THE UNIVERSITY OF QUEENSLAND

AUSTRALIA

Thesis Project Proposal

Feature Selection for Brain Computer Interface using Particle Swarm Optimisation and Genetic Algorithm

Timothy Mahoney (42360366)

Supervised by: Marcus Gallagher

March 22, 2018

Table of Contents

1 Introduction

Brain computer interfacing (BCI) is a complex problem which needs to overcome many challenges before it can be considered a viable device control mechanism. However, recent advances in computing technology, and subsequent machine learning research, has led to an increase in interest in the field. Recently, engineers at Johns Hopkins University have developed a BCI capable of controlling an upper limb prosthesis [1], even Elon Musk's latest venture is in BCI technology with his startup company *Neuralink* [2]. Due to the high spatial resolution and frequency bandwidth of electroencephalographic (EEG) signals, one hurdle slowing BCI development is determining what data is useful and what can be ignored.

1.1 Aim and Scope

This project aims to create an optimal feature subset from EEG data capable of reducing computation time and hardware requirements while still maintaining classification accuracy. The feature space will consist of the mean and variance of the five EEG frequency bands from epochs of 64-channel EEG data collected from the motor movement / imagery BCI PhysioNet Database [3] [4]. After filtering and artefact removal, the features will be selected using Particle Swarm Optimisation and Genetic Algorithm where the objective function minimises redundancy and maximises relevance of the features. The feature selector will be a filter type, i.e. independent of the classification method. Three classifiers will then be used to evaluate the feature subset: a Support Vector Machine, Random Forest and an Artificial Neural Network.

The dataset being used was created in 2004 by the BCI2000 instrumentation system and contributed to the PhysioNet database [3] [4]. The dataset contains over 1500 one- and two-minute EEG recordings from 109 participants. The system used 64 channel EEG with a sampling rate of 160 Hz. It consisted of 14 experimental runs including two baseline recordings (eyes open, then closed) and three repeats of four tasks. The datasets being considered for this project are of tasks 1 and 2, whereby the subject would *actually* close their fists then opening (motor movement), or they would *imagine* closing their fists then opening (motor imagery) for periods of time following visual cues. This results in three classes: no movement, left-handed movement and right-handed movement. The aim of the BCI for this project is to classify these events using the EEG recordings.

MATLAB will be the primary software package being used to import, sample, optimise and classify the data. This is because the author is familiar with the software package and there is a substantial pool of supporting documentation available for correctly implementing most, if not all, of the processes involved with this project.

It should be noted that there are assumptions being made in this project. The features being extracted from the EEG (ie. variance and mean of epochs in frequency bands) are typically used for BCI implementations, however previous studies have used other more advanced feature extraction methods to compute temporal frequency (de) synchronisation metrics about EEG events [5] [6]. This project will not consider these feature extraction methods but may compare results to explore their importance. Other studies have investigated the effects of user specific parameter selection under the motivation that EEG responses can be dependent on the subject [7]. Due to the nature of the dataset, implementing the BCI in this way may produce less meaningful results as the number of experimental runs per person is too small. Lastly, the focus topic for this project is the optimisation methods being implemented, hence the classification algorithms are simply used to evaluate the performance of the features selected and will not be investigated in detail.

1.2 Proposal Layout

The report clearly defines and explains the background of a Master of Engineering Science thesis project to be conducted by the author at the University of Queensland. A thorough exploration of background material is covered in Section 2, followed by an investigation into previous similar studies in Section 3. A risk analysis is displayed in Section 4 and Section 5 covers the project plan, including milestones, assessment dates and a timeline.

2 Background

This project requires a reasonably large amount of background coverage due to the multidisciplinary nature of BCI technology. As such, this section will cover an introduction to EEG sampling, features and filtering, followed by defining a BCI and common features. Next the types and potential benefits of feature selection are outlined, and the two proposed methods of optimisation to be used in this project are explained. This is followed by a brief introduction to the classifiers being used to assess the feature selections' performance against previous studies.

2.1 Electroencephalography

Electroencephalography (EEG) is the recording of action potentials originating from the brain. It is most often used in the study of brain functionality or the diagnoses of neurological disorders [8]. EEG signals are sampled using several small electrodes placed on the scalp, which can detect minute potential differences un-invasively on the surface of the skin [8]. Through proper filtering and signal processing techniques, the summation of electrical activity from the brain cells can be isolated and displayed allowing professionals to examine neural activity [8].

The exact placement of the electrodes is a generally subject to the standardised International 10-20 System, whereby the distance between left-right and front-back electrodes is either 10% or 20% of the total skull distance [9]. The electrodes are labelled according to the area of the brain, or lobe, that they're reading from. These are pre-frontal (Fp), frontal (F), temporal (T), parietal (P), occipital (O) and central (C) as seen in Figure 1 [8] [9]. The number of electrodes in total varies, systems can sample anywhere from 1 to 256 EEG signals in parallel [8].

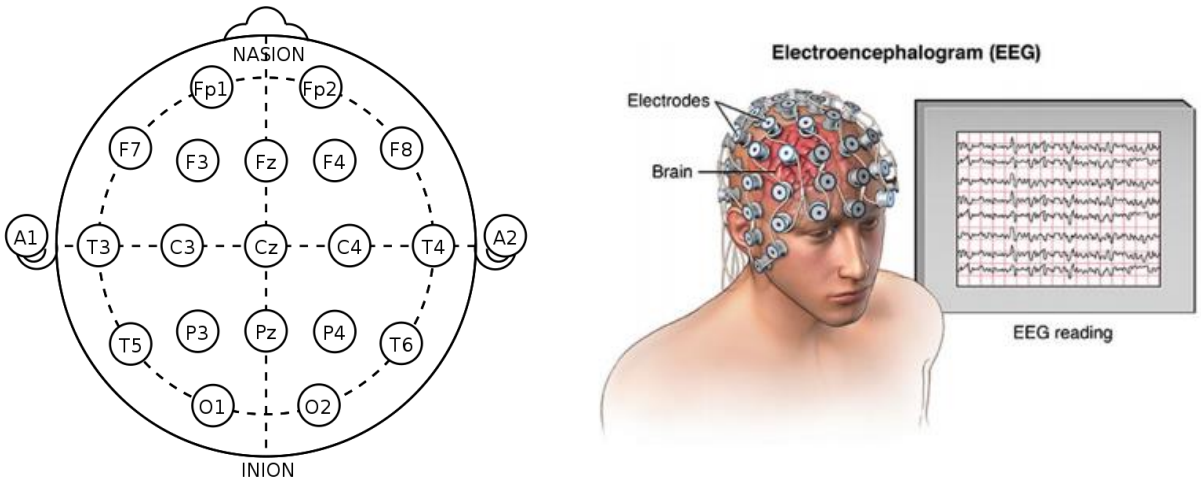

Figure 1 (left) Standard EEG labelling of each lobe. (right) Examples of EEG positioning and recordings [8]

2.1.1 Frequency Bands

The frequency components of EEG are considered some of the most important criteria for interpreting brain function [8]. The bandwidth of these components is clearly defined: delta (0.5 – 4Hz), theta (4 -8Hz), alpha (8 – 13Hz), beta (13 – 30Hz) and gamma (>30Hz). The presence or absence of each component is generally associated with a particular type of brain function or state [8]. For instance, alpha waves are typically observed during sensory-motor activity. There is evidence to suggest that in addition to the amplitude of each bandwidth indicating a neurological event, the (de) synchronisation between particular bands may correlate to the mind processing information [5] [6].

2.1.2 Artefacts

Signals included in EEG recordings that do not originate from neural potentials are referred to as artefacts. They can be introduced from other physiological events such electromyographic signals (EMG) from muscles, electrooculographic signals (EOG) from eye movements or even electrocardiographic signals (ECG) from the heart. They can also exist as noise from poor electrode contact or 50Hz mains coupling. EEG signals can easily be corrupted by artefacts as their amplitude is in the range of 10-100μV, hence filtering techniques have been thoroughly explored in previous studies to establish clean signals for professional examination or further processing [8].

2.2 Brain Computer Interface

A brain-computer interface (BCI) aims to establish a pathway between neuropsychological signals and an external device [10]. One potential goal of such an interface is to provide a substitute method of communicating or controlling extremities for people with neuromuscular impairments [8] [10]. The ability to classify mental imagination events would allow patients to regain some degree of freedom whether it be through wheelchair or neuroprosthesis control, or communication. With proper training a person could repurpose their neural pathways for any number of applications to improve quality of life. However, for BCI systems to be considered a useful alternative control method they need to achieve critical performance requirements for robustness, consistency and safety [10].

Mason and Birch described a general framework for BCI systems as a way of standardising ideas and terms for these devices to aid with their development and comparison [11]. The framework consists of 6 steps [11] [8]:

Brain activity measurement - the sampling of the raw EEG data

Pre-processing – digital to analogue conversion, initial filtering and artefact removal

Feature extraction – retrieve meaningful features of the EEG data which can discriminate between mental states being used to control the system

Classification – employ some algorithm which can classify mental states based on the extracted features.

Translation into a command – relate the result of classification with a response in the target device.

Feedback – sensory feedback to the user not only allows them to know that the device is working, but also helps train their brain to use it.

The proposed project will focus on the area of feature extraction but will include elements of pre-processing and classification methods.

Commonly used mental responses for controlling a BCI are motor imagery (MI), where the subject imagines moving without performing the movement, and motor movement (MM), which is the brains response to an actual motor movement being made. Both tasks will be considered as potential inputs into a BCI system.

2.2.1 Feature Extraction

One of the most critical elements of a functional BCI is effective feature extraction. Each lobe within the brain is responsible for different mental functions, and as discussed in Section 2.1.1, each frequency band is related to particular mental states. For this reason, extracted features are usually some metric of information of particular frequency bands recorded from one or more mental lobes. Often a subset of channels and/or frequency bands are chosen for analysis based on prior knowledge or assumption regarding their physiological function. However, a more optimal selection of features may be extracted using feature selection methods to maximise the classification performance of a BCI system.

2.3 Feature Selection

Consider the case where a 128 channel EEG was divided into its 5 frequency bands and the mean and variance of each band was taken as the feature set, resulting in a vector 1280 features in length. A feature vector this large introduces several problems in the classification stage, such as [12]:

1. Over complicating the classification model
2. Long training time
3. Slow real time BCI performance
4. Increased risk of overfitting
5. Curse of dimensionality – small ratio of data to number of features

Feature selection attempts to counter these problems by selecting a subset of features without incurring much loss of information [12]. The selection process follows the premise that certain features may correlate with each other (termed: *redundancy*), whereas others may have little to do with classification

so including them within the feature vector results in no information gain (*irrelevance*). Feature selection algorithms are essentially a search problem for a feature set which is optimal concerning some evaluation score. Score metrics depend on the specific algorithm being applied, the three types are: filter, wrapper and embedded [13]. Filter algorithms select the feature subset independent of the prediction method, whereas wrappers assess the subset using the predictor and embedded methods optimise the subset in the process of training the predictor [13]. This project will use a filter approach and implement two search optimisation algorithms to determine a suitable feature subset.

2.4 Computational Optimisation

Computational optimisation uses algorithms to minimise or maximise an objective function. In the case of EEG feature selection, the objective function would be the metrics we chose to optimise for (e.g. minimise redundancy, maximise relevance) over the space of feature subsets. If the space of subsets includes every combination of the feature vector proposed earlier (1280 features), there exists over 10^{30} combinations of features. Brute force search methods over this space are not computationally feasible and as such optimisation problems involving large spaces over many dimensions have led to the creation of novel search algorithms to find *good enough* optimal solutions, these are referred to as *heuristic* algorithms. These methods aim to computationally balance *exploration* and *exploitation* of the search space. Exploration is the wide spread search of the space, often at random, in order to avoid focussing on what could only be local optima, whereas exploitation attempts to converge on the ideal solution based on information the algorithm has already discovered. This project will compare the effectiveness of two such heuristic methods: Genetic Algorithm (GA) and Particle Swarm Optimisation (PSO). Previous studies implementing these, or other optimisation methods, in relation to EEG classification will be discussed in Section 3.

2.4.1 Genetic Algorithm

As is the case with many optimisation methods, Genetic Optimisation is an algorithm designed to mimic the principals of a naturally occurring system. The method takes the process of evolution, such that the genetic code of the parents is inherited by the child, and simulates a ‘survival of the fittest’ scenario over a series of generations [14]. With respect to this project, the genetic code of the parents/child would be a combination of EEG features and the measure of ‘fitness’ could be the metrics of redundancy and relevance for that particular combination. The method starts with randomly generated parents in the search space. The next generation is computed with a combination of three *genetic operators* [14]:

> **Reproduction operator:** poor solutions are neglected, whereas fit solutions are passed down. The number of candidates is preserved by including more children with strong performance to replace the those lost. The resulting candidates are now part of what is

referred to as the *mating pool*. It can be seen that the reproduction operator does not result in any new solutions, they are formed in the remaining two operators.

Crossover Operator: there are a few variations of this operator, but generally two candidates (the parents) from the mating pool are chosen randomly and a subsection of their genetic structure is switched to form children. This operation supplies new candidates with genes that have been previously shown to perform well in the reproduction stage, in the hope that they may perform better than their parents.

Mutation Operator: the remaining codes are now subject to mutation, where there exists a small probability that one or more of their genes may be altered. This introduces exploration into the algorithm, allowing it to traverse previously unseen areas of the search space in pursuit of better optima.

New generations are computed continuously until some stopping criteria is met, for example a certain number of generations or a high score threshold. The process is described quite well by Deb, K. who uses the example of minimising materials used to produce a cylinder of set volume as a function of diameter and height [14].

2.4.2 Particle Swarm Optimisation

Initially, the particle swarm model was introduced in 1995 by Kennedy and Eberhart [15] to describe the behaviour of swarm phenomena in nature, such as schools of fish, ant colonies or birds. It has since been adapted and successfully implemented as a search optimisation algorithm for many applications [15]. The model is simple, yet powerful. The premise is that each particle in the swarm alone is not an effective optimiser, but through communication with other particles the swarm as a whole can find a valid heuristic optimum. The algorithm is as follows [15]:

1. Randomly distribute n particles throughout the search space
2. Evaluate the objective function for each particle
3. The particle with the current found global optimum tells the other particles and each particle moves towards it. The velocity of that movement for each particle is a function of three terms: the particles current velocity, the particles personal best optimum and the global optimum, including scalars of uniform random distribution. Respectively, these terms are referred to as the particles' *inertia*, *cognitive component* and its *social component*.
4. Repeat steps 2 and 3 until some stopping criteria is met. The solution is the best optimum found by the particles.

As the particles traverse the search space, their constant communication and slight randomness in velocity allows them to determine if a better optimum solution has been found elsewhere, thereby

avoiding local optimum solutions. PSO implementation in MATLAB has been well documented. This project will conduct a comparison between PSO and genetic algorithm in optimising feature selection of EEG, the resulting feature sets will then be used with various classifiers to examine their effectiveness.

2.5 Classification Algorithms

This project's main focus is with the optimisation techniques being applied to the feature set extracted from the EEG data and will hence not go into too much detail concerning the development and tuning of classifiers. However, as the proposed feature selection filter is independent of the method of classification it will be used with, it would be interesting to see how the final feature combination performs when input into a variety of supervised classification algorithms. The classifiers will be a Support Vector Machine (SVM), a Random Forest (RF) and an Artificial Neural Network (ANN). These methods were chosen as they have been used previously with EEG classification [16] [17] [18] and MATLAB documentation on their implementation is extensive.

2.5.1 Support Vector Machine

Support Vector Machines are simple discriminant classifiers which construct an optimal hyperplane in the sample space between classes from a training set. The resulting hyperplane is then the discriminant used to classify new data points. Usually, the best hyperplane is taken as the one that maximises the separation of the classes.

2.5.2 Random Forest

A decision tree is a chain of nodes branching to each other, where each node decides which subsequent node it branches to will be selected based on some parameter. The aim is that upon reaching an end node, the tree will have enough information about the data's parameters, so it can be classified. The trees are supervised learners that try to maximise information gain per decision node allowing the tree to separate the classes. A Random Forest is a collection of decision trees that vary randomly in terms of their structure and node parameters. Any new data given to the forest is classified based on the votes it receives from each tree resulting in a higher probability of correct classification.

2.5.3 Artificial Neural Network

An artificial neural network mimics the principals of neural pathways in the brain. A single neuron takes weighted inputs from one or more neurons in a previous layer, where it can then signal other neurons in subsequent layers. The weights between the neurons are updated as learning progresses, tuning the

network to accept data features as inputs and predict its class based off supervised learning from a training set. A wide variety of network topologies exist, however this project will only be considering simple feedforward networks with a small number of layers.

3 Previous Studies

Feature selection for EEG analysis has been the subject of many previous studies as BCI control still presents a real challenge in biomedical engineering today. Exploring these papers gives insight into the importance of accurate feature selection as well as confirming that this project has a solid foundation consisting of other similar project which have produced meaningful results. In addition, the study discussed in 3.4 provides benchmark values to compare with the results of this project as it makes use of the same dataset.

3.1 Automated Feature Selection for Brain Computer Interfaces 2003

This paper by Shroder et al. [7] was part of the first IEEE conference preceding for neural engineering in 2003. The motivation for the study was that appropriate features can be dependent on the subject and the mode of control which the subject is trying to employ, hence a robust algorithm for feature selection would be useful in tuning a BCI for a particular subject and purpose. Shroder et al. use a wrapper selection method, including a Genetic Algorithm selector and an SVM classifier. The results were positive, in most cases the GA was able to determine the optimal solution which was previously computed by brute force method. The improved feature selection was then able to create a better SVM classifier than simply using all EEG channels. While there are similarities with this project, the two are quite different. In this case, data was collected from 5 subjects moving a curser on a screen, which differs from this project's dataset of fist clenching for 109 subjects. In addition, the feature selection method used by Shroder et al. was a wrapper, whereas this project implements a filter.

3.2 Feature Selection for Brain Computer Interfaces 2009

Koprinska, from the University of Sydney 2009, evaluated the performance of five filter feature selection methods for BCI EEG data [19]. The methods were information gain ranking, correlation-based feature selection, ReliefF, consistency-based feature selection and 1R ranking, none of which are being considered in this project. She also uses ten different classifiers to evaluate the results, which included RF and SVM. The results varied across the filters and classifiers, however relative to other top competitors with the same dataset, she performed quite well. Unfortunately, the data used is not the Physionet MM/MI dataset so the results of this project can't be directly compared, however Koprinska, among other researchers, does provide a benchmark in EEG classification accuracy to aim for.

3.3 Particle Swarm Optimisation for a Brain Computer Interface

Satti et al. propose the use of class correlation analysis as a score metric when optimising parameter selection of EEG channels for use with a BCI with PSO [10]. Their motivation was that using correlation analysis would be much quicker than classification accuracy to assess the combination sets presented by the particles. Essentially Satti et al. aim to use a filter parameter selector as opposed to a wrapper and produce comparable classification accuracy. The results outline that feature selection is able to eliminate redundant data and allows features characterizing a classifiable event to be more prominent and detectable. Despite this, the wrapper method still appeared to be more accurate by 2%. Satti et al. argues that this is expected as the wrapper method is optimising for classification accuracy, the very metric he used to assess the two methods, whereas the filter method is only optimising for class correlation. However, they conclude that a mere 2% accuracy loss is worth the large decrease in computational time achieved with the PSO method. Interestingly, this paper uses the same dataset as Koprinska.

3.4 EEG Classification with Advanced Feature Extraction and Machine Learning

This paper by Alomari et al. uses the same dataset proposed for use with this project [20] [3] [4]. They propose a system which focuses on feature extraction, rather than selection. The features include Event Related Synchronization and Desynchronization (ERS/ERD) of the beta and alpha bands, as well as Movement Related Cortical Potentials (MRCP), which were then parsed through Independence Component Analysis to eliminate artefacts. The remaining feature set is then evaluated using an ANN and SVM. The results show high classification accuracy: 97.1% using SVM and 86.5 using ANN. This paper sets a benchmark for this project to aim for and may provide insight into the importance of feature extraction vs. feature selection. On one hand, utilizing knowledge regarding the problem may provide insight into what particular features to extract from the data and what to ignore. On the other, the very nature of EEG data is not entirely understood, each channel is the summation of potentials across many thousands of neurons which could have any number of functions. This could indicate that choosing features based on optimisation may produce better classification performance, at the cost of intuitive understanding.

4 Risk analysis

This risk analysis will briefly outline the three risks involved with this project, the level of risk and actions that will be taken to minimize this risk. The risk level will be assessed against the following risk matrix:

		Severity			
Likelihood		Negligible	Marginal	Critical	Catastrophic
	Rare	Low	Low	Moderate	High
	Unlikely	Low	Low	Moderate	Extreme
	Possible	Low	Moderate	High	Extreme
	Likely	Moderate	High	High	Extreme
	Certain	High	High	Extreme	Extreme

4.1 Breach of Copyright

Copyright laws could be breached by including datasets and results from other studies in this project.

Likelihood: Rare
Severity: Marginal
Risk level: Low

Mitigations: All ideas and data used within this report which are not owned by the author will be referenced correctly in accordance with the University of Queensland referencing and plagiarism policy as well as per request by the original owner/s. In addition, datasets will only be collected from reputable sources.

4.2 Data Loss

Unforeseen circumstances such as damage to or loss of the authors computer, data or other records could lead to loss of project progress. As the project progresses, the severity of this risk increases to a critical level.

Likelihood: Unlikely
Severity: Critical
Risk level: Moderate

Mitigations: All data and electronic records will be committed regularly to a git repository dedicated to this thesis project. This includes but is not limited to: any MATLAB script, function files or data sets, electronic notes and assessment items. Care will be taken to minimise risk of malware being downloaded. Computers being used to work on the project will not be left unattended while thesis material is able to be accessed.

4.3 Underestimation of Time Requirement

This project includes many steps which require careful implementation and validation before advancing to the next stage, as discussed in Section 5. Incorrect implementation or underestimation of the time required to complete anyone of these stages could lead to limitations in the project due to time constraints.

Likelihood: Possible
Severity: Critical
Risk level: High

Mitigations: A well defined project plan has been outlined in the following section which includes milestones designed to motivate completion of the thesis within a manageable timeframe. This plan has been designed to expect delays and provide adequate time to overcome obstacles. In addition, regular correspondence with the thesis supervisor Prof. Gallegher or one of his tutors/associates has been agreed upon to verify progress and provide guidance if required.

5 Project Plan

This project plan organizes the required workload into five milestones with achievable timeframes. Each milestone corresponds with a significant subsection in a series of tasks which are required to be completed and thoroughly validated in order to ensure the correct implementation of each subsequent task. The tasks parallel 3 segments of the BCI framework presented in Section 2.2: preprocessing (Milestone 1), feature extraction (Milestone 2, 3 and 4) and classification (Milestone 5). The milestones aim to complete sections of the project in correlation with assessment pieces, such that there will be enough completed to present by the seminar date and there will be enough time to write the conference paper, thesis paper and prepare for the demo.

5.1 Milestone 1: Preparation of the Dataset

Commence: 26th of March
Complete: 23rd of April
Duration: 4 Weeks

This first step is vitally important to the project as any inconsistencies introduced here will carry through the pathway, limiting the power of the results. Milestone 1 involves downloading and importing the dataset into MATLAB, filtering the data and removing artefacts. Research will be conducted into appropriate artefact removal techniques, as well as suitable filtering bands and methods. Epoch length will be considered. In addition, careful consideration will also be required to pair movement annotations with the correct epochs. Methods of storing the data and the exact structure will also be a concern considering the size of the dataset. The goal of this milestone is to achieve consistently segmented and correctly annotated epochs of EEG data over every channel for each subject.

5.2 Milestone 2: Extract Features

Commence: 23rd of April
Complete: 14th of May
Duration: 3 Weeks

Milestone two consists of taking the clean EEG data and segmenting it into frequency bands with minimal loss of information. This can be achieved two ways: fast Fourier transform (FFT) and discrete wavelet transform (DWT). An analysis into the advantages of each considering this application will need to be conducted. This milestone will then be implemented in MATLAB to attain the mean and variance of each band for each epoch. These will likely be the features used in this project. As this

milestone will not take very long to achieve, it may include brief research into other potential features attainable from the data, e.g. skew and kurtosis.

5.3 Milestone 3: Feature Selection Using PSO

Commence: 14th of May
Complete: 13th of August
Duration: 13 Weeks

Milestones 3 and 4 may run one after the other or in parallel. Milestone 3 will involve a revisit of the PSO algorithm, any variations which may improve feature selection in this project and an examination of any previous studies using PSO for a similar purpose. Parameters for consideration would include the number of particles to place randomly in the space as well as stopping criteria. Specifics on how to implement PSO in MATLAB would also need to be investigated, such as previous examples and possible toolboxes. Research into redundancy and relevance calculations will also need to be done to ensure their proper implementation. The milestone is completed when the PSO algorithm effectively selects a subset of features to be used for classification. Milestones 3 and 4 combined are likely to constitute the majority of the project as optimisation is the focus topic, and hence have been allotted an appropriate expected duration.

5.4 Milestone 4: Feature Selection Using GA

Commence: 14th of May
Complete: 13th of August
Duration: 13 Weeks

Much the same as Milestone 3, Milstone 4 aims to use MATLAB to implement the Genetic Algorithm for feature selection optimisation of the EEG data. To ensure the algorithm is applied properly, it will need to be revisited in greater detail. A major consideration with this algorithm is how to present the data such that it can be likened to binary code. As with PSO, the number of candidates and stopping criteria will also be explored. Again, this milestone is completed when the algorithm successfully produces a subset of features optimised in terms of relevance and redundancy, for use with the proposed classifiers.

5.5 Milestone 5: Classification

Commence: 13th of August

Complete: 17th of September

Duration: 5 Weeks

Milestone 5 consists of coding the classification algorithms (SVM, RF and ANN) in MATLAB and using them to classify the EEG classes from the features selected in Milestones 3 and 4. As machine learning is not the focus of this project, limited time will be allotted for designing the ideal or optimal classification algorithm. However, there are a wide range of parameters which will need to be considered in order to initialise these algorithms, these include but are not limited to: distance metric and kernel function for SVM, the number of trees, tree structure and stopping criteria for RF, the network size/depth, activation function and stopping criteria for ANN, and the split between training/validation/testing datasets. These classifiers will be applied to both the subset of selected features as well as the total set of features to compare accuracy and computation time.

6 Conclusion

This proposal has clearly outlined the motivations, scope and aims of this project. Considerate background research was conducted to establish a solid understanding of the concepts involved across the disciplines of biology, engineering and computer science. Investigations into previous studies indicated that this particular focus on a filter type feature selector using PSO and GA for EEG MI/MM classification has not been conducted before to the authors knowledge. However, outcomes from similar studies provide confidence in this projects achievability as well as supplying a benchmark to compare results with. Potential risks to the project have been considered and steps put in place to mitigate them, and a detailed project plan has been implemented to ensure satisfactory progress. To conclude, this project is designed to explore useful techniques within the field of BCI development with the hope that further technological advances could lead to a viable alternative device control mechanism.

7 Bibliography

[1] "Brain-Computer Interfaces," JHU Neuroengineering & Biomedical Instrumentation Lab, 2008. [Online]. Available: http://neuroengineering.bme.jhu.edu/ecog.htm. [Accessed 22 March 2018].

[2] "Neuralink," Neuralink Corp, 2018. [Online]. Available: https://www.neuralink.com/. [Accessed 22 March 2018].

[3] A. L. G. L. H. J. I. P. M. R. M. J. M. G. P. C.-K. S. H. Goldberger AL, "EEG Motor Movement / Imagery Dataset, PhysioBank, PhysioToolkit, and PhysioNet: Components of a New Research Resource for Complex Physiologic Signals. Circulation 101(23):e215-e220," 13 June 2000. [Online]. Available: https://www.physionet.org/pn4/eegmmidb/. [Accessed March 2018].

[4] G. M. D. H. T. B. N. W. J. Schalk, "BCI2000: a general-purpose brain-computer interface (BCI) system.," *IEEE Trans Biomed Eng,* vol. 51, no. 6, pp. 1034-1043, 2004.

[5] Y. Kim, J. Ryu, K. K. Kim, C. C. Took, D. P. Mandic and C. Park, "Motor Imagery Classification Using Mu and Beta Rhythms of EEG with Strong Uncorrelating Transform Based Complex Common Spatial Patterns," *Computational Intelligence and Neuroscience,* vol. Online, 2016.

[6] D. McFarland, L. Minor, T. Vaughan and J. Wolpaw, "Mu and Beta Rhythm Topographies During Motor Imagery and Actual Movements," *Brain Topography,* vol. 12, no. 3, pp. 177-186, 2000.

[7] M. Schroder, M. Bogdan, T. Hinterberger and N. Birbaumer, "Automated EEG feature selection for brain computer interfaces," in *Conference Proceedings. 1st International IEEE EMBS Conference on Neural Engineering 2003,* Capri Island, Italy, 2003.

[8] S. Suily, Y. Li and Y. Zhang, EEG Signal Analysis and Classification: Techniques and Applications, Gewerbestrasse: Springer International Publishing, 2016.

[9] V. Towle, J. Bolaños, D. Suarez, K. Tan, R. Grzeszczuk, D. Levin, R. Cakmur, S. Frank and J.-P. Spire, "The spatial location of EEG electrodes: locating the best-fitting sphere relative to cortical anatomy," *Electroencephalography and Clinical Neurophysiology,* vol. 86, no. 1, pp. 1-6, 1993.

[10] A. R. Satti, D. Coyle and G. Prasad, "Spatio-spectral & temporal parameter searching using class correlation analysis and particle swarm optimization for a brain computer interface," in *IEEE International Conference on Systems, Man and Cybernetics*, 2009.

[11] S. Mason and G. Birch, "A general framework for brain-computer interface design," *IEEE Transactions on Neural Systems and Rehabilitation Engineering,* vol. 11, no. 1, pp. 70-85, 2003.

[12] Y. Saeys, I. Inza and P. Larrañaga, "A review of feature selection techniques in bioinformatics," *Bioinformatics,* vol. 23, no. 19, pp. 2507-2517, 2007.

[13] I. Guyon and A. Elisseeff, " An Introduction to Variable and Feature Selection," *Journal of Machine Learning Research,* vol. 3, pp. 1157-1182, 2003.

[14] K. Deb, "An introduction to genetic algorithms," *Sadhana,* vol. 24, no. 4-5, pp. 293-315, 1999.

[15] R. Poli, J. Kennedy and T. Blackwell, "Particle Swarm Optimization," *Swarm Intelligence,* vol. 1, no. 1, pp. 33-57, 2007.

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

[17] M. Bentlemsan, E. Zemouri, D. Bouchaffra, B. Yahya-Zoubir and K. Ferroudji, "Random Forest and Filter Bank Common Spatial Patterns for EEG-Based Motor Imagery Classification," in *Intelligent Systems, Modelling and Simulation*, Langkawi, Malaysia, 2014.

[18] A. Subasi and E. Ercelebi, "Classification of EEG signals using neural network and logistic regression," *Computer Methods and Programs in Biomedicine,* vol. 78, no. 2, pp. 87-99, 2005.

[19] I. Koprinska, "Feature Selection for Brain-Computer Interfaces," in *Proceedings of the Workshop on Advances and Issues in Biomedical Data Mining, Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD 2009)*, Bangkok, 2009.

[20] M. Alomari, A. Samaha and K. AlKamha, "Automated Classification of L/R Hand Movement," *International Journal of Advanced Computer Science and Applications,* vol. 4, no. 6, pp. 207-212, 2013.