://bbci.de/competition/iii/desc_V.html. Last accessed 19 March 2016.

[Dayan and Abbott 2001] Peter Dayan and L.F Abbott. Theoretical Neuroscience: Computational and Mathematical Modeling of Neural Systems. Massachusetts Institute of Technology, 2001.

[Delorme et al. 2010] Arnaud Delorme, Christian Kothe, Andrey Vankov, Nima Bigdely-Shamlo, Robert Oostenveld, Thorsten O. Zander, and Scott Makeig. MATLAB-Based Tools for BCI Research, D.S. Tan, A. Nijholt (eds.), Brain-Computer Interfaces, Human-Computer Interaction Series, Springer-Verlag London Limited 2010.

[Emotiv 2014] Emotiv EPOC Specifications 2014.
https://emotiv.com/product-specs/Emotiv%20EPOC%20Specifications%202014.pdf. Last accessed: 9 March 2015.

[Fisch 1999] Bruce J. Fisch. Fisch and Spehlmann's EEG Primer: Basic Principles of Digital and Analog EEG. Elsevier Ltd., 1999.

[Fok 2011] Sam Fok, Raphael Schwartz, Mark Wronkiewicz, Charles Holmes, Jessica Zhang, David Bundy, and Eric Leuthard. An EEG-based Brain Computer Interface for Rehabilitation and Restoration of Hand Control Following Stroke Using Ipsilateral Cortical Physiology. The 33rd Annual International Conference of the IEEE EMBS Boston, Massachusetts USA, August 30 - September 3, 2011.

[Garrett et al. 2003] Deon Garrett, David A. Peterson, Charles W. Anderson, and Michael H. Thaut. Comparison of Linear, Nonlinear, and Feature Selection Methods for EEG Signal Classification. IEEE Transactions on Neural Systems and Rehabilitation Engineering, Vol. 11, No. 2, June 2003.

[Hasan and Gan 2011] Bashar Awwad Shiekh Hasan and John Q. Gan. Conditional random fields as classifiers for three-class motor-imagery braincomputer interfaces. J. Neural Eng.8(2011) 025013 (5pp).

[HCRF-lib 2015] HCRF library.
https://sourceforge.net/projects/hcrf/. Last accessed 21 March 2016.

[Hodgkin and Huxley 1952] A. L. Hodgkin and A. F. Huxley. A quantitative description of membrane current and its application to conduction and excitation in nerve. J. Physiol. (I952) I I7, 500-544.

[Iturrate et al. 2009] I. Iturrate, J. Antelis, and J. Minguez. Synchronous EEG Brain-actuated Wheelchair with Automated Navigation. IEEE International Conference on Robotics and Automation, May 2009, Kobe Japan.

[Ivan 2016] Classification.
http://en.proft.me/2015/12/24/types-machine-learning-algorithms/. Last accessed 26 March 2016.

[KIN 2016] KIN450-Neurophysiology.
https://kin450-neurophysiology.wikispaces.com/Multiple+Sclerosis. Last accessed 28 May 2016.

[Koller and Friedman 2009] Dephne Koller and Nir Friedman. Probabilistic Graphical Models: Principles and Techniques. MIT Press, 2009.

[Mayer-Kress 1994] G. Mayer-Kress. Localized measures for nonstationary time-series of physiological data. Integrative Physiological and Behavioral Science, July 1994, Volume 29, Issue 3, pp 205-210.

[Sutton and McCallum 2010] Charles Sutton and Andrew McCallum. An introduction to conditional random fields. Foundations and Trends in Machine Learning Vol. 4, No. 4 (2011) 267373.

[Mirowski 2006] NeuroSimulator.
http://cs.nyu.edu/ mirowski/pub/mathneurophysiology/. Last accessed 26 March 2016.

[Morency et al. 2007] Ariadna Quattoni, Sybor Wang, Louis-Philippe Morency, Michael Collins, and Trevor Darrell. Hidden Conditional Random Fields. IEEE Conference Computer Vision and Pattern Recognition, 2007. CVPR '07. June 2007.

[Neurofeedback 2016] Anatomy and physiology.
https://neurofeedback.wikispaces.com/12+Anatomy+and+Physiology. Last accessed 26 March 2016.

[Penfield and Rasmussen 1950] Wilder Penfield and Theodore Rasmussen. The Cerebral Cortex Of Man: A Clinical Study Of Localization Of Function. Macmillan, 1950

[Pfurtscheller et al. 2004] Reinhold Scherer, Gernot R. Müller, Christa Neuper, Bernhard Graimann, and Gert Pfurtscheller. An Asynchronously Controlled EEG-Based Virtual Keyboard: Improvement of the Spelling Rate. IEEE Transactions On Biomedical Engineering, Vol. 51, No. 6, June 2004

[Pfurtscheller et al. 2003] B. Obermaier, G. R. M*ü*ller, and G. Pfurtscheller. "Virtual Keyboard" Controlled by Spontaneous EEG Activity. IEEE Transactions on Neural Systems and Rehabilitation Engineering, Vol. 11, No. 4, December 2003.

[Pfurtscheller et al. 2005] G. Pfurtscheller, Horst Bischof, Felix Lee, and A. Schl*ö*gl. Characterization Of Four-Class Motor Imagery Eeg Data For The Bci-Competition 2005. J. Neural Eng.2 (2005) L14l22.

[Pfurtscheller et al. 2006] G. Pfurtscheller, C. Brunner, and A. Schl*ö*gl, F.H. Lopes da Silva. Mu rhythm (de)synchronization and EEG single-trial classification of different motor imagery tasks. NeuroImage 31 (2006) 153 - 159.

[Presage 2015] Presage Predictive Text Platform. http://presage.sourceforge.net/. Last accessed 11 March 2015.

[Purves et al. 2004] Edited by Dale Purves, George Augustine, David Fitzpatrick, William Hall, Anthony-Samuel LaMantia, James McNamara, Mark Williams. Neuroscience, 3rd Edition. Sinauer Associates, Inc, 2004.

[PyMVPA 2012] Curve fitting. http://v04.pymvpa.org/examples/curvefitting.html. Last accessed 26 March 2016.

[Quattoni et al. 2007] Ariadna Quattoni, Sybor Wang, Louis-Philippe Morency, Michael Collins, and Trevor Darrell. Latent-Dynamic Discriminative Models for Continuous Gesture Recognition. IEEE Transactions On Pattern Analysis And Machine Intelligence, Vol. 29, No. 10, October 2007.

[Roula et al. 2012] M. A. Roula, J. Kulon, and Y. Mamatjan. Brain-Computer Interface Speller Using Hybrid P300 And Motor Imagery Signals. The Fourth IEEE RAS/EMBS International Conference on Biomedical Robotics and Biomechatronics Roma, Italy. June 24-27, 2012.

[Saa and Çetin 2011] Jaime F Delgado Saa and Müjdat Çetin. Hidden Conditional Random Fields For Classification Of Imaginary Motor Tasks From Eeg Data. 19th European Signal Processing Conference (EUSIPCO 2011).

[Saa and Çetin 2012] Jaime F Delgado Saa and Müjdat Çetin. A latent discriminative model-based approach for classification of imaginary motor tasks from EEG data. J. Neural Eng.9(2012) 026020 (9pp).

[Saa and Çetin 2013] Jaime F Delgado Saa and Müjdat Çetin. Discriminative Methods for Classification of Asynchronous Imaginary Motor Tasks From EEG Data. IEEE Transactions On Neural Systems And Rehabilitation Engineering, Vol. 21, No. 5, September 2013.

[Sassa 2014] South African Social Security Agency.
http://www.sassa.gov.za/index.php/knowledge-centre/statistical-reports. Last accessed 12 March 2015.

[Sharbrough et al. 1991] F. Sharbrough, G.E. Chatrian, R. Lesser, M. Nuwer, T. Picton. American Electroencephalographic Society guidelines for standard electrode position nomenclature. J Clin Neurophysiol. 1991 Apr, 8(2):200-2.

[Statssa 2011] Poverty Trends in South Africa.
http://beta2.statssa.gov.za/publications/Report-03-10-06/Report-03-10-06March2014.pdf.
Last accessed 12 March 2015.

[Statssa 2013] General Household Survey.
http://beta2.statssa.gov.za/publications/P0318/P03182013.pdf. Last accessed 11 March 2015.

[Taylor 2006] Jill B. Taylor. My Stroke of Insight: A Brain Scientists Personal Journey. Viking, 2006.

[Wang et al. 2004] Y. Wang, H. Yuan, D. Zhang, X. Gao, and Z. Zhang. BCI Competition, data set IVa (motor imagery, small training sets).
http://bbci.de/competition/iii/results/berlin_IVa/YijunWang_desc.pdf. Last accessed 31 March 2016.

[Wikipedia action potential 2016] Action potential.
https://en.wikipedia.org/wiki/Action_potential. Last accessed 28 May 2016.

[Wikipedia Bayesian network 2016] Bayesian network.
https://en.wikipedia.org/wiki/Bayesian_network. Last accessed 28 May 2016.

[Wikipedia electroencephalography 2016] Electroencephalography .
https://en.wikipedia.org/wiki/Electroencephalography. Last accessed 28 May 2016.

[Wikipedia electrophysiology 2016] Electrophysiology.
https://en.wikipedia.org/wiki/Electrophysiology. Last accessed 28 May 2016.

[Wikipedia Hodgkin-Huxley 2016] HodgkinHuxley model.
https://en.wikipedia.org/wiki/Hodgkin-Huxley_model. Last accessed 28 May 2016.

[Wikipedia neuron 2015] Neuron doctrine.
https://en.wikipedia.org/wiki/Neuron_doctrine. Last accessed 28 May 2016.

[Wikipedia signal processing 2016] Sampling (signal processing).
https://en.wikipedia.org/wiki/Sampling_(signal_processing). Last accessed 28 May 2016.

[Wikipedia spinal cord 2016] Spinal cord.
https://en.wikipedia.org/wiki/Spinal_cord. Last accessed 28 May 2016.

[Zalkow 2016] Spectrograms.
http://www.frank-zalkow.de/en/code-snippets/create-audio-spectrograms-with-python.html. Last accessed 26 March 2016.