UNIVERSITY OF SYDNEY

HONOURS THESIS

Machine Learning for an EEG-based Brain Control Interface

Author:
Saahil RELAN

Supervisor:
Dr. Graham BROOKER

A thesis submitted in fulfilment of the requirements for the degree of Bachelor of Engineering (Mechatronics)(Space)(Honours)

for

School of Aerospace, Mechanical and Mechatronic Engineering

November 2, 2017

Statement of Student Contribution

I, Saahil RELAN, declare that this thesis titled, "Machine Learning for an EEG-based Brain Control Interface" and the work presented in it are my own. I confirm that:

- I worked with Dr. Daniel Hermens and Dr. Graham Brooker to define the objectives of this thesis.
- I developed a research design and experiment to answer the questions raised in this paper.
- I developed the programs required to train the machine learning algorithms.
- I developed a Java based BCI that interacts with the Muse headset.
- I developed the numerous programs, filters and tools required to analyse all the EEG datasets and the outcomes of cross-validations.
- I carried out relevant Literature Research in the field of Machine Learning.
- I carried out numerous experiments to map subjects' brain activities.
- I carried out longitudinal tests to extend my analysis for possible future applications.
- I composed the following honours thesis.

Signed:

Saahil Relan

Signed:

Dr. Graham Brooker

Date:

Disclaimer

Under the University of Sydney Guidelines and pursuant to Section B Clause 2190.0 of the Human Research Ethics Committee (HREC) Online Application Questions, my investigation holds 'negligible risk'. Therefore the methodology and subject testing performed in this dissertation does not require HREC approval.

Abstract

Utilising machine learning algorithms to interpret EEG-based signals has revolutionised the design of assistive devices. It has improved the lives of people with disabilities, and redefined modern society's research into brain computer interfaces (BCIs). Therefore extensive funding and academia has been directed at developing practical EEG-based systems that can perform in real time and enable users to effectively manipulate digital devices with only their thoughts.

This study takes an original approach towards investigating the capabilities of a modern BCI by using 39 subjects and five colour based tests to examine how accurately and reliably it can classify the brain activity of subjects envisaging yellow. The findings reveal that with a consumer grade EEG headset, and SVM-based algorithm, a real time BCI can achieve classification accuracies above 75% at a 95% confidence interval, and produce an accuracy range of 85-95% if parameters are optimised. Longitudinal studies on a male and female subject are also performed showcasing similar results and delineating modern algorithms capacity to deal with the presence of uncontrolled variables.

Outcomes of this investigation lay the foundations for further research that attempts to determine the potential of machine learning algorithms and explore more effective ways to implement them. These can include using a wide array of thoughts to control complex devices or find the extent to which BCIs can perform live sessions accurately and reliably. Such systems will have the potential to extend the functionality of the human mind giving us, as a society, access to a new mechanism through which we can interact effortlessly with the digital environment around us.

Acknowledgements

I would like to firstly acknowledge and show real appreciation for Dr. Graham Brooker's time, expertise and passion that he has kindly shared with me throughout the year. In his truly inspiring words: 'It was just my job.', he reminds me that humility is underrated, under appreciated and always in short supply. A special mention also to Dr. Daniel Hermens, who believed in my idea from the very first day I stepped in his office, and always had faith in me. Without his initial support it would not have been possible to start, let alone complete this thesis.

To my whole family. I thank them for all their continued support throughout the year and otherwise, always, in everything I strive to do. They always do what they feel is best for me and play a key role in shaping who I am today. A special thanks goes out to Bianca given everything that she does for me. She always gives her time and never thinks twice to give me a helping hand in times of need.

...On a less serious note to my friends at university who have been with me at during my highs and lows. I've learnt so much spending my time with all of you because in some way or another you have contributed to this thesis. Whether it was a concept you taught me long time ago, a trick you showed me at some point or a new way to live life staring down the end of an empty tequila bottle, this thesis has a piece of all of you embedded into it.

Mitchell Rain Colak. You've been a friend and mentor from the day we sat in Engineering Mechanics in 2013. We've been through thick and thin at University and I couldn't have asked for a better man to accompany me on the journey. Thank you for always being there, imparting your ever-lasting wisdom and sharing this unforgettable experience!

Lastly I express sincere gratitude to all my other friends, particularly those who showed real interest and have really helped make this thesis concept become a reality.

Contents

List of Figures

List of Tables

List of Abbreviations

ANN	Artificial Neural Network
API	Application programming interface
BCI	Brain Computer Interface
ECoG	Electrocorticography
EEG	Electroencephalogram
ERP	Event-Related Potential
fMRI	Functional Magnetic Resonance Imaging
LDA	Linear Discriminant Analysis
MEA	Micro-Electrode Arrays
NIRS	Near Infra-Red Spectroscopy
NMP	Neuro-motor prosthesis
PCA	Principal Component Analysis
SVM	Support Vector Machine
WEKA	Waikato Environment for Knowledge Analysis

1 Introduction

1.1 EEG

Brain computer interfaces (BCI) were initially developed to aid physically challenged individuals with locked-in syndrome to overcome their difficulties through providing a means of communicating, with a "mind reading" machine (Alonso-Valerdi, Salido-Ruiz, and Ramirez-Mendoza, 2015). Today, in the rapid-growing technological climate, research into BCIs has cultivated applications for mind reading using different mental states that were previously inconceivable (Lotte et al., 2015). A BCI is a communication system between the brain and a machine where electrical brain signals are transmitted from within the scalp to trigger a command, or control a digital system. A typical BCI consists of a measurement device to capture cognitive activity, a filtering/decoding system and an activation module to generate commands (Myrden and Chau, 2017). Research into BCI's in the last two decades has demonstrated a clear trend towards medical diagnosis and improving the quality of life of physically challenged individuals (MAK). However, more recent investigations established their focus on understanding how to manipulate the level of control humans can establish with machines (Kübler et al., 2006).

There are various brain imaging modalities used for BCI, with electroencephalography (EEG) being extensively adopted by academics and corporations, warranted by its non-evasive techniques, portability, low-cost and combined demonstration of high temporal resolution (Ge et al., 2017). EEG is the process of detecting the electrical potential of neurons, named raw EEG signals, through the placement of electrodes on the scalp, with a filtering process adopted to attenuate unwanted noise. Non-cerebral signals (called artefacts) caused by voluntary and involuntary movements of muscles have previously posed a challenge in attempting to extract meaning from brain signals as they tend to generate unwanted noise and by contaminating recorded data. (Kaur and Kaur, 2015). However, the methodologies and equipment used in this report, demonstrate society's progression into the field of machine learning and how simple features can be utilised with machine learning algorithms to extract meaningful information and enhance our interaction with computer interfaces.

EEG's usually consist of many wet electrodes (32 and above), large filtering systems and have complex preparation procedures associated with their use. This has prevented the progression of research into them and established hurdles into their viability as a consumer grade product. However the invention of EEGs with fewer dry electrodes and small electronic digital filters has pushed the boundaries of their capabilities as a consumer product. Consequently, I have been able to conduct my research on an EEG device (Muse 2016) that provides the flexibilty to be used with significant ease and wirelessly with a smartphone device.

1.2 Machine Learning

The demand for developing machine-learning algorithms to analyse information in the 21st century has particularly risen due to the beginning of the data era, with the invent of cloud computing, despite academics utilising artificial intelligence systems since the 1950s for scientific curiosity (Copeland, 2016). With technology continuing to obey Moore's Law, today's microprocessors have accelerated the commercial use of such algorithms, as they can execute millions of instructions per second, overcoming the challenge of performing machine-learning in *real time*; a reality that was not possible 25 years ago. In 2011, a McKinsey report (James et al., 2011) announced that there was a shortage of 190,000 people in the United States alone who possessed deep analytical skills, and data storage was exponentially rising with already an estimated 13 exabytes stored around the world. Therefore the adoption of machine learning systems, that could process and analyse data, had become a priority and the pinnacle of research and growth in data analytics around the world.

Machine learning algorithms are mathematically derived algorithms that are implemented by machines to identify patterns and rules from complex data sources that are subsequently used to to make predictions and decisions in uncertain environments (Murphy, 2012). Machine learning can be broken down into two classes of learning: supervised and unsupervised learning. Supervised learning is mapping known inputs of a variable x to a known output variable y for the machine to establish its own pattern/trend. Unsupervised learning requires the machine to generate its own outputs based on the model data provided. In both cases, algorithms employ statistical methodologies to formulate predictions without the confirmation of complementary statistical proofs. Supervised learning takes the form of classification, where outputs take the form of discrete values, or regression, where output values are continuous. Unsupervised learning, on the other hand, involves grouping input data based on their similarities (L'Heureux et al., 2017).

1.3 EEG, Machine Learning and Thesis

In this paper, supervised machine learning marries EEG with algorithms trained using data retrieved from headsets to predict whether a person is thinking about switching on a light bulb or not. Here, switching on a light bulb entails a subject thinking about the colour yellow and evoking some level of concentration. As the colour yellow can be perceived in multiple varieties, in this study, 39 subjects participate in five tests that require them to think about the colour yellow[1] in the presence of visual and aural stimuli, whilst an EEG headset records their brain activity. As brain states can be extremely varied at different times of the day, brain patterns can be very unpredictable. Thus longitudinal tests are performed upon 2 subjects, who attempt these five tests, at seven random times, in order to evaluate the impact of different brain states in an uncontrolled and natural environment. These include a multitude of emotional states which form part of our everyday lives when attempting simple tasks, and thus must be studied.

Consequently, this study is exceptionally results driven with the methodology atypical in nature, conducted in an uncontrolled environment, to test the capabilities of a modern day EEG headset and machine learning algorithm in the presence of artefacts. It investigates the following questions with a Java-based BCI system adopting a low-cost Muse (2016) headset and open source WEKA[2] machine learning algorithms:

- How reliable and accurate are machine learning algorithms today?
- How reliable are consumer grade EEG headsets?
- Can we use cheap consumer grade headsets with modern machine learning algorithms to develop an effective BCI?

This paper initially outlines the research that has been conducted in the EEG and BCI space through a literature review, specifically alluding to developments and findings of recent studies whilst identifying limitations that may have prevailed in authors' investigations. It then moves on to describing the methodologies used to design the experiment, in Chapter 3 including high and low level designs for each component of the BCI, alongside the statistical approach used to analyse data acquired. The statistical information is presented in Chapter 4 to identify and discuss the presence of trends in subjects' cross validation results and indicate which machine learning parameters formulate the most effective outcomes. Finally, the results are discussed within the framework of the literature examined in Chapter 2, limitations of the experiment are summarised and suggestions for future work are illustrated in Chapters 5 and 6.

[1]One of the tests asks subjects to think about the colour Blue as a contrast to yellow
[2]Waikato Environment for Knowledge Analysis

2 Literature Review

This literature review has been conducted in the BCI space, to analyse the various methodologies adopted to determine the effectiveness of EEG headsets and their associated machine learning algorithms. Research was undertaken to explore the multitude of techniques that are being used by the academic community to produce accurate and reliable BCIs, forming the foundations upon which the proposed aims and procedures in this thesis are built. This review will investigate studies by independently examining each stage of a BCI, namely *brain imaging modality*, *feature extraction* and *classification* and then finally, exploring potential and existing applications of such systems today.

2.1 Brain Imaging Modality

2.1.1 Types of Brain Imaging Modalities

Brain Imaging modality is the method of recording neural activity through either invasive techniques including using micro-electrode arrays (MEA) and electrocorticography (ECoG), or through non-invasive techniques including functional magnetic resonance imaging (fMRI), near-infrared spectroscopy (NIRS) and electroencephalography (EEG). Each approach demonstrates arguably supportive reasoning for BCI implementation but current research lays focus on non-invasive techniques emphasising the need to avert issues concerning operational risks, electrode sustainability and the constant need for a team of experts to maintain functioning of the system (Gerven et al., 2009).

A study performed by Hochberg et al. (2006), incorporates a neuromotor prostheses (NMP) to allow a quadriplegic subject to perform challenging tasks, such as playing a video game, to demonstrate the effectiveness of invasive methods as they requires little training from the subject. It is conducted by implanting electrodes directly to the brain to detect voluntary neural activity at the motor cortex and asynchronously sending electrical signals to a simplified computer interface. It addresses the benefits of harnessing invasive procedures indicating that significant learning is not required, data is comparatively (to non-invasive systems) easier to process and signals can be used to reanimate any paralysed limbs as they directly engage the motor cortex of the brain for control. Leuthardt et al. (2004) agrees with this notion, summarising ECoG as an 'excellent modality' providing better noise to signal ratios, wider frequency range measurements, superior stability and signal quality and higher spatial resolution compared to that of EEGs. However unlike Leuthardt et al. (2004), Hochberg et al. (2006) discusses noteworthy shortcomings of using invasive procedures including: limited scalability to other parts of the brain to provide improved system learning, cosmesis, cost and efficacy considering the incorporated levels of risk.

Studies conducted in the past decade recognise the benefits of using EEG and thus have made it the most common non-invasive method for recording neural activity with Hwang et al. (2013) finding that 60% of BCI research implemented EEGs in the period of 2007-2011 with this number growing significantly. Whilst some studies have utilised fMRI and NIRS methods, typical drawbacks including poor temporal resolution and lacking the ability to measure neural activity throughout the full brain have deemed them less popular (Gerven et al., 2009).

2.1.2 Electrode Properties

EEG's benefits of affordability, safety, mobility and high temporal resolution rely on the existence of its limitations. While EEG's applications resonate largely in providing disabled healthcare, its use also extends to healthy users as an extension of their physical capabilities. Social acceptance thus is a prevalent issue as EEGs lack aesthetics and require strenuous preparation time. Preparation requires the use of conductive gel, introducing issues of possible short circuits between electrodes, and hence increased impedance of EEG recordings as the gels dry up with time (Ferree et al., 2001).

To address these concerns, Lin et al. (2008) demonstrate that dry electrodes can be designed to achieve readings that are as equally accurate as wet electrodes and overcome these limitations. Lin et al. (2008) conduct an experiment to evaluate the feasibility of using a neuroprosthetic system to monitor fatigue levels of subjects in long-term driving scenarios through continuous feedback of their task performance. In this study, the authors concurrently explore the efficacy of a dry electrode implementation by undertaking this experiment with wet and dry electrodes. While they design their own micro-electromechanical (MEMS) EEG sensors, they contrast the impedance of their dry electrode with regularly used wet electrodes and highlight the lower impedance and higher signal quality of the dry electrodes at frequency ranges of 0.5-150 Hz (see Figure 2.1).

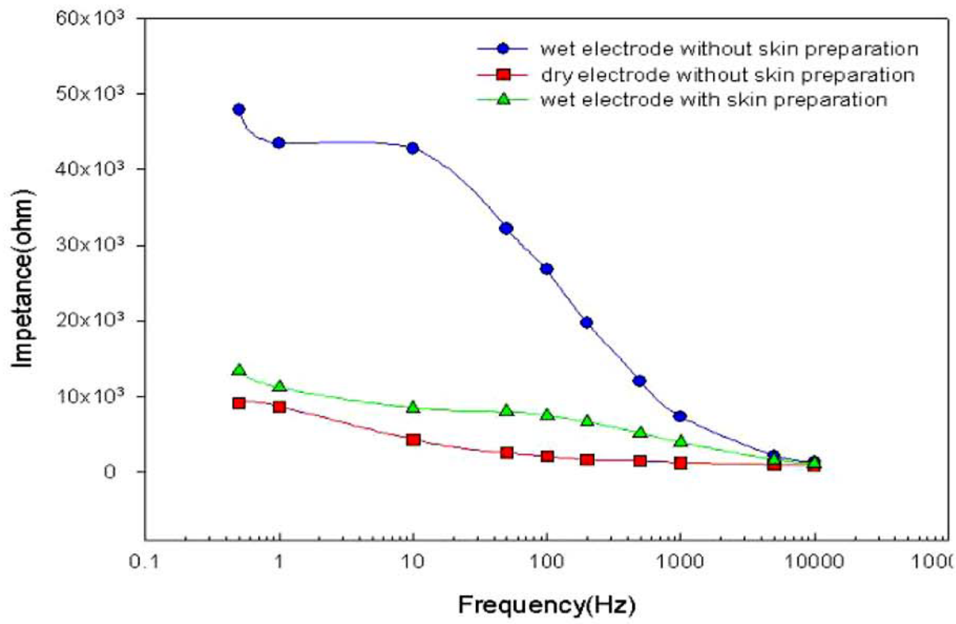


FIGURE 2.1:
[1]Comparison of Impedance for Wet and SEMS Dry Electrodes (Lin et al., 2010)

Lin et al. (2008) used 10 subjects to participate in the investigation, as they drove a car in a VR-simulated environment for 1 hour whilst investigators recorded their neural activity, with 5 electrodes on their forehead. Data processing involved using: a 90s moving average filter to eliminate variance, logarithmic scaling to linearise the EEG amplitudes, Karhunen-Loeve Principal Component Analysis (PCA) to extract features and multiple linear regression to estimate the driving error time. Their results can be observed in Figure 2.2 where authors described the wet and dry electrode signals to be "virtually identical" and exemplify their similarities with the MEMS dry EEG sensors producing an overall mean accuracy of 89.6% compared with 89.1% with the wet electrode.

Despite these positive results, the authors inadequately identify issues surrounding the use of their MEMS EEG as it is of high cost, causes pain and discomfort for patients as a result of penetrating 2 layers of the skin, exposes the risk of infection and is subject to impedance mismatch between electrodes as microscopic structures in MEMS are known to frequently split (Lopez-Gordo, Sanchez-Morillo, and Valle, 2014). Further, as the MEMS design uses micro-sized needles it consequently will generate poor readings in the presence of hair, as hairs are of thickness 50-100μm (Bhushan, Wei, and Haddad, 2005). Lin et al. (2010) only tests their electrodes on the scalp, in the absence of hair and don't discuss the challenges encompassing artefacts in their signals (i.e. muscle movement, blinking) - a common issue experienced by academics in this field.

[1]Impedance is spelt incorrectly by this source

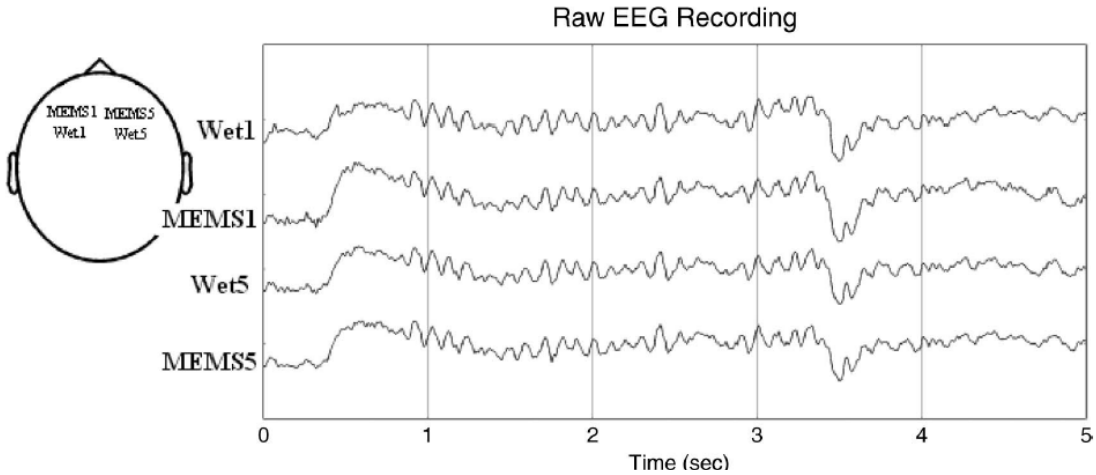
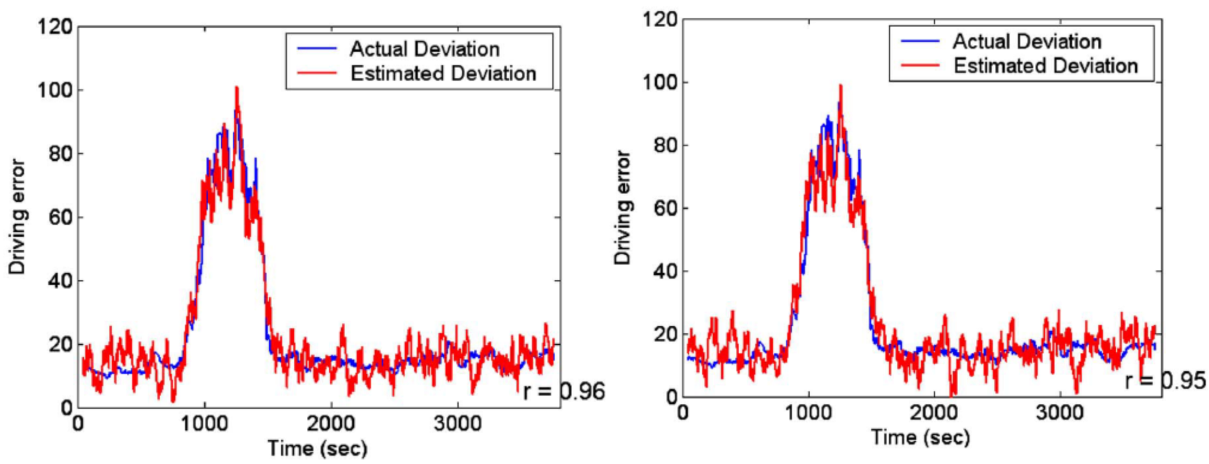


FIGURE 2.2: Dry vs Wet Electrode Raw EEG Signals (Lin et al., 2008)



(A) Wet-Electrode Estimate

(B) Dry-Electrode Estimate

FIGURE 2.3: Wet vs Dry Electrodes (Lin et al., 2008)

Lin et al. (2011) attempts to address the shortcomings of studies that employ MEMS electrodes by proposing a dry electrically conductive urethane electrode, covered in a conductive fabric to tackle other frequent issues such as sweat. An impedance spetroscopy methodology is employed, performing 19 tests on five subjects, to analyse and contrast the impedances of the wet and the proposed foam electrode on the forehead and test it's ability to deal with hairy regions of the skin. To test the signal quality, a linear correlation coefficient function was used between the wet electrode and dry electrode signals, and subjects were instructed to walk with the EEG headsets to investigate the challenges of artefacts' presence in signal data. Results indicate that the impedance of the dry electrode is the same, if not better than the wet electrode due to the foam's ability to flex and maintain contact with the skin (Figures 2.3a and 2.3b). It further verifies the inverse relationship between the impedance of the electrode and its surface area, a discovery made earlier by Ferree et al. (2001), and showed signal quality was correlated to the wet electrode data, in this case by 96%. Noise was also less prominent using the dry electrode after 17mins of walking, when there were overall more distinct artefacts present in the both signal electrodes' data.

Comparatively, Guger et al. (2012) assess the implementation of the dry electrode against the wet electrode in the P300 BCI spelling task further described in section 2.2, spelling "L-U-C-A-S". Implementing the LDA algorithm to detect the P300 ERP peaks captured some discrepancies that resonated in the result trials between the dry and wet electrode. The P300 ERP peaks from the dry electrode were noticeably lower, and suffered signal drifts below 3Hz, yet did not draw a large impact on the final classification of results and accuracies of the task.

2.2 Feature Extraction and Classification

2.2.1 Paradigms and Accuracy

Depending on the experimental paradigms used, different brain signals are typically induced and the characteristics of those brain signals are defined as a *features* or *signatures*. Features are usually analysed using machine learning algorithms and then cross validated to generate results indicating the quality of the features and/or the effectiveness of the machine learning algorithm. Features for EEG paradigms are typically categorised into evoked and induced responses. An evoked response is one where a stimulus is presented or a task is performed and the brain signals are recorded, with signals being phase-locked to the event whereas an induced response is one where brain signals are recorded based on (usually) a spontaneous stimulus-related change where the significant change in amplitude of the brain signals during this event is recorded (David, Kilner, and Friston, 2006). Features derived from induced responses are known as event related potentials (ERP).

The P300 signature is an ERP whereby a positive deflection of brain waves occurs 300ms after a subject has been exposed to some stimuli and is generated using a variety of paradigms. Using visual stimuli to evoke the P300 have demonstrated advantages in BCI studies as they have revealed less training time and illustrated a higher transfer rate of information of up to 70bits/min compared to paradigms generating evoked responses such as motor imagery (Parini et al., 2009). Research around the P300 feature has mainly been implemented by using the Farwell-Donchin paradigm (Farwell and Donchin, 1988) , asking subjects to spell a word and select each letter of the word when the letter is visible to them in a matrix of size 6x6 containing random numbers of letters with rows and columns being flashed at random, see Figure 2.4. Curtin et al. (2012) presented the use of this paradigm to train subjects to navigate in a 3-D virtual environment using an EEG headset. Their experiment implemented the Farewell-Donchin paradigm with directions and commands in the form of icons rather than letters and numbers, where subjects would count the number of times a specific icon appeared. Functions such as look left, right, turn around etc, were implemented as icons and then were subsequently used as tools for users to navigate through simple and complex mazes. The methodology utilised a 40-channel EEG headset, recording measurements from 9 electrode sites sampling at 1kHz and employing a band pass filter. Processing the data using a stepwise LDA, analysing the P-300 ERP recordings, yielded a mean accuracy of 91.4% of correct actions/intentions over 9 subjects after 10 sequences of recorded data. Although this highlights a positive outlook of the P300 feature, the reliability of this experiment is limited, as it lacks indication about the time delays encountered when users made decisions, and houses a biased subject class, constrained to 9 right handed males within a small age bracket of 19-23. Additionally, it, alongside most other studies used healthy subjects from similar demographics to perform their analysis. Therefore they draw a questionable methodology, particularly for those who have the intent of studying BCIs to improve the quality of life for those suffering from disabilities (Manyakov et al., 2011).

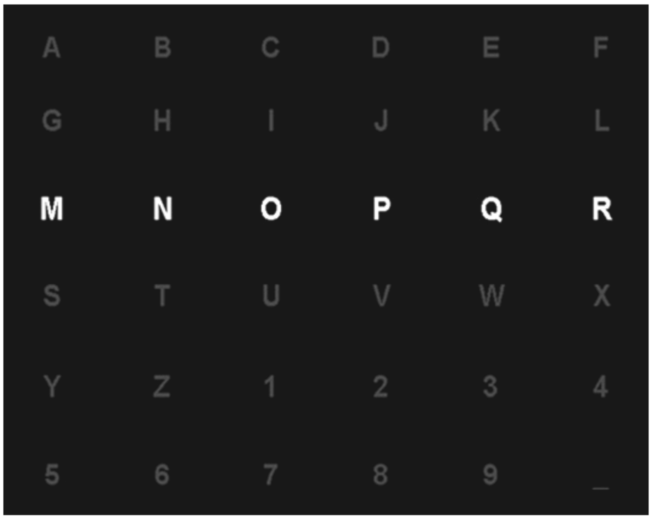


FIGURE 2.4: Illustration of the 6x6 Farewell-Donchin paradigm (Fazel-Rezai and Ahmad, 2011)

A comparative study was conducted by Li et al. (2010) who used motor imagery and analysed the P300 features, with a class of users (5 male, 1 female) aged 20-31, to provide 2-D cursor control with an EEG acquisition device. Users' brain signals for vertical movement, were recorded by asking them to focus on a displayed up or down button inducing ERP P300 signals, and for horizontal movement, were recorded by asking them to imagine left/right hand movement. Their study found that the motor imagery paradigm was less useful in training the machine, having difficulties in recognising horizontal movement compared to the P300. They also presented that the average accuracy was 90.75% and the average time taken to hit a target with the cursor was 28 seconds. These results denote that algorithms for P300 detection are time consuming and yet similarly to Curtin et al. (2012)'s study, fail to justify a generalisation, by using predominantly right hand-males within a small age group. Brunner et al. (2010) is a more comprehensive study, as it used a balance of 6 females and 8 males ages 23-41 and explored BCI Illiteracy, a concept whereby brain signal analysis fails to generate positive results for certain individuals. By exploring this concept with the subject class, it concluded that using a hybrid paradigms (using more than one paradigm) may reduce the effects of BCI illiteracy.

2.2.2 Universality

Universality describes being able to apply a concept to all subjects. BCI illiteracy hinders universality in the BCI field, with many studies reporting that an estimated 20% of subjects could not demonstrate BCI performance levels for any effective control (Blankertz et al., 2008), (Sannelli et al., 2008). Typically, subjects are deemed illiterate if their classification accuracy falls below an acceptable rate, usually defined at 70% as Perelmouter and Birbaumer (2000) highlight this as the minimum accuracy required to achieve reasonable levels of communication. Blankertz et al. (2006) recognises the importance of better training methods, subject instructions and more efficient signal processing but claims that this will only reduce BCI illiteracy rather than achieve universality.

Allison et al. (2010) presented a study to extend the possibility that BCI illiteracy may only occur if users are being tested for a single BCI feature. In this study the authors implemented a different approach by comparing individual features (ERD/ERS and SSVP) with a hybrid BCI, incorporating ERD/ERS and SSVEP BCIs speculating that the additional information throughput from both paradigms may overcome universality. The process involved placing 14 subjects in front of a television and asking them to imagine opening and closing their left/right hand using the imagery cues on the television, for the ERD/ERS exercise. In the SSVEP exercise, subjects were asked to focus on the active left/right oscillating LED light and in the hybrid task, subjects were asked to conduct both of the these tasks at the same time. Results were cross validated and the authors found that for each individual task, there were 5 subjects who were BCI illiterate in each task, with 3 of those illiterate in both tasks. The hybrid BCI on the other hand, was able to, on average, classify the activities more accurately than each individual task, as demonstrated in the box plot (Figure 2.5), and eliminate any form of BCI illiteracy.

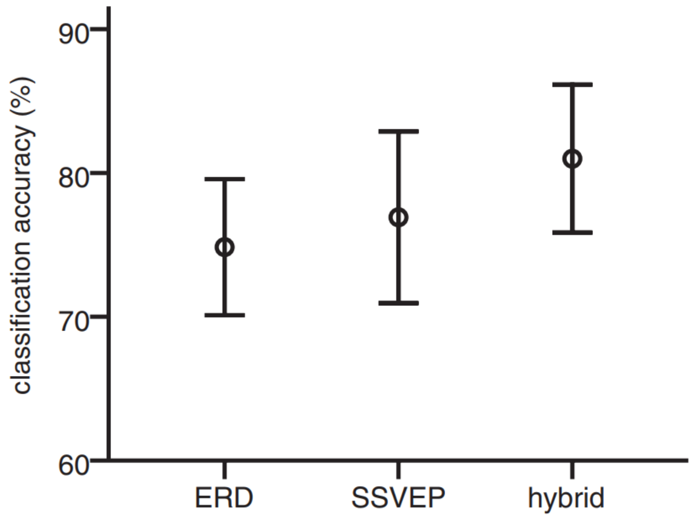


FIGURE 2.5: Classification Accuracies of Hybrid vs ERD vs SSVEP (Allison et al., 2010)

It should be noted that a hybrid BCI can present the challenge of finding complementary interfaces as Hinterberger et al. (2004) discovered that using a secondary system (in their case auditory communication) generated inimical properties in the data to the primary system (visual communication) and therefore reduced the accuracies of their overall results. In this instance, the authors believed the success of alleviating BCI illiteracy with a hybrid system was attributed to the increased concentration of the subjects as they had to perform two tasks instead of one. Furthermore, these findings can be deemed controversial as subjects may have alternated between tasks rather than performing them simultaneously and since hybrid tasks were conducted last, users may have become more accustomed to the exercise. For more conclusive findings, the authors would have had to ask subjects to attempt the hybrid scenario on a different day or utilised other hybrid tasks. Nonetheless, investigators highlighted that for one of the subjects, the machine learning algorithms were able to effectively ignore ERD features as the SSVEP features provided better information for the machine to generate increased accuracies. Consequently, the hybrid system was able to pick the best of all the information generated to yield optimal results for each individual subject.

Other hybrid developments have been explored through the analysis of evoked and induced response with Sato and Washizawa (2016) establishing that using N100 as a feature (a negative ERP generated 100ms after stimuli) alongside a P300 analysis, improved the classification accuracy of the BCI system. As the experiment was conducted by reducing the number of counting tasks in the Farwell-Donchin paradigm, because N100 is evoked rather than induced based, there was also a reduction in user fatigue - a common concern raised when training subjects for BCI investigations.

2.3 Classifiers' Accuracies and Associated Features

With signal processes improving their abilities to deal with the anomalies and setbacks of EEG signals, drastically over the last two decades, academics have shifted their focus to finding optimal combinations of classifiers and feature sets to generate higher levels of classification accuracy in a BCI system (AlZoubi, Calvo, and Stevens, 2009). Therefore the different systems and their properties are studied to establish a thorough understanding of which combinations are currently most optimal for BCIs.

In order for BCIs to process data into meaningful commands, machine learning algorithms need to be implemented through the use of one of the two methodologies: regression or classification. Regression models such as logistic linear regression are developed by using probabilities to determine a logistic function and estimate the relationship between independent and dependent variables (Ilyas et al., 2016). Classification algorithms are used to classify the EEG signals into numerous or binary outputs where the outputs are discrete unlike continuous outputs of a regression output.

NTo comprehensively understand which classifiers have been most successful Lotte et al. (2007) have reviewed the various classification algorithms to compare their various properties and performances. The classifiers tested in this study are Linear Discriminant Analysis (LDA), Support Vector Machine (SVM), Artificial Neural Networks (ANN), Non-Linear Bayesian Classifiers and Nearest Neighbour Classifiers. Lotte et al. (2007), reveals that linear classifiers such as LDA are commonly used to great success in many BCI applications but face limitations when implemented with EEG data that is non-linear in nature. Neural networks are therefore the choice of method proposed by the authors, but highlight the model's nature to overtrain the data and therefore lack versatility when testing various subjects. Bayesian classifiers and Nearest neighbour classifiers are typically not used in BCIs as, respectively, they are unable to perform in real-time and suffer from the curse of dimensionality.

Using the aforementioned investigated properties, Lotte et al. (2007) used existing literature to find the accuracies of different classifiers and their respective features in the following categories: movement intention based BCI, mental imagination based BCI, motor-imagery based BCI and the Farwell-Donchin P300 spelling based BCI. Motor imagery performed the weakest, ranging in accuracies from 61-89% and was outperformed by the movement intention and mental imagination categories that demonstrated accuracies over 90% using ANN and linear classifiers. They remarkably found that classifiers in the P300 category achieved accuracies of 100% multiple times using raw EEG data features with Linear and

Gaussian filters, assuring the practicality of this result by emphasising that these algorithms are robust, simple and regularised. The following literature similar endorsed such results: Hoffmann et al. (2005), Rakotomamonjy et al. (2005), Kaper et al. (2004), Bostanov (2004)) but showcased misleading interpretations, as the studies' methodologies were identical (dismissing the element of robustness) and less than 5 subjects are tested to signify generality. In particular, Hoffmann et al. (2005) examined a physically challenged subject who was not able to exhibit the same definitive response as his tested healthy counterpart, suggesting this was a byproduct of his disability. Additionally, left-handed subjects and females were omitted from the testing pool, who typically display disparate signals to healthy male subjects.

With respect to other specific classifiers, the article demonstrates Bayesian Classifiers to be the worst performers for BCI investigation except in the motor imagery category. Neural networks were able to display near-perfect accuracies exercising a variety of features including gamma, beta, alpha and raw EEG waves with no bias towards a specific feature though Li, Xu, and Zhu (2015) found that they were able to differentiate between gamma wave patterns more explicitly when subjects were in a "concentrating" state of mind. Other studies such as Hiraiwa, Shimohara, and Tokunaga (1989) utilised readiness potentials[2] as a feature, from voluntary joystick movement to achieve 100% accuracy, and Palaniappan (2005) who computed spectral power from EEG readings and applied a single layer ANN reporting classification results of up to 97.5%.

Even so, with ANN drawbacks suffering from lengthy computing times, linear classifiers are acknowledged as the most suitable for real-time BCIs accommodating most features as different features do not cause drastic impacts on final results. Literature recognises that algorithms are powerful enough to address specific BCI feature issues including data outliers (Chiappa and Bengio, 2003), curse of dimensionality (Kaper et al., 2004), time information and smaller training sets (Raudys and Jain, 1991).

[2]Bereitschaftspotential

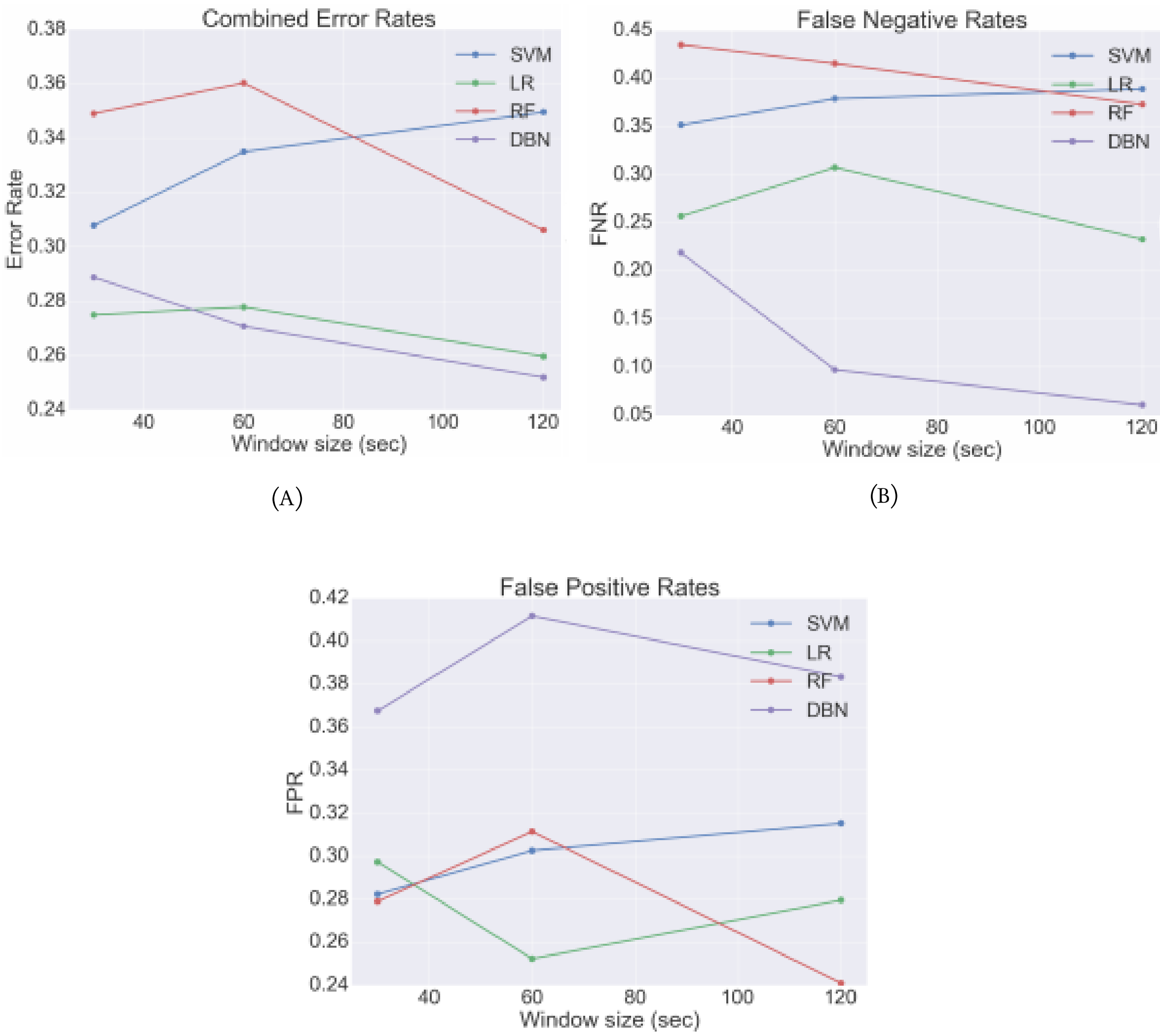
2.4 BCI and Applications

The BCI serves a dual purpose. That is to improve the quality of life for those who suffer from some form of disability by enabling them to communicate with others, and to expand the capabilities of society by extending the applications of the brain. Research today incorporates new technology into its methodologies and assessing the extent to which these new technological advancements are able to enhance the capabilities of a BCI.

Bashivan, Rish, and Heisig (2016) explored the capabilities of newly developed EEG technology by attempting to distinguish two mental states of subjects: logical/rational and emotional with a commercial Muse headset. Thirteen subjects were placed in a quiet room and were shown academic videos and then cat videos to exhibit 'focused' and thereafter 'elated' responses. To generate features from their raw EEG data, a 6 level discrete wavelet transform (DWT) was used. DWT is the process of applying a linear convolution between a signal and low and high pass filter to generate approximation co-efficients computed by using a high pass sub-band (named a detailed level) and a low pass sub-band (named the approximation level). These co-efficients can then have another DWT applied to it again, which means a 2 level DWT has been applied (giving a lower frequency range in DWT Level 3). In this study a 6 level DWT is applied and for each co-efficient generated, there was a corresponding frequency band associated with the EEG Raw Data. For example, DWT Level 7 corresponded to a frequency range of 0 - 3.44Hz, which indicated this band was comprised of delta waves. Using these co-efficients, the authors found the average and variance of power for each frequency range to utilise as features for classification. They employed SVM, AAN, and Logistic regression to classify their data.

Their data results showed that SVM achieved the lowest classification error at 25.5% noting that classification errors tended to reduce as the window recording period increased. However, when performing a 13-fold cross validation[3] the AAN and logistic regression modeled better results than the SVM. Where false positive rates measures categorising data as a cat video (incorrectly) Figure 2.6 highlights that the ANN (named DBN) has a bias towards categorising cat videos where as logistic regression provided a more balanced error.

[3] See Glossary for definition



(A) (B) (C)

FIGURE 2.6: Error Rates of various Classification Algorithms using a Muse Headset (Bashivan, Rish, and Heisig, 2016)

Sources of these errors were varied and encompassed issues that were particularly attributable to acquisition of data from electrodes as electrical activity generated large volumes of noise caused by facial muscles such as squinting, blinking and jaw movements. Further, the minimal number of electrodes, four meant that the volume of data collected was also limited and in the presence of artefacts makes it more challenging to obtain relevant data. The study may have found that implementing a different feature methodology or using multiple features in the presence of such noise, to be a more efficient method of determining the mental states of subjects.

A more successful study by Rebsamen et al. (2006) incorporated robotic systems to control a wheelchair with EEG data analysis for patients suffering from motor neuron disease. Here, a combination of the P300 paradigm supplemented with motor imagery was used for feature extraction and classified using an SVM algorithm. The classification score generate by this algorithm was converted to a probability and compared to a threshold to estimate whether a P300 signal had been recorded. If a low threshold was used, the system would compute higher classification errors but would have a faster response. Therefore a trade-off had to be made between speed and accuracy, specifically as the wheelchair was designed to operate in real time. Additionally, it was found that heavy concentration had an effect on machine error due to the large presence of alpha and beta waves, consistent with findings from Vourvopoulos and Liarokapis (2014), who found users unable to control a robotic system used to navigate in a virtual world, when trying too hard.

2.4.1 Application of BCI with consumer grade headset

Studies trending towards high accuracy BCI application indicates the increased demand for consumer grade EEG headsets as literature has evidently demonstrated that subjects are generally restless during test conduction and in those cases, noise generation has generally been a prevailing issue (Hari and Puce, 2017), (Kropotov, 2010). To address such concerns, wireless headsets that are more comfortable, portable and possess lower resolution electrodes to capture stronger signals have been increasingly used in more recent investigations. Two studies by Li, Xu, and Zhu (2015) and McCrimmon et al. (2017) illustrate this modern theme focusing on EEG's suitability and reliability in more useful real-time BCI applications.

These investigations used headsets that made use of four electrodes as opposed to 32 (standard EEG cap), similarly to Bashivan, Rish, and Heisig (2016) but used electrode placement to complement the method of training for the machine learning algorithms. McCrimmon et al. (2017) used four electrodes that were placed over the motor cortex of the brain (Figure 2.7) to recognise the SSVEPs induced by the brain whilst also exploring the differences between a 32 channel expensive EEG and a four channel cheap EEG headset. Li, Xu, and Zhu (2015) quite similarly, used a Muse headset with four electrodes but only collected data from the electrodes sitting on the frontal lobe as subjects trained the machine learning algorithms through concentrating and relaxing - an exercise where the primary involvement is induced by the frontal lobe.

McCrimmon et al. (2017) tested the performance of the custom four electrode EEG by asking subjects to open and close their fist for six seconds and hypothesised that it would demonstrate results similar to the conventional 32 channel EEG headset. This was a consequence of McCrimmon et al. (2015)'s findings as they discovered no significant loss in decoding performance [4] of the signals when comparing 32 channel to four channel electrodes.

[4] decoding SSVEP patterns

They used a microcontroller to decode the presence of motor-task related modulation [5] using alpha and beta wave bands as they were the most distinguishable features when conducting any motor task and discarded the first second of data to eradicate any brain activity induced whilst transitioning between mental states (open to closed fist). Decoding involved using LDA and PCA to extract hand movement features from the highly dimensional data and then implementing a Bayesian classifier to compute the probability of a state movement. The investigation was also repeated if the classifier found that accuracy was below 85% and in these odd cases, an additional two minutes of calibration was performed.

Their cross-validation results showed that overall the custom headset was able to achieve comparable accuracies to the conventional headset but compromised in computation time as it exhibited delays of up to 2.5 seconds. The custom headset displayed cross validation accuracies around a mean of 93.6% while the conventional BCI was able to detect a mean accuracy 96.2% (with lower variance in test results). These results, though promising, portrayed limited reliability as the motor task performed is quite simple but the decoding process is complex and 'mobiles may struggle to decode the SSVEP' (McCrimmon et al., 2017). The authors' decision to remove the one second state transition and re-test any data that scored less than 85% is also not a realistic demonstration of how BCIs would be used in the real world as headsets would have to cope with unwanted data. Further, the authors suggest that more complex tasks and elaborate movements require a different approach as the computing power required for their analysis would be too high.

Contrarily, Li, Xu, and Zhu (2015)'s use of the Muse headset with the SVM algorithm and Feedforward Neural Network model (predicting 0.5s forward) revealed more assuring results at a conviction rate of 88% piggybacking off selective results and favourable method construction. This assurance stemmed from a live application of their EEG model, asking subjects to control a simulated aeroplane in a video game by concentrating to move it up and relaxing to move it down, and then requesting feedback from their participants about their experience - which was overall positive. However, as algorithms produce poorer results when subjects tire, the methodology favoured positive results as it incorporated long rests between each test, and similarly to McCrimmon et al. (2017) the authors only accepted test samples that achieved accuracies above 80% - they were both biased against any form of weak results. Therefore it is apparent that while EEG BCI investigations are highlighting more usable systems that be incorporated into our everyday lives, the methodologies used question the reliability of their results in the real world.

[5] power decrease in alpha/beta wave bands

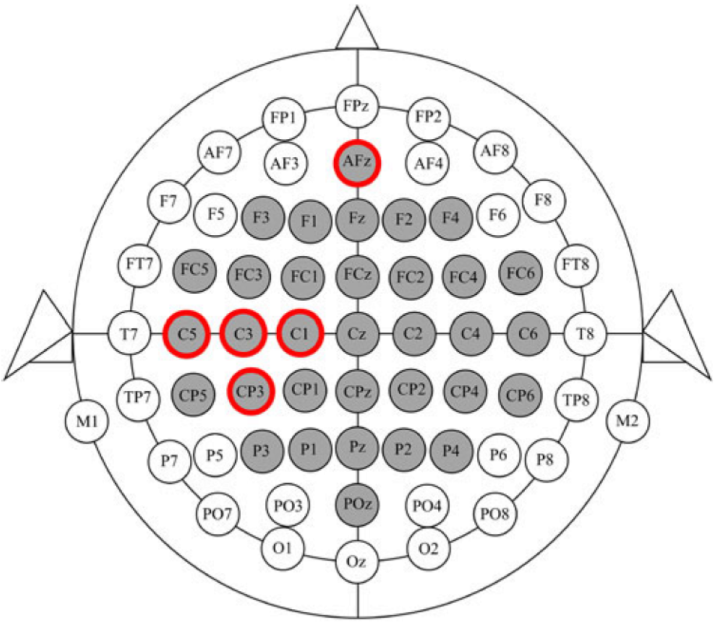


FIGURE 2.7: Electrode locations for the international 10-10 EEG system. Conventional 32 channel EEG electrodes in grey, custom EEG in red. (Bashivan, Rish, and Heisig, 2016)

2.5 Literature Discussion

It is clearly evident that there are absences in literature regarding the practical applications of BCIs with mixed results in feature selection and machine learning algorithms and practical application missing the mark on real world every day use. As these experiments demonstrate highly scientific attributes, focus has been diverted to finding the best features and classifiers in achieving accurate results rather than addressing issues surrounding how:

- EEG can become more integrated into everyday life
- Efficient machine learning algorithms perform in people's natural environment, and
- More natural tasks can be exploited with the application of machine learning and EEG.

The lack of studies in the implementation of commercial grade EEG products such as the Muse used by Bashivan, Rish, and Heisig (2016), is most likely due to their recent introduction to the public. Further exploration of these products in this research paper alongside open source machine learning methods will reduce some of the gaps that are prevalent in the study of BCI systems today.

3 Methodology

There components required to prepare, conduct and analyse the outcomes of this thesis were all designed purposefully to comprehensively examine the usefulness and reliability of today's consumer-grade EEG headsets. Programs were written in various programming languages capturing the strengths that different languages provide, to maximise process efficiency and overcome a wide array of language-specific hurdles. The BCI in particular was compiled in *Java* for possible extension of this study, whereby live testing could be implemented on a mobile device. The components and the associated languages include:

1. Construction of the Brain-Computer Interface - *Java*
2. Conduction of Experiments - *Java*
3. Data Filtering and Results Construction - *VBA, Matlab*

3.1 Brain-Computer Interface Setup

3.1.1 High Level Structure

The BCI was designed with a more complex structure than one would expect from a simple device in order to retrieve data. The 2016 Muse headset limited the utilisation of direct interfacing between a PC and the headset as the SDKs were not written to extract raw EEG data for the new low-energy bluetooth technology implemented in this edition. Further, the Muse developers advised not to use Windows' APIs in attempt to establish a connection as there is 'large variation in third-party hardware' (Developers, 2017). Therefore the structure used to connect the Muse headset is diagrammatically represented below in Figure 3.1. It illustrates that the Muse headset directly interfaces with an Android App that was specifically designed to read all the raw EEG data that was being recognised by each electrode on the headset. This data was then live streamed across a Wi-Fi network (using a specific port) to a computer. The computer would then ping the target server port to establish the existence of any data, and then send to the Java programs where they were stored in a CSV file for further analysis.

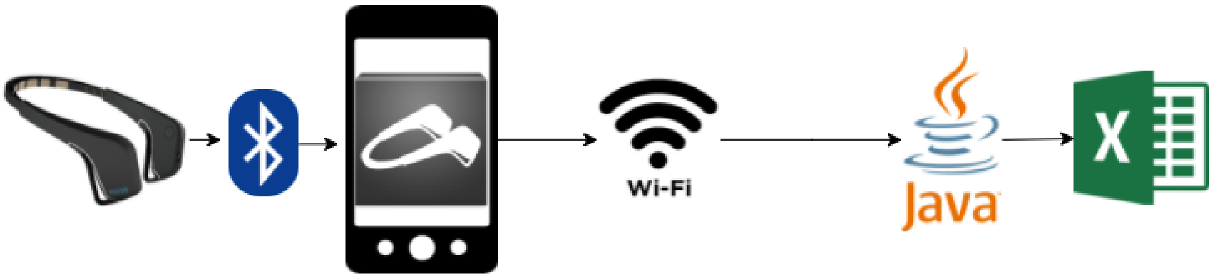


FIGURE 3.1: High Level Structure of BCI with Muse Headset

3.1.2 EEG Headset - Muse Headset

The selection of the Muse EEG headset was based on a variety of factors that were required to make the experiment as "practical" as possible. With the intent of using the headset in future applications, literature has usually criticised the potential of dissemination and support for healthy *and* physically challenged subjects. These have included 'ease and convenience of daily use, cosmesis, safety, reliability,usefulness of the BCI applications in the user's daily life, and the need for ongoing expert technical oversight' (Mak and Wolpaw, 2009). Consequently, the criteria entailed finding a headset with minimalistic features incorporating a convenient fitting process and a high quality and robust design, that could be easily programmed to extract raw EEG data. The additional criterion, *Simple Design*, was added to this list to accommodate usability of the product in daily applications to the extent where its employment has little to no physical or aesthetic impact on a person. Table 3.1 highlights the choices of headsets and their ability to fulfill the set criteria.

Headset	Ease of Wear	Preparation Time	Robustness	Simple Design	Electrode Quality	Raw Data Extraction	Movement Noise
Emotiv Epoc	x	x	✓	x	✓	x	x
NeuroSky	✓	✓	✓	✓	x	✓	x
Emotiv	✓	✓	✓	✓	✓	x	✓
OpenBCI	x	x	✓	x	✓	✓	x
Muse	✓	✓	✓	✓	✓	✓	✓

TABLE 3.1: Criteria for Headset Selection

From Table 3.1, is evident that the Mindwave, Insight and Muse headsets are the most suitable choices for the experiments. However, the Neurosky was designed with a single electrode that was quite sensitive to any head movements, which meant subjects would have to remain absolutely still and any experimental hypotheses designed to test the optimal location of a single or number of electrodes could not be performed. Additional reviews from Berg (2012) supported this notion as it described ear movement and excessive blinking impacting the quality of the sensor's data. The Emotiv Insight was the clear choice with it's features exhibiting far superior qualities to its counterparts including 5 electrodes with 14 bit resolution, but the developers provided limited tools to access raw data; users had to pay for the volume of data used. The Muse headset, displayed in Figure 3.2, was therefore chosen, giving free access to raw data with 4 electrodes in the AF7, AF8, TP9 and TP10 positions, refer to Figure 3.3 - Fpz refers to a reference electrode position. The electrode positioning is advantageous, as it enabled the comparison of using temporal lobe electrodes and frontal lobe electrodes in isolation.

(A)

(B)

FIGURE 3.2: The Muse 2016 EEG Headset used in experiments

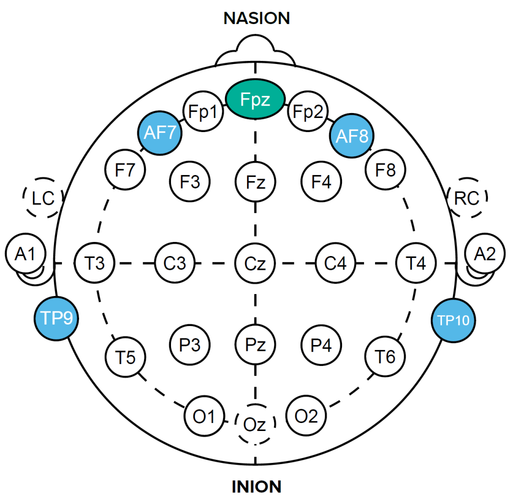


FIGURE 3.3: Muse electrode locations by 10-20 International Standards (Developers, 2017)

3.1.3 Muse Monitor App

The Muse Monitor App was designed by a Muse enthusiast for extracting raw data from the Muse headset. As it embeds many practical features such as Wi-Fi streaming, live data view, pre-processing into brain wave frequencies and a signal strength indicator, it was a suitable choice for implementation as opposed to attempting to build a custom android application. Muse Monitor uses Low Energy Bluetooth Technology available on a smartphone to conduct communication with the headset and then allows the implementation of the Open Sound Control (OSC) protocol to stream data between the smartphone and the computer over a Wi-Fi network.

It sends OSC packets to the computer in a bundle containing multiple messages with the first 2 bytes defining the bundle number and the time tag. In each bundle there are a number of messages that illustrate what the incoming data is and its corresponding data type through a specified *Address Pattern* and *Type Tag String*. Thus all the data that is sent can be identified by the computer. This is particularly important as not all information is sampled and sent at the same rate from the headset. E.g. Raw EEG data is sampled at a rate of 220Hz whilst Band Powers are sampled at a rate of 10Hz. *Band Powers* for a particular frequency range (ie. Alpha waves at 9-13Hz) are calculated as 'the logarithm of the sum of the Power Spectral Density of the EEG data over that frequency range' (Developers, 2017).
To establish a connection between the smart phone and the computer, both devices must be on the same network and the local host server must be listening on port 5000.

3.2 Java Programs

Java programs were designed with the classes listed in Table 3.2 for three main functions, two of which are described in this section. These three functions are to:

1. Present visual stimuli for participants of the experiment to engage in
2. Acquire the incoming data over the duration of the experiment
3. Train the Machine Learning algorithm and cross-validate the results to find the accuracies at which binary outputs can be achieved

The program starts by asking the user to name the test and training data files for experimentation/testing. Once provided it gives the user a choice to either train the data or test the data. Training the data is effectively running the experimentation on the subjects whereas the testing of data is to cross validate the training files or to conduct live testing [1] using

[1] live testing is not discussed in this thesis, but is designed in the Java based BCI

the specified training files. Once the user has chosen, the oscP5 protocol will start, starting the server and opening port 5000 to listen for any incoming messages that are being sent by the Muse headset. Whilst these functions execute the user is required to elect which programs they would like to use. If they are experimenting (training), they will select one of four visual stimuli. These stimuli are as follows:

- *Yellow Vs Red*: The screen will go blank with the exception of an 'x' in the centre and will alternate between the colours, yellow and red, at random time intervals (0.5s - 4s). The participant counts the frequency of yellow whilst focusing on the cross in the centre.
- *Yellow Vs Colours*: Similar to *Yellow Vs Red*, a blank screen with an 'x' at the centre will change colour. However in this test, several colours are used, as opposed to only red and participants are instructed to count the frequency of yellow observed.
- *Yellow Vs Visual Distractions*: A small box is presented at the centre of the screen with a black 'x' in the middle. The background outside of the box remains black whilst the colour of the box changes at random time intervals (0.5s-4s). Concurrently, two other boxes of random size and colour continue to appear and disappear around the screen, while the participant counts the frequency of yellow for the box in the centre of the screen.
- *Blue Vs Red*: An identical stimulus to *Yellow Vs Red* however participants are instructed to find the frequency of blue.
- *Yellow Vs Red with Audio*: Although this is not one of the four selectable options, it is included as an experimental stimulus. The *Yellow Vs Red* test is selected here and in the background the soundtrack 'Sorry' by Justin Bieber is played for participants.

Classes	Function
csv2arff	Converts *.csv* files into *.arff* for WEKA to process.
CSVWriter	Writes the information into the relevant *.csv* files based on the type of program running. It also writes the headings at the start of the files.
GUI	Displays all the GUIs and visual stimuli for the test subjects. It handles the actions for all objects on the GUI when they are activated (ie. Buttons). and indicate the progress of programs through dialogue boxes.
infoPassed	Determines the information that is going to be written into the information files. It also builds the strings to be written into the files.
mainApp	Initialises all the classes and runs the programs. It also indicates the status of the connection including when data is being received from the headset.
mlalgo	Creates all the instances from the training files and applies the machine learning algorithms. It will test data as well using either the cross validation method or by parsing a test and a training file.
MuseData	Sets the subjects state of mind (Concentrating vs Relaxed) and performs the assigning and retrieving of data to and from specific variables.
MuseOscServer	Initialises the interrupts in the event an oscMessage has been received and sets the data-specific flags to true, informing the user that the requested data has been received.

TABLE 3.2: Java Classes used in the development of Experimental Testing and Cross Validation

3.2.1 Data Acquisition

The ***MuseOscServer*** class is responsible for establishing a connection with the Muse headset. In order to conduct this, the Muse Monitor application on the mobile device must be running, connected to the Muse Headset and streaming data to the computer's IP address at Port 5000 (instructions located on the applications website). An external Java library *oscP5* is imported to define OSC specific functions, namely *oscEvent* and *oscMessage*. *oscEvent* is a function that takes oscMessage as an input and determines the type of data that has been received. It sets the flags corresponding to the data received and places it into a structured array labelled "musedata". Each row of this structure represents each electrode of the Muse headset. The Raw EEG, absolute alpha, beta, gamma, delta and theta waves are stored as floats for 7 point precision.

When all required data is received and the array is populated, the function *write2file* in ***infoPassed*** will write all the data to the training file allocated by the user. The training file is a comma separated value (csv) file that stores raw data as presented in Figure 3.4 below, to be further cross-validated once data acquisition is complete. As data comes in at different frequencies, the Java program writes to the training file one line at a time (ie. with ALL of the electrode data). More frequently sampled information such as Raw EEG values are updated as new values are received, to prevent stale data when writing to files. Further, to avoid time lags in reading, writing and processing, multi-threading is implemented throughout the code. This allows the training data to be labelled when published, indicating whether the participant is observing the colour yellow (blue).

Alpha0	**Beta0**	**Gamma0**	**Delta0**	**Theta0**	**RAW0**	**A**	**B**	**G**	**D**	**T**	**R**	**A**	**B**	**G**	**D**	**T**	**R**	**A**	**B**	**G**	**D**	**T**	**R**	**State**
1.0445	1.18	1.180913	1.4436	0.6885	868.3	0	0	#	#	#	#	0	0	#	1	0	#	1	1	1	1	1	#	Relaxed
1.0426	1.16	1.182203	1.4243	0.6783	853.8	0	0	#	#	#	#	0	0	#	1	0	#	1	1	1	1	1	#	Relaxed
1.0426	1.13	1.178524	1.4204	0.6875	850.6	0	0	#	#	#	#	0	0	#	1	0	#	1	1	1	1	1	#	Relaxed
1.0426	1.09	1.178524	1.4204	0.6875	858.6	0	0	#	#	#	#	0	0	#	1	0	#	1	1	1	1	1	#	Relaxed
0.7719	0.64	0.725373	0.4856	0.0252	849.4	1	1	0	2	1	#	1	#	#	1	1	#	1	1	1	1	1	#	Concentrating
0.7889	0.66	0.740187	0.669	0.1452	869.5	1	1	0	2	1	#	1	#	#	1	1	#	1	1	1	2	1	#	Concentrating
0.8128	0.68	0.748823	0.7956	0.24	843.3	1	1	0	2	1	#	1	#	#	1	1	#	1	1	1	2	1	#	Concentrating
0.8318	0.69	0.746856	0.8642	0.2918	850.2	1	1	0	2	1	#	1	#	#	1	1	#	1	1	1	2	1	#	Concentrating

FIGURE 3.4:
[2]3 seconds of Raw EEG Data, highlighting data from only Sensor TP9 - Subject 3)

[2]Relaxed is a state in which the user does not visualise yellow, Concentrating is a state in which the user visualises yellow

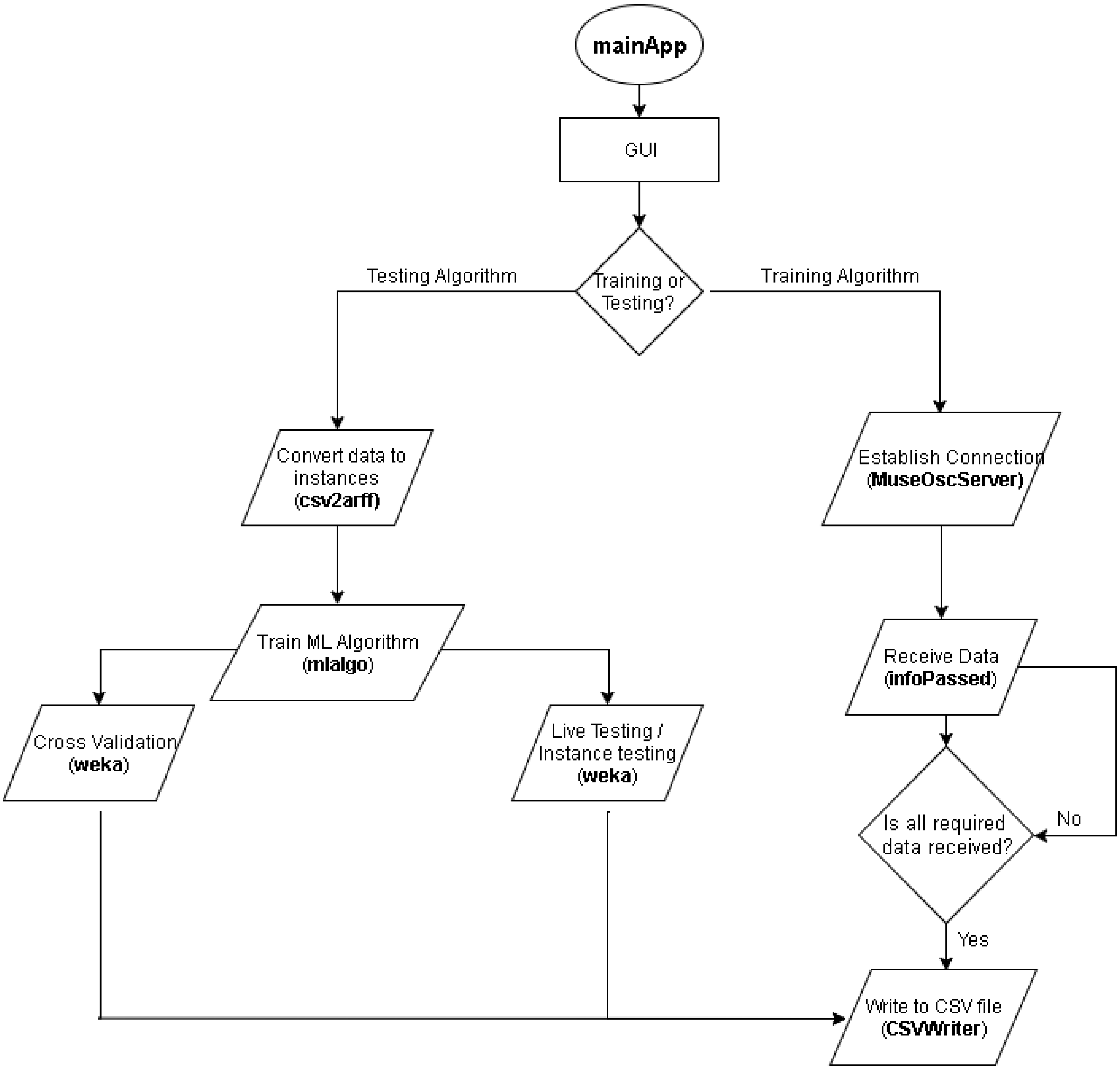


FIGURE 3.5: High Level Overview of Training and Testing ML Algorithms for further Cross Validation

3.2.2 Development of Visual Stimuli

Training Programs Selection

The training programs were designed to encapsulate the notion of recognising and thinking about a light bulb through the use of colours and shapes on a screen. Recognising a light bulb switching on and off would become a variant of the commonly used neurological examination: Go/No-go whereby participants are to recognise when to "go" or "not go" based on the stimulus that is presented to them. Here, responses were elicited by displaying yellow amidst other colours and distractions, to encourage thoughts of a light bulb switching on particularly to simulate an environment where there are activities regularly transpiring around them. The distractions simulated include the presence of alternative colours, moving objects and aural presence. As Yoto et al. (2007) demonstrate in their paper, colour has a significant impact on people's brain activity as it arouses particular neurological responses such as elevated alpha activities in the frontal lobe when subejcts are anxious or having rising levels of attention. This fosters predictable patterns that can be more easily sought out for the machine to recognise. It should be noted that the program selection showcased is not to elicit the most prominent response, rather to establish a series of recurring brain activity that a person would exhibit in a natural environment and state of mind. The *Yellow Vs Red* stimulus was designed as a standard control which would most likely indicate differences in brain activity because red usually causes anxiety and heightens participants alpha and theta waves (Yoto et al., 2007). *Yellow Vs Red* is a written ***GUI*** where a timer is used to switch the colour of the screen from yellow to red and vice-versa. The timer will generate a random number every half a second and perform a modulo operation with the number three. In the instance that the random number is perfectly divisible by 3, the colour will change (to red/yellow) whilst the subject counts. After 120 iterations (60 seconds), the timer will stop and display the number of times that the screen turned yellow. Subjects who have poorer concentration levels, or are more easily distracted will calculate incorrectly and will be instructed to concentrate harder on the upcoming tests.

Yellow Vs Colours is constructed to make it harder for subjects to count the yellow appearances as they cannot pre-empt the next colour or the frequency. This test was designed in a similar manner to *Yellow Vs Red* using the modulus three algorithm to determine colour change. Colour changes would however, occur on any even random number. If the number was even and divisible by 3 then the screen would turn yellow.

Yellow Vs Visual Distractions was the most complex of tests, for subjects to undertake and to develop. The *GUI* class had an in-built subclass named *Square* which would construct an array of 2 square objects which would be utilised to move around the screen. This class defined the colour of the square, and the location of the square on the screen, taking into account that different computer screens have different resolutions. In the instance that the test may have to be replicated in alternative environments with different screen resolutions, the screen bounds were found and then utilised to ensure that these moving squares always remained within the screen. Once again, the modulo 3 algorithm was used to alternate colours for the 2 moving squares and the centre box. Upon initial testing, it was found that subjects did not find the moving squares distracting enough and thus the test was upgraded to have 3 separate timers and incorporating size changes for each of the two moving squares. The implementation of 3 different timers meant that squares were moving more randomly and would be able to instigate increased concentration levels for individuals while they counted the colour frequencies.

Blue Vs Red and *Yellow Vs Red with Audio* were designed on the backbone of Yellow Vs Red simply changing the code to display blue contrary to yellow and commence music simultaneously to the beginning of the test. The *Blue Vs Red* stimulus was used to test the effectiveness of the colour yellow. Although yellow is justified with thoughts encapsulating a light bulb, it is possible that other colours may be better at driving a result than yellow. Audio distractions are also implemented as a stimulus to emulate the manner of background noise that subjects would ordinarily encounter through the day when performing medial tasks such as flicking a light switch. It also serves the purpose of finding the impact on a participant's temporal lobes as it is subjected to stimulus that precisely activates this region.

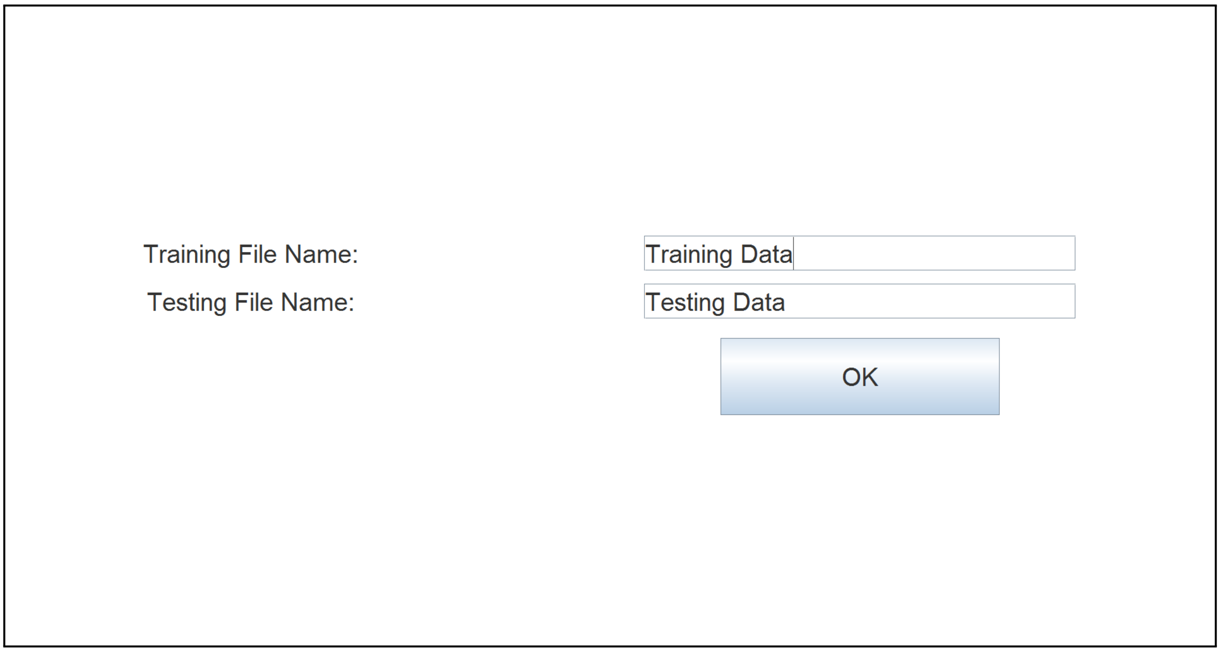


FIGURE 3.6: File Namer GUI

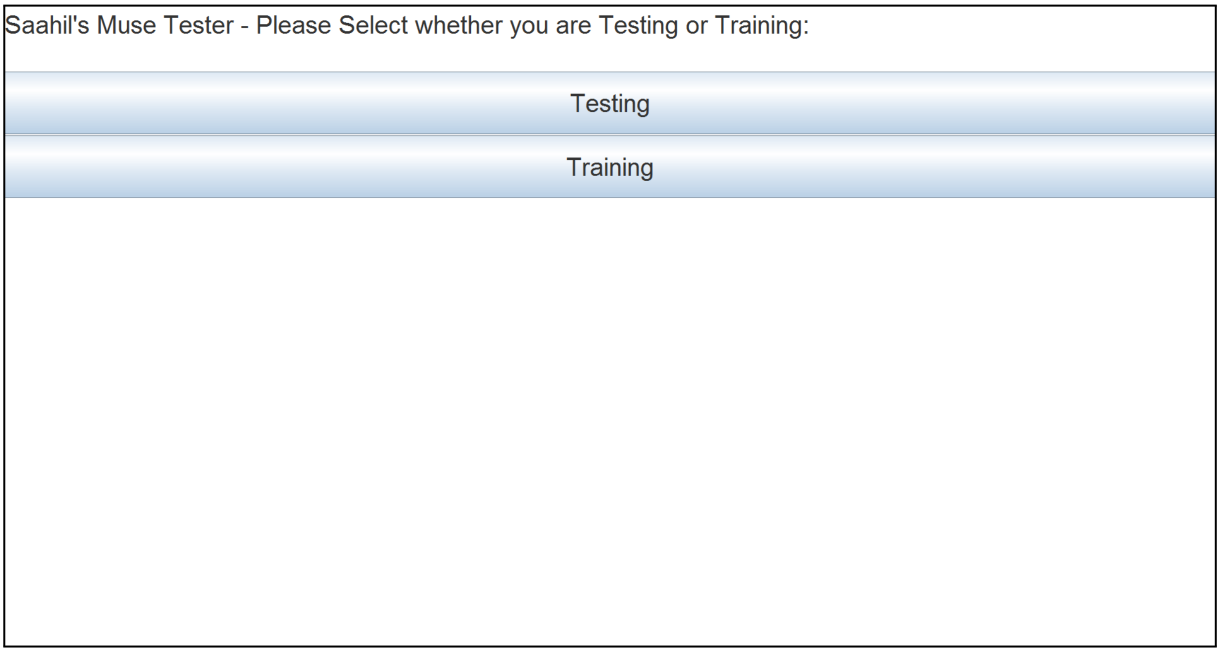


FIGURE 3.7: Training vs Testing the Machine Learning Algorithm GUI

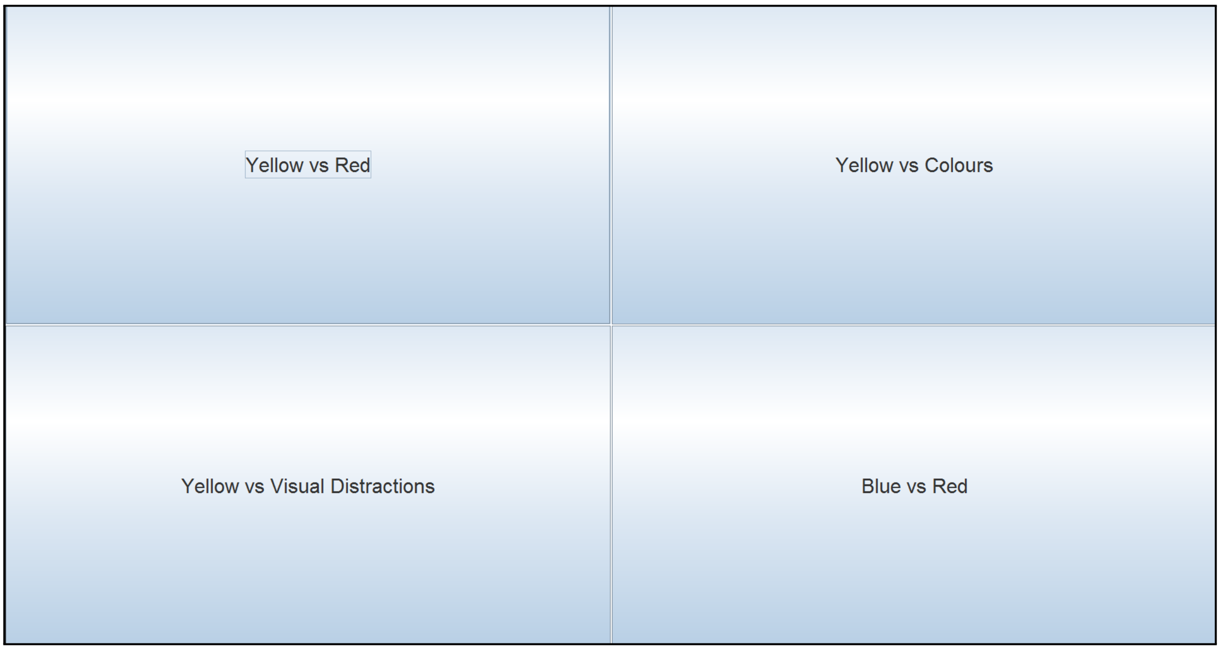


FIGURE 3.8: Training Stimuli Selection GUI

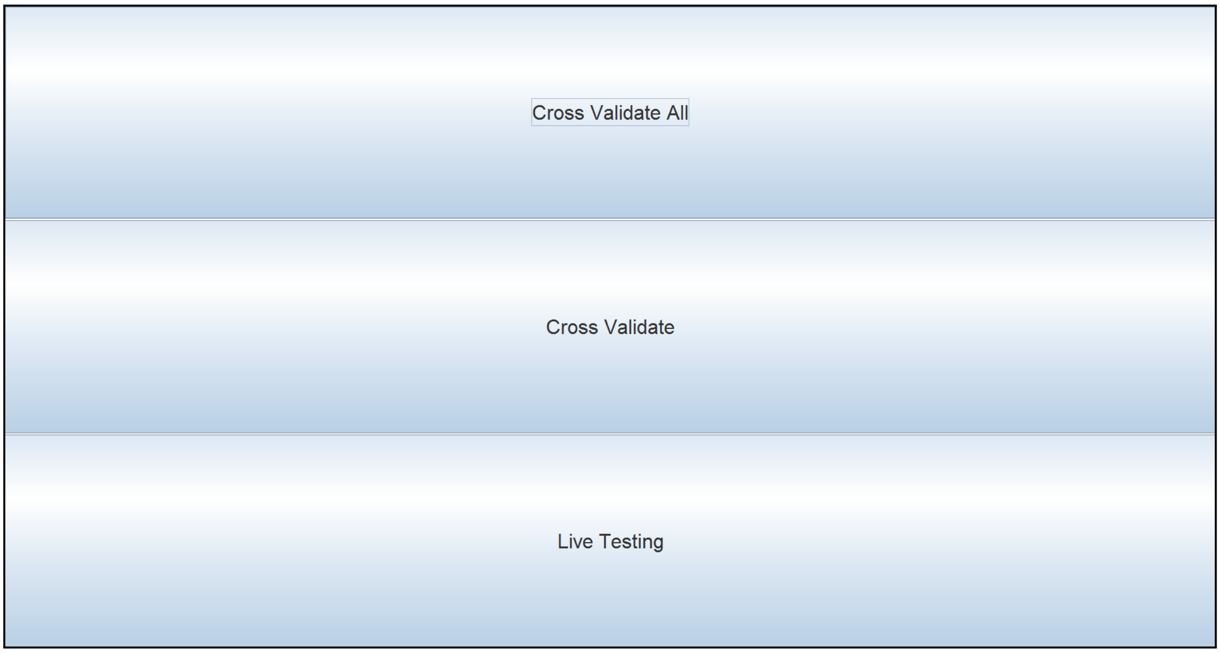
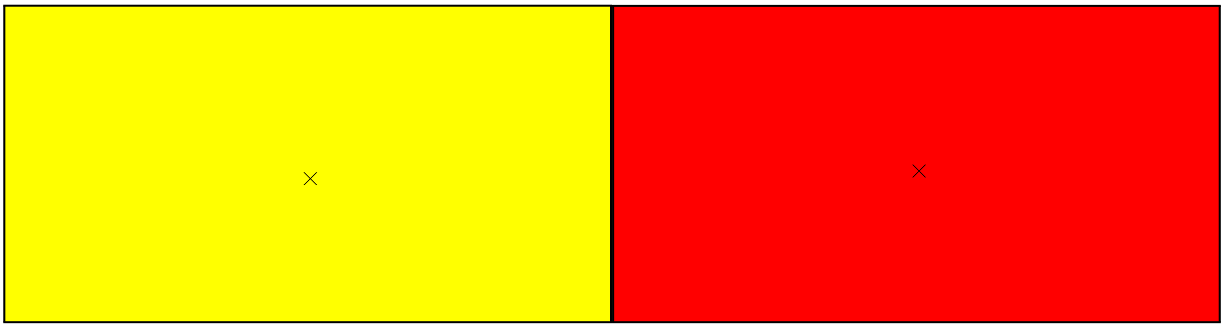
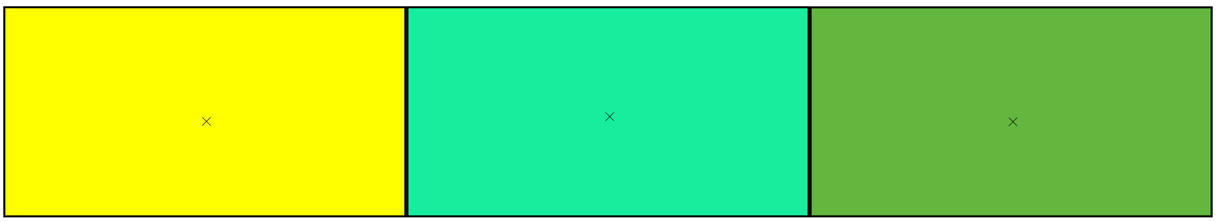


FIGURE 3.9: Testing Stimuli Selection GUI



(A) (B)

FIGURE 3.10: Subject Machine Learning Training Exercise: *Yellow Vs Red*



(A) (B) (C)

FIGURE 3.11: Subject Machine Learning Training Exercise: *Yellow Vs Colours*

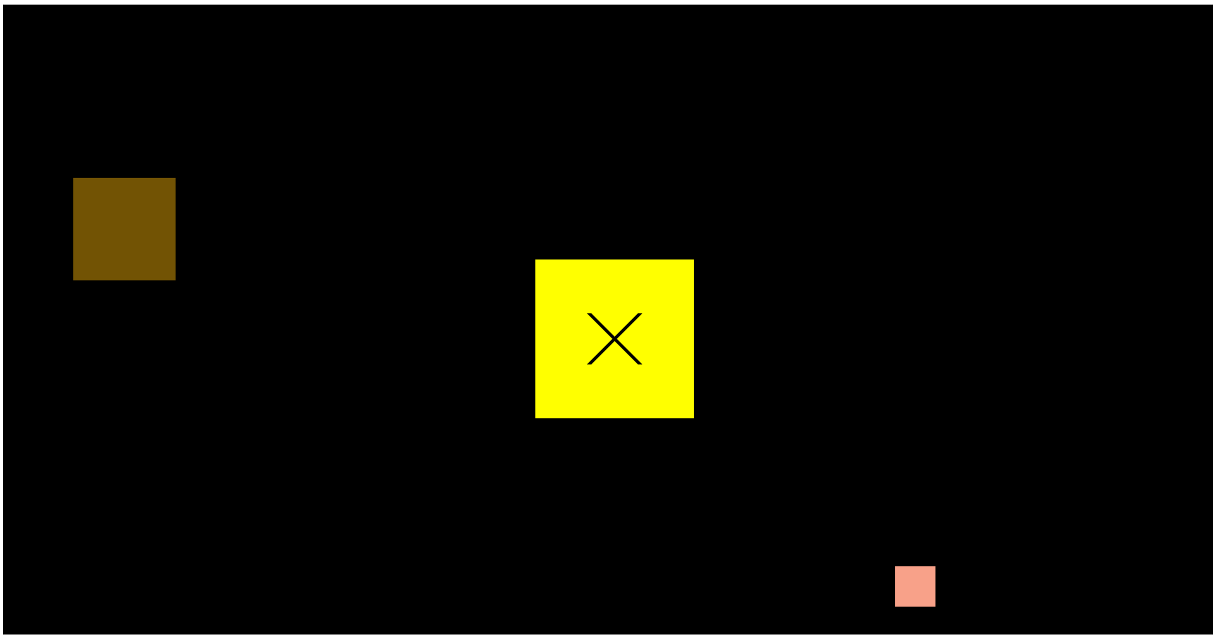
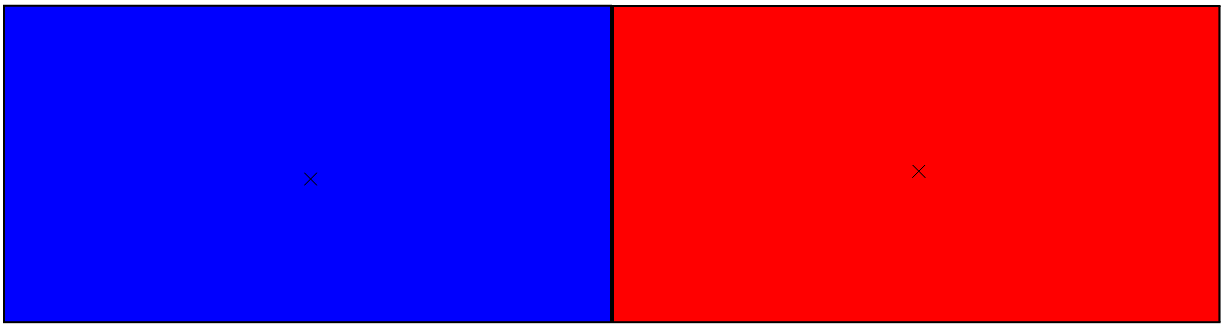


FIGURE 3.12: Subject Machine Learning Training Exercise: *Yellow Vs Visual Distractions*

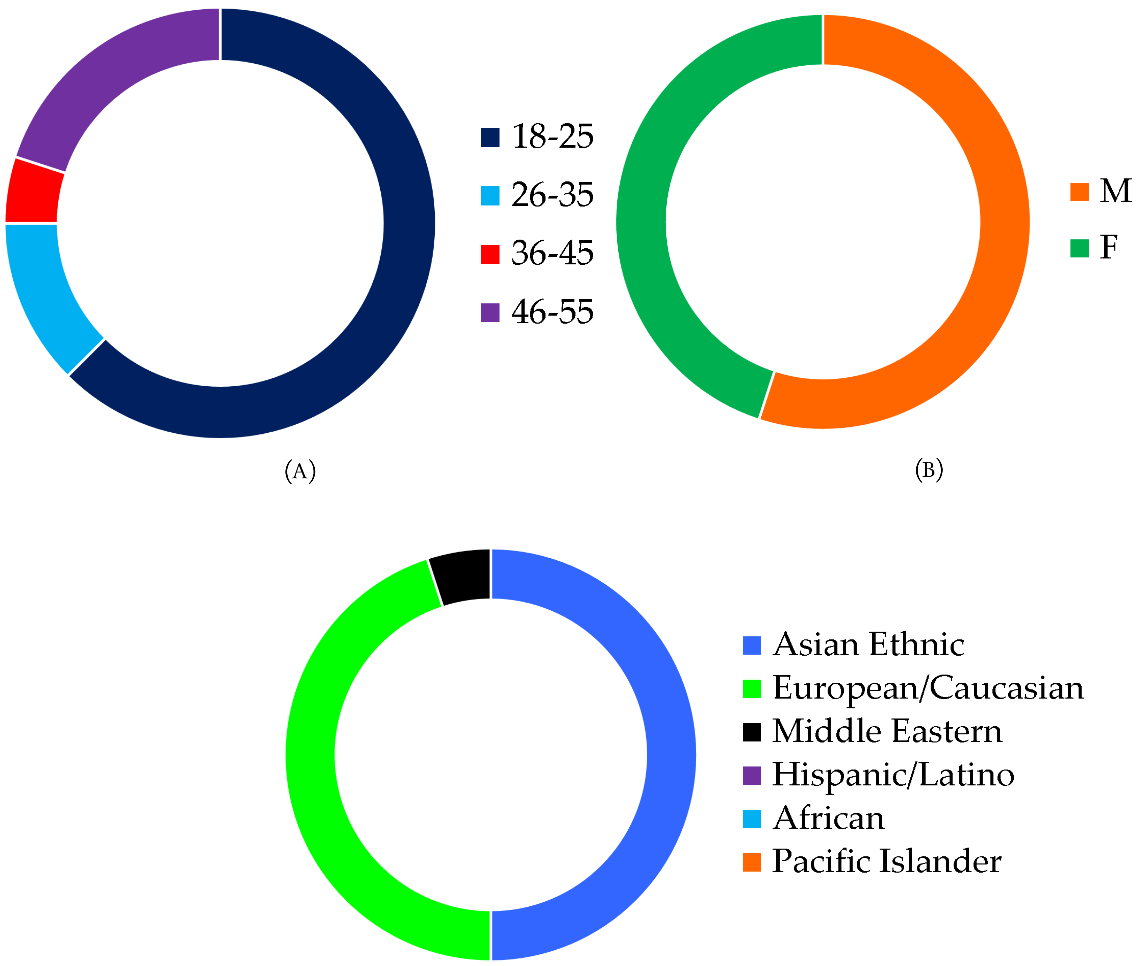


(A) (B)

FIGURE 3.13: Subject Machine Learning Training Exercise: *Blue V Red*

Subject Testing

For this process, a total of 39 subjects were used altogether. As a wide variety of literature has been criticised for having subject bias' including increased number of males and a younger population, the 39 test subjects have been uniformly chosen with their demographics shown in Figures 3.14a, 3.14b and 3.14c.



(A) (B)

(C)

FIGURE 3.14: Subject Characteristics: Age, Gender, Demographic

It should be noted that the training data collected from subjects undertaking the visual tests encompass a variety of uncontrolled variables which will be examined in the results and discussion. This lenient approach may not abide by scientific reasoning, that is, to narrow down variables that have a specific impacts but it does illustrate how a low-cost BCI system with consumer grade hardware can fare in an amalgamation of environments. The community will never be able to simulate all the environments in which products will be exposed to and for this purpose the methodology used to engage subjects is not within strict parameters. Furthermore, 5 randomly selected test subjects were exposed to 120 seconds of testing as opposed to 60 seconds of testing to investigate the effects of an increased number of observations.

Firstly, no subject was informed about the purposes of the test and the results that were trying to be achieved beforehand. This ensured that subjects could not "cheat" or intentionally produce outcomes to skew data in/against this papers' hypotheses. They were placed in an environment exposed to limited background noise, though not absolutely quiet.

Each task was explained before the visual cues were presented and participants were told to remain calm before the test began. The process of being calm involved discussing their week whilst they took deep breaths for 1min. This ensured that all subjects were not anxious when instructed to count the yellow frequencies. Calming down participants of the thesis trials were more so aimed at cases with elder persons as they had assumed their intelligence was being tested and seemed nervous.
The following questions were asked prior to subjects participating in the thesis tests to investigate possible sources of outliers when conducting machine learning training:

- What is your age?
- Would you classify yourself as Middle Eastern, Asian Ethnic, European (Caucasian), Hispanic/Latino, Pacific Islander or African?
- Are you right handed or left handed?
- Is there any history of mental disease in the family?
- Did you get adequate amount of sleep last night?
- Do you suffer from any mental illnesses?
- Would you classify your attention span as short (1min), medium (5min), long(>10min)?

3.3 Longitudinal Tests

Longitudinal tests were also performed on two subjects, one male adolescent (Age $\leq$25) and one female adult (Age $>$ 25). In this extended time series test, these two subject participants were asked to perform the 5 tests, seven times over the span of 2 weeks at variable times of the day. They were asked to do it at any time to capture a spectrum of moods and to also incorporate the impact of varying times of the day. The aim of this exercise was to test how well the machine learning algorithm can perform at various times of the day under participants' moods. Quantifying participants' moods is subject to interpretation and the time of day parameter most likely would be somehow correlated to this. Therefore I did not analyse trends around those variables, rather used the best universal machine learning parameters found in normal subject testing to compare how longitudinal tests performed next to one-off tests. The pre-processing and cross validation methods outlined below remain the same for longitudinal tests.

3.4 Machine Learning Algorithms and Data Processing

3.4.1 Pre-Processing

Before any of test subjects' raw data can be placed into the machine learning algorithm, to find how accurate the machine learning algorithm will be in generating a binary output, faulty or corrupt data is removed from all the files. Matlab is used to pre-process this data on a large scale with Microsoft's VBA allowing csv data filtering to remove specific rows of data on an individual test scale.

Matlab looks for the all the training files that are provided in the folder and calls a macro within an Excel file. This macro is designed to find rows where more than 1 data point in that row has a value `0.0000` (4 precision floating point). This result occurs when the Muse headset has not retrieved a strong enough signal from the headset and therefore writes the value `0` into the *csv* files. This is the main cause for fluctuations in number of observations presented in Results.

To encompass the effects of frontal lobe and temporal lobe impacts independently they were analysed with their own filtered sensor data (ie. without the inclusion of other sensors in each instance). Therefore two new data sets were created including either only TP9 and TP10 sensor outputs or AF7 and AF8 sensor outputs. Microsoft Excel/Matlab was used to remove any redundant sensors' data for this additional analysis.

3.4.2 WEKA and selection of Machine Learning techniques

WEKA is the collection of open-source machine learning tools that has been used in this study for pattern recognition and feature selection. Features are usually referred to as attributes (in particular with ML programming) and each line of data that the algorithm uses to train the software is called an instance. WEKA has been designed to seamlessly integrate into Java programs with extensive Java documentation providing insight into methods of implementation for machine learning techniques within the thesis' customised software. WEKA's algorithms include Simple Linear Regression/LDA, MultiLayer Perceptron (Neural Networks), Bayes and Linear SVM. These are some of the most widely used algorithms however Linear SVM was chosen in order to provide highly accurate results in real-time with the Muse Headset. The reasonings for this selection are outlined.

Simple Linear Regression regresses the outputs of y against x to forge a statistical linear relationship, which can be optimised by transforming either variable through simple functions. It has the main advantage of low computational power and thus speed, but as it is quite simple it makes a large number of assumptions that cannot be easily applied in the instances that are provided by the EEG headset. The normal distribution assumption as an example will be violated when using raw data signals. During the process of the tests, the likelihood of any of the features tested (alpha, beta, gamma, raw waves etc) holding a mean value about which values are distributed is extremely low.

Neural networks are one of the highest performing algorithms using complex structures and relationships between nodes to generate accurate results. They also allow algorithms to be more tailored and customised to suit the requirements of the data, through additional of neural layers and nodes. However even with powerful computers, the time taken to arrive at these results is out of the scope for a real-time BCI system. Over-fitting is also an issue that recurs in neural network algorithms and due to the dynamics nature of brain activity to identical visual stimuli it was inappropriate to select any variant of neural networks.

Naive Bayes Classifiers are also widely used to specifically categorise instances of data accurately based on conditional probability particularly as they use a "probabilistic approach" and because of their high speech. The independence assumptions associated with this algorithm, though, are illogical as brain activity is usually correlated with numerous shared factors including the stimuli provided. Further, Bayes algorithms perform poorly in real-time as similarly to neural networks, they perform too slowly for an efficiently functioning BCI.

Lastly linear SVM, or known as SMO (in WEKA-only), uses a linear decision boundary to classify training data by maximising the functional margin. That is, the distance from the decision boundary to training samples nearest to it. However with the use of a kernel, data points can be artificially mapped into a 3-D space, giving rise to a 3-D/non-linear boundary. This is known as‘kernel trick’. A kernel is an arbitrary function that defines where each data point lies in the additional dimensions. The kernels utilised in this thesis are the Gaussian (RBF Kernel) and Polynomial Kernel (PolyKernel) as the PolyKernel tends to under-fit data because of its simplicity and the RBF Kernel slightly over-fits data because of its complexity. Further the parameters associated with these kernels can be adjusted to alleviate the challenges of over or under-fitting. Consequently it is able to bypass sluggish computation time commonly attributed to higher dimensionality problems and thus SMO algorithms are most suitable within this higher dimensionality brain-machine learning environment.

ML Algorithms	Advantages	Disadvantages
Linear Regression	• Low computational power • High speed	• Simple • Makes a large number of assumptions
Neural Networks	• Complex, thus achieve exceptional results • Can flex parameters	• Slow • High computational power
Naive Bayes Classifiers	• Based on conditional probability • Tend to achieve quite accurate results	• Illogical independence assumptions • Perform poorly in real time
SVM/SMO	• Speed • Can use kernels to increase complexity of algorithm	• Have to select appropriate kernel for individual use

TABLE 3.3: Comparing the Pros and Cons of Various Machine Learning Algorithms

3.4.3 SMO Parameters

The SMO complexity parameter is used to control how soft margins are between the decision boundary and the nearest training points. The C parameter is defined as the characteristic which informs the algorithm how much the user wants to avoid misclassifying data points. A larger value of C will choose very small margins, attempting to classifying all the points correctly. Think of this as a line of best fit that weaves in between points to ensure it has adequately separated the training data as opposed to drawing a just straight line with one or two data points that have been misclassified. Figure 3.15 highlights a neat comparison between the performance of the SMO algorithm where complexity parameters are high and low.

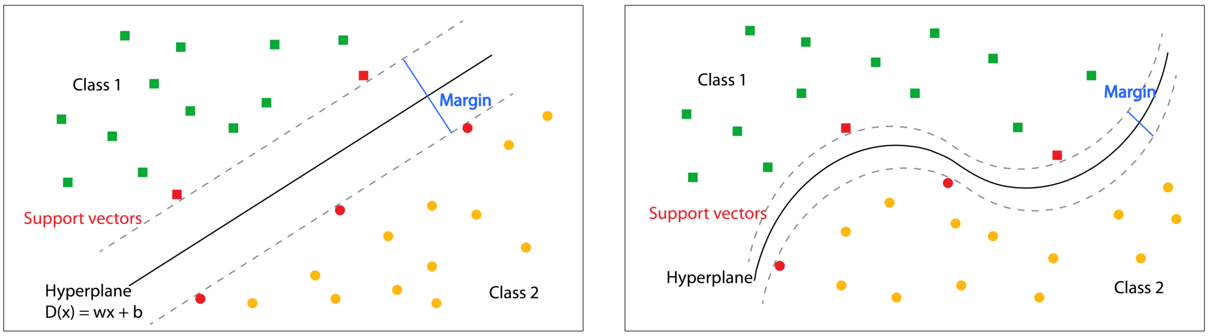


FIGURE 3.15: Contrasting a lower complexity value (left) with a higher complexity value (right) (*Classifier Parameter Optimisation* 2017)

The gamma parameter defines the influence of each training point on the final decision boundary. A smaller value of gamma will only capture the influence of points very close to the initial estimated decision boundary, and thus will not capture the complexity of the data as the region of influence from each data point is limited.

It is important to find a balance between these two parameters (RBFKernel) or in the case of PolyKernel, polynominal degree and complexity factor, in order to classify the data in the most effective way. High complexity values and small gamma values may draw a "perfect" decision boundary given the training set, but will fail when new data is introduced to the mode. This is the result of over-fitting the SMO model. A combination of parameters are used to find the most suitable in each test subject and in each test conducted within this investigation.

3.4.4 Implementing the Cross Validation Algorithm

The class used to test the training data is **mlalgo** with methods used to convert data to the supported formats, cross validate the model with a multitude of parameters and then further store the results in a suitable manner.

The first step in training the data is to convert it from the *.csv* file it is in to a *.arff* file. The University of Waikato that developed the ML algorithms describes this file as "an ASCII text file that describes a list of instances sharing a set of attributes" (Waikato, 2017). The SMO algorithm then splits the trained data into its attributes using the data in the *.arff* format provided and then trains the model. The model can be trained in a number of ways, depending the on the type of parameters it is given. These parameters are set to optimise the hyperplane for the support vectors. Optimise is not defined as the best line of fit for each instance but the best combination of parameters that will identify classes the most accurately for all data files. The parameters used for each of the kernels aforementioned are the exponent parameter, complexity parameter, and gamma parameter. The definitions of these parameters are presented in Glossary. Thus when conducting the machine learning process, the gamma and complexity characteristics were tested at various powers of 10 starting at 0.001 and moving up by magnitudes of 100x (ie. 0.001, 0.1, 10) and their results stored in a *crossvalidation.csv* file. Note: the gamma parameter is only relevant to the Gaussian kernel and the exponent parameter is only relevant to the Polykernel. The exponent parameter was tested at 3 levels, linear, quadratic and cubic with higher exponent levels expected to over-fit the data or decelerate the speed at which cross validation tests were conducted.

Cross validation is the process of splitting the data into x equal portions, using 1 of those portions for testing, and the remaining x-1 for training. This process is repeated x times with each portion used individually for testing and the average result written into the *crossvalidation.csv* file. The format of this file is in comma separated values for the purpose of reading it in other programs (Matlab) and being able to manipulate it as easily as possible. Figure 3.4 highlights how results are displayed. A five fold cross validation method was used.

Training File Name	Kernel	Complexity Parameter	Gamma (RBF)	Exponent (Poly)	Classification Accuracy	Number of Instances in File
Alex S Yellow V Colours	PolyKernel	0.001	0	1	51.56739812	638
Alex S Yellow V Colours	PolyKernel	0.1	0	1	64.4200627	638
Alex S Yellow V Colours	PolyKernel	10	0	1	71.63009404	638
Alex S Yellow V Colours	PolyKernel	0.001	0	2	51.56739812	638
Alex S Yellow V Colours	PolyKernel	0.1	0	2	69.27899687	638
Alex S Yellow V Colours	PolyKernel	10	0	2	81.66144201	638
Alex S Yellow V Colours	PolyKernel	0.001	0	3	62.69592476	638
Alex S Yellow V Colours	PolyKernel	0.1	0	3	77.27272727	638
Alex S Yellow V Colours	PolyKernel	10	0	3	88.55799373	638
Alex S Yellow V Colours	RBFKernel	0.001	0.001	0	51.56739812	638
Alex S Yellow V Colours	RBFKernel	0.1	0.001	0	51.56739812	638
Alex S Yellow V Colours	RBFKernel	10	0.001	0	51.56739812	638
Alex S Yellow V Colours	RBFKernel	0.001	0.1	0	51.56739812	638
Alex S Yellow V Colours	RBFKernel	0.1	0.1	0	51.56739812	638
Alex S Yellow V Colours	RBFKernel	10	0.1	0	71.47335423	638
Alex S Yellow V Colours	RBFKernel	0.001	10	0	51.56739812	638
Alex S Yellow V Colours	RBFKernel	0.1	10	0	53.44827586	638
Alex S Yellow V Colours	RBFKernel	10	10	0	91.53605016	638
Alex S Yellow V Colours	RBFKernel	0.001	1000	0	51.56739812	638
Alex S Yellow V Colours	RBFKernel	0.1	1000	0	51.56739812	638
Alex S Yellow V Colours	RBFKernel	10	1000	0	51.56739812	638

TABLE 3.4: Extract of Cross Validation Results (All Sensors)

3.4.5 Matlab Post Processing

After the cross validation tests have run, the data is stored in structures to observe the data trends in the EEG results. Initially, all the data from the *crossvalidation.csv* file is stored in structs for each subject. That is, 39 structs are created with properties of each participant. A substruct is created within each subject struct to store data fields associated with each training test file the participants took part in. These fields include the maximum accuracy that could be achieved for each training set per subject, and the parameters that generated the best results. The results were graphed and trends were found then analysed using statistical analysis.

3.5 Statistical Analysis

3.5.1 Test Subjects

To analyse test subject data the following variables were found for each of the data sets:

1. Maximum Accuracy achieved over all tests
2. Mean of all cross validation data by training test
3. Mean of all subject data by training test
4. Median of all subject data by training test
5. Interquartile range of all subject data by training test

3.5.2 Machine Learning Parameters

Machine learning parameters were analysed by sorting data into the types of kernels used, the complexity factor, gamma parameter (RBF Kernel only) and exponential (PolyKernel only). The mean accuracy was found for each combination of parameters used to cross validate the training tests over the total 39 participants. The variance of these cross validation results were found and utilised to hypothesise the probabilities that these parameters were able to generate their associated mean results. The hypothesis was tested by using the t-statistic distribution method whereby the t statistic for each parameter combination was found using the following equation:

$$t_{n-1} = \frac{X-\mu}{\frac{\sigma}{\sqrt[2]{n}}} \tag{3.1}$$

n is the number of participants, μ is mean, σ is the standard deviation and t_{n-1} is the t-statistic. Using the t-statistic, I could test whether the machine learning parameters had a mean that satisfied a suggested hypothesis. If the t-statistic was greater than the value 1.684 representing 38 degrees of freedom at a 95% confidence interval ($t_{38,0.05}$)), then I could be 95% confident that the outcomes had a higher mean than the tested mean hypotheses. The tested mean hypotheses were ($H_0 : \bar{x} > 75\%$, $H_0 : \bar{x} > 80\%$, $H_0 : \bar{x} > 85\%$. The results for each parameter combination were recorded and likelihood for each mean hypothesis was defined as the number of parameter combinations that would ensure the tested mean hypothesis was true; i.e. if all the parameter combinations for the *Yellow Vs Red* test showed that their mean was greater than 75% then the data bucket "Likelihood > 75%" would indicate 100%. Results of this hypothesis testing were tabulated and plotted. If the hypothesis was not rejected, the cell was formatted green. Additionally, the number of tests that were not rejected were counted, see Appendix B.

4 Results

All the results discussed in this section are based on the output of the machine learning algorithm produced by the 5-fold cross validation test across each of the training tests. The results show that males and females do not indicate very significant disparities between their results whilst adolescents (Age $\leq$ 25) and adults (Age $>$ 25) delineate significant differences including improved performance measures in adolescents. The *Yellow Vs Vd* test indicated least favourable results in generating an accurate binary output. The RBF Kernel though, with a parameter combination 10-10 (complexity-gamma), indicated that it was able to produce consistently produce accurate results, highlighting a success rate in excess of 90% for each training test within each of the 3 data sets used ("All Sensors", "Frontal Lobe", "Temporal Lobes"). The results of the longitudinal test depict the capabilities of the SMO algorithm through its' high rate cross validation results, exceeding accuracies of 85% across all tests, given subjects were exposed to a multitude of uncontrolled variables.

4.0.1 Method of Analysis

The raw data collected from each training exercise is analysed and presented in this chapter within a frame work that tests the three components of the BCI design: the participants' characteristics, the muse headset technology and the machine learning algorithm. The participants' characteristics are split broadly into age and gender with differences in their participation environment, background and behaviour incorporated into the overarching analysis of the BCI results. These include raw brain signal analysis, cross validation results and participants' individual responses to each of the training tests. Further, the performance of machine learning parameters is investigated through the conduction of various cross validation tests and their post-statistical study. The role of this study was to find which set of parameters generated the best results for either a specific class of test subjects or for the general population. Since the Muse headset consists of frontal electrodes (FP7 and FP8) and temporal electrodes (TP9 and TP10), acquiring different signals based on the visual and aural stimuli, their data should ideally present interesting artefacts when analysed individually. Thus the results of the machine learning parameter and subject characteristics analysis is presented in this chapter using 3 sets of data (**All sensors, Frontal lobes only and Temporal lobe only**).

4.1 All Sensors

4.1.1 Age

Participant Feedback by Age

Feedback from older participants indicated issues surrounding the complexity of tests, whilst younger participants' issues were more emotional and abstract. Older participants found it difficult initially to focus in the *Yellow Vs Vd* exercise as it was a significant step up in difficulty. Younger participants found all the tests quite comfortable and only found the *Yellow Vs Red with Audio* exercise difficult as they enjoyed the music and were more distracted particularly if they recognised the music. They also seemed to have a shorter attention spans, and were more impatient with 10 of the 25 adolescent subjects showing disinterest after 30 seconds and desiring to fidget or engage in a different activity.

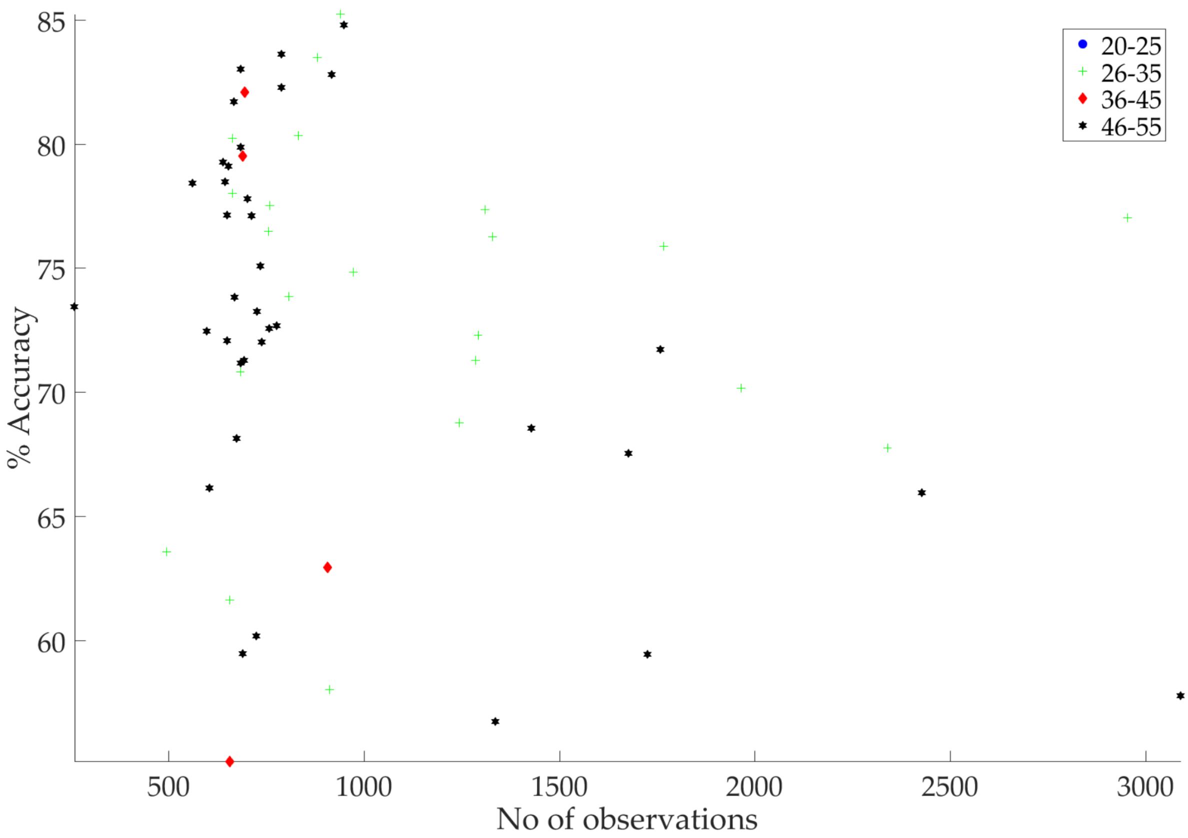


FIGURE 4.1: **All Sensors:** Combined Cross Validation Results - sorted by Age

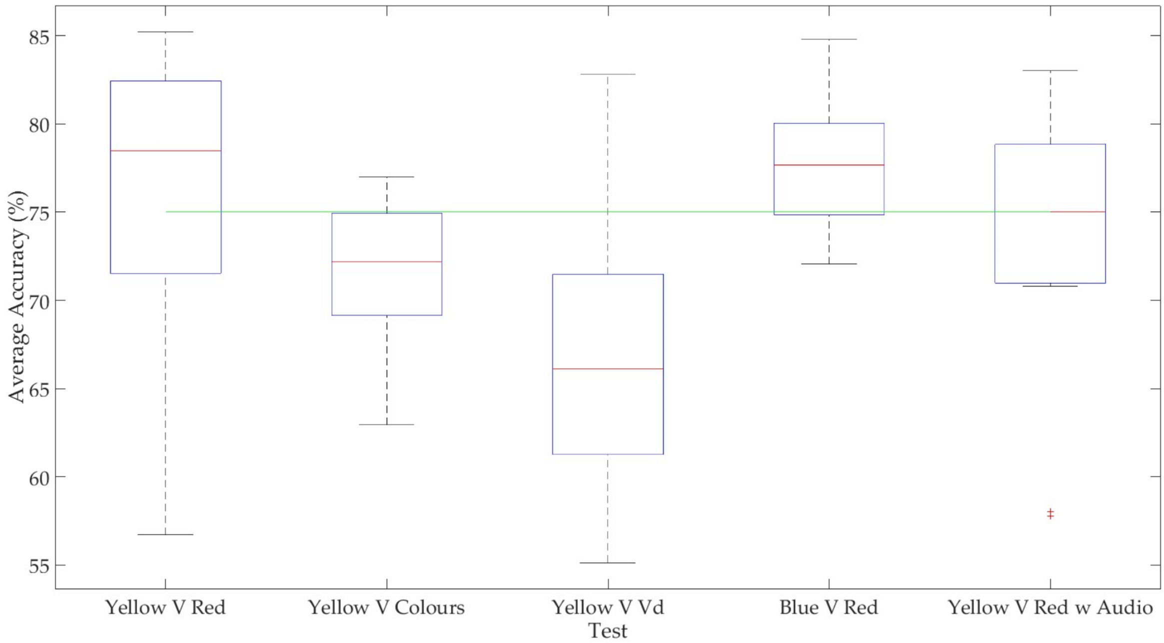


FIGURE 4.2: **All Sensors:** Average Cross Validation Results for Subjects over 25 years of Age

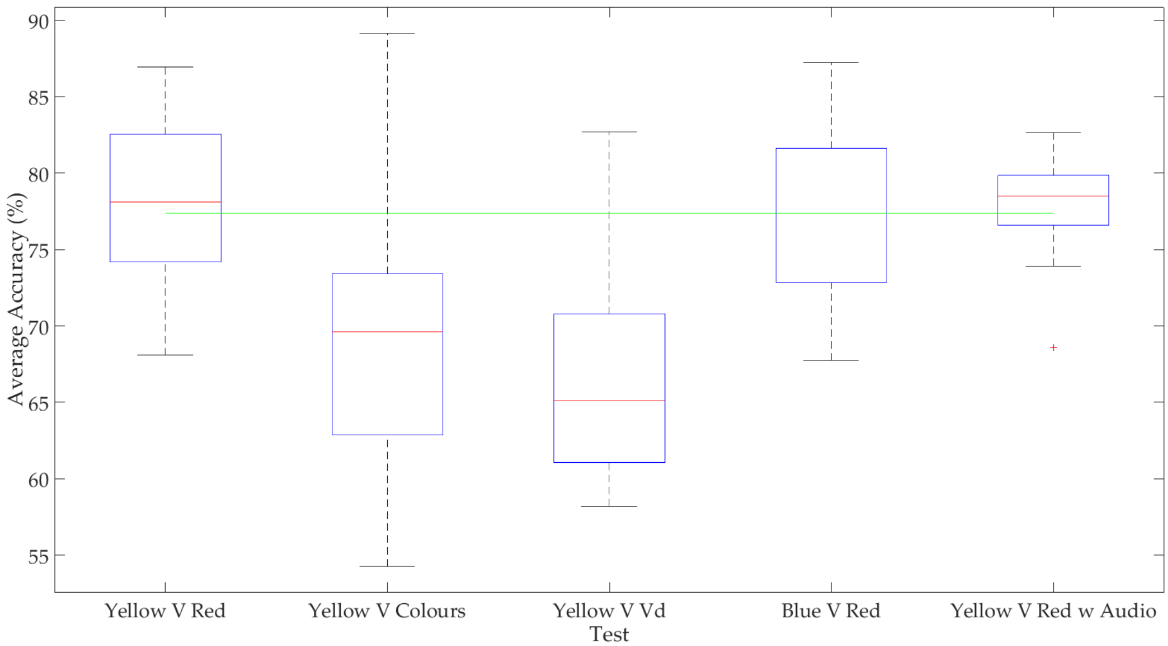


FIGURE 4.3: **All Sensors:** Average Cross Validation Results for Subjects under 25 years of Age

Overall Participant Test Trends

Figures 4.1 demonstrates the average[1] results of each cross validation test with Figure 4.2 and 4.3 showcasing that commonly poorer results were achieved with *Yellow Vs Vd* exercise. Adults performed more consistently in the *Blue Vs Red* test while adolescents performed more consistently in the *Yellow Vs Red w Audio* exercise.

It is made abundantly clear upon visual inspection in Figure 4.1 that the variance of results across ages is of high order ranging from 54% to 94%. At first glance it is almost impossible to acknowledge the existence of any trends, but Figures 4.2 and 4.3 clarify how these accuracies are clustered within each test. Figure 4.2 demonstrates that poorer results are achieved using the *Yellow Vs Vd* test where the quartiles range from 63% to 72% and the median is 10% lower than the total median of results (75%) for participants over the age of 25. This is a unique result as the *Yellow Vs Vd* exercise was expected to generate superior results as the brain produces an increased volume of beta and gamma waves when subjects are engaging in more complex tasks that require concentration (Li, Xu, and Zhu, 2015). The differentiation between this increase would have justified an increased accuracy for each subject. However this was not the observed case as the highest achieved accuracy was achieved by the *Yellow Vs Red* exercise. It should be noted that the *Blue Vs Red* exercise was found to have equally strong results, but showed more consistency between subjects. By initial observation, it can be argued that this test therefore is a more reliable test for producing a binary response from a participant but forgoes the adjusted characteristics of participants from the time they started the five round of tests to the time they conducted this test. They may have been more tired when conducting this exercise, following the *Yellow Vs Vd* exercise, or may have felt more comfortable doing an easier exercise, a coherent rationale given their feedback about the more complex tests. The evidence reiterates this with both tests succeeding the *Yellow Vs Vd* generating median results that are significantly better and the *Yellow Vs Red w Audio* exercise exhibiting slightly weaker outcomes than the *Blue Vs Red*.

Figure 4.3 demonstrates that results between adolescents and adults are alike in nature yet adolescents are able to achieve better performing results than adults. The *Yellow Vs Red* training exercise shows higher average accuracy than most of its counterpart tests but interestingly, the spread of this test among adults was quite high whilst the *Yellow Vs Colours* exercise generated the largest variance of results in adolescents. Accumulating better results from the *Yellow Vs Red w Audio* exercise using younger participants suggests that younger participants are able to control a binary output more efficiently in the presence of aural distractions. It is possible that this is a consequence of younger generations becoming more accustomed to drowning out background noise to focus on social media/mobile applications.

[1] *The average* defined as the average result of all parameters used for each SVM cross validation

Or it simple represents the byproduct of a still maturing, temporal lobe. This is contrasted to elder subjects whose accurate and reliable results from the *Blue Vs Red* and *Yellow Vs Vd* exercise convey that binary feedback is better achieved through the use of blue as a stimulus rather than yellow, but may, still, potentially exhibit poorer outcomes in a noisier environment.

4.1.2 Gender

Participant Feedback by Gender

There were large disparities in the practical testing between male and female mainly that can be attributed to differences in hair interference, head size, skin product usage and confidence. Whilst the electrodes did a good job in detecting signals overall, they struggled to acquire signals for females that had lots of hair over their ears, where the electrodes to capture temporal lobes would sit. The thickness of hair was directly correlated to the difficulty in acquiring the signal as certain female subjects with thin hair were able to generate strong signals overall. It was also found that younger girls who had applied make-up or had used skin cream, generated worse results than many male counter parts and in turn did not provide a large number of observations through tests. Li, Xu, and Zhu (2015) used the Muse headset in their experiments and highlighted that alcohol wipes were utilised on foreheads before headset to avoid this, and while this may have eliminated issues, it would not have encompassed the spirit of testing under everyday conditions. Further, electrodes TP9 and TP10 were specifically most prone to this issue whilst FP7 and FP8 coped with this issue seamlessly deducing that the conductive rubber electrodes were most likely the source of complications.

On the other hand, males did not face these obstacles when wearing the headset and possibly because of their confidence with electronic items (less time was spent explaining how to use the headset) and stronger signals were observed when any subject was generally more confident about using the headset. Subjects' head sizes alternatively were a primary concern particularly with female subjects as the headset's band was not tight enough to maintain sufficient contact with the frontal electrodes. This was usually solved by placing the muse headset at an acute angle on the forehead rather than at 90°, but reduced the number of observations for the participant. Additionally, subjects with a more oval-shaped head (mainly females) found difficulty in upholding full contact - particularly with electrodes FP7 and FP8. It seemed the Muse Developers had designed the headset to fit a particular perfectly circular shaped head that is not shared by the entire general population and thus questions its reliability as a consumer headset for everyday use.

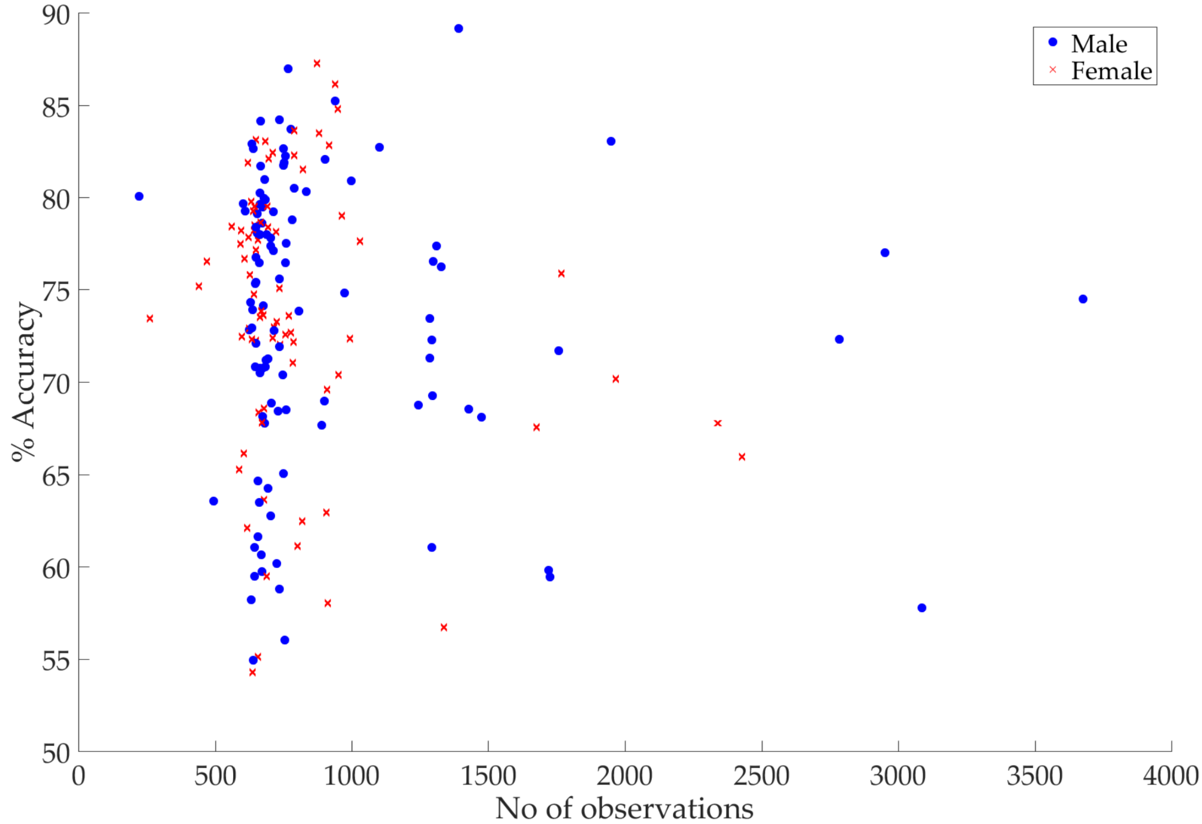


FIGURE 4.4: **All Sensors:** Combined Cross Validation Results - sorted by Gender

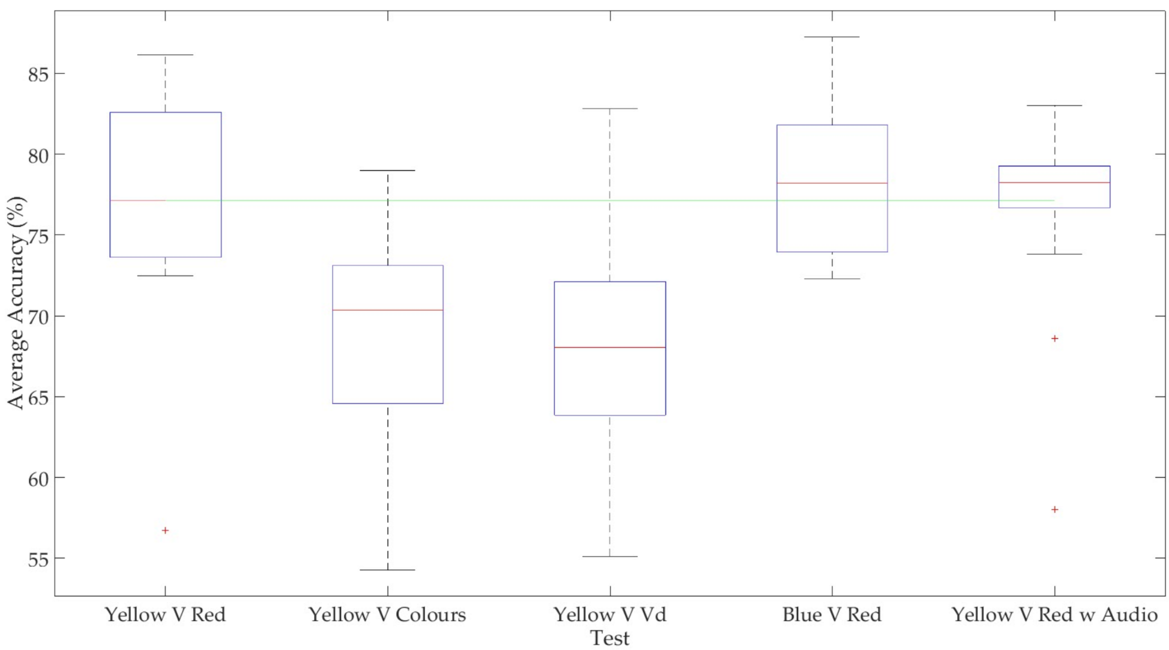


FIGURE 4.5: **All Sensors:** Average Cross Validation Results for Female Subjects

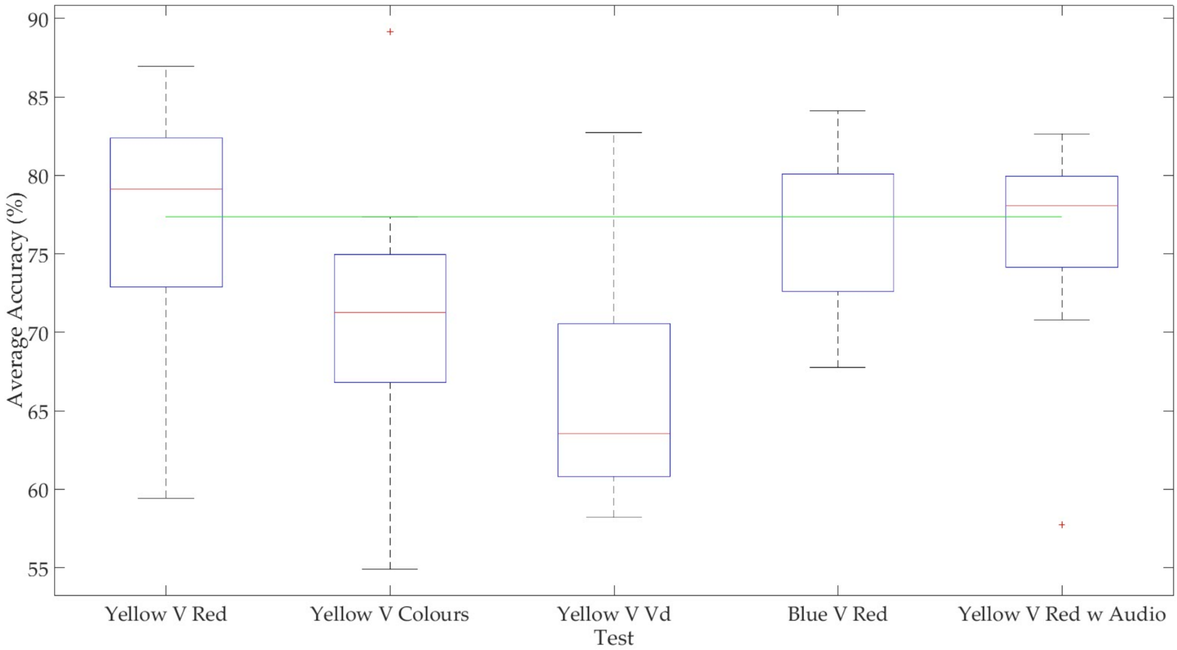


FIGURE 4.6: **All Sensors:** Average Cross Validation Results for Male Subjects

Overall Participant Test Trends

Figure 4.5 and 4.6 indicate surprisingly little differences in their trends with results shown in Figure 4.4 congruent among the results of male and female. That is, no bias towards specific variance or mean/median to generate the results highlighted in section 4.1.1. Comparing the male and female results in Figure 4.4 present the repudiation of the idea that more data will induce more accurate results. While some males were able to keep full contact with the headset for longer periods of time they were not able to produce more accurate results with the machine learning algorithm, specifically Figure 4.7 showcasing the results of a male in this test who produced the worst results of the class despite having 3x the number of instances recorded for the test. It was believed that the weak contact females have with the Muse headset would lead to results that are less accurate and reliable, but these figures indicate otherwise.

An interesting proposition about society's perception of the human brain was conjured through the raw data analysis[2] of male and female subjects', particularly those who participated in the same environment. This proposition is described using notable two observations within the context already discussed:

1. Males' brain waves exhibit more disparate amplitudes and frequencies than females amongst themselves

2. Despite these significant disparities the results between males and females remain widely similar.

[2] amplitude and frequency of brain waves

Figures 4.8 and 4.9[3] are 2 randomly chosen subjects[4] (Male and Female) used to present the idea that the human brain, despite its complexities and idiosyncrasies, has many features that are constant between males and females. This is vital to the community's understanding of machine learning algorithms as many machine learning experiments that are conducted exclude the use female subjects claiming to eliminate distortions from an additional variable within their studies.

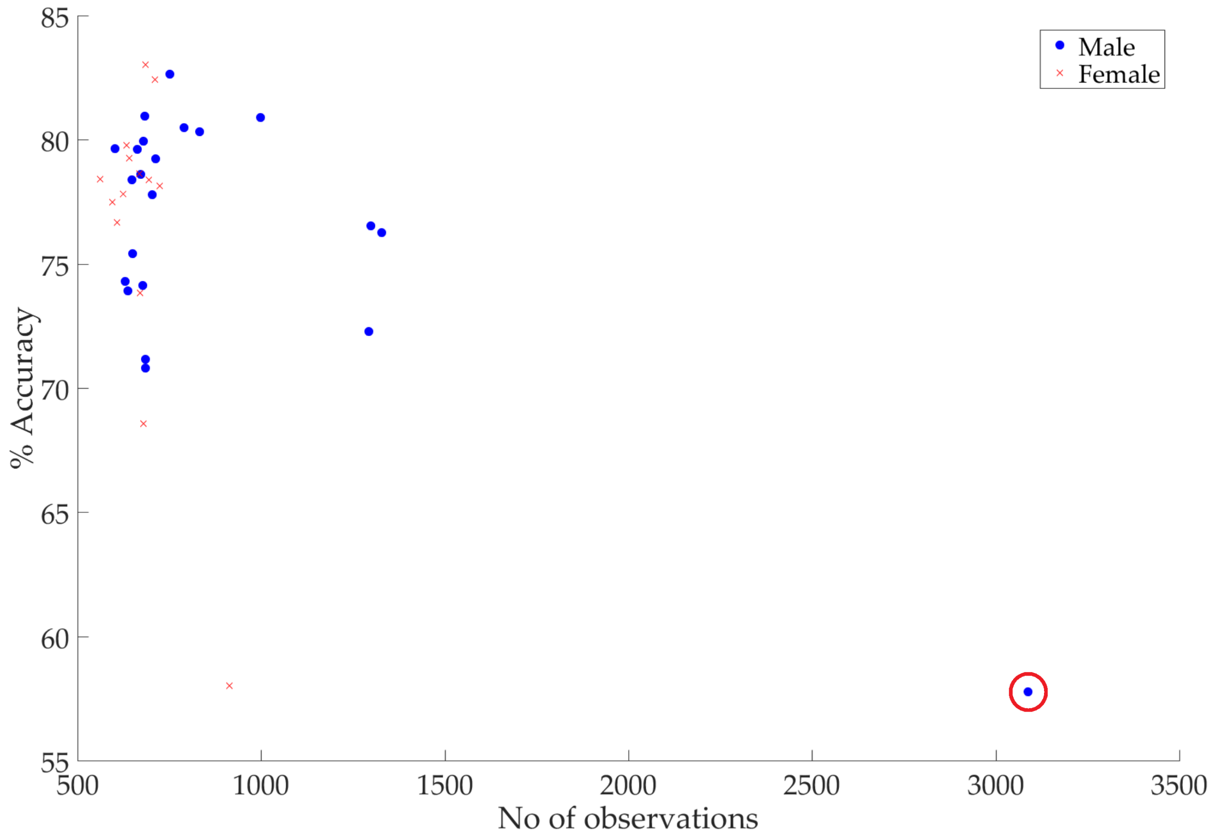


FIGURE 4.7: **All Sensors:** Yellow versus Red with Audio Cross Validation Results with observations-to-accuracy outlier

[3]The **relaxed** state is when a participant does not recognise yellow. The **concentrating** state is when the participant recognises yellow.

[4]refer to Appendix A

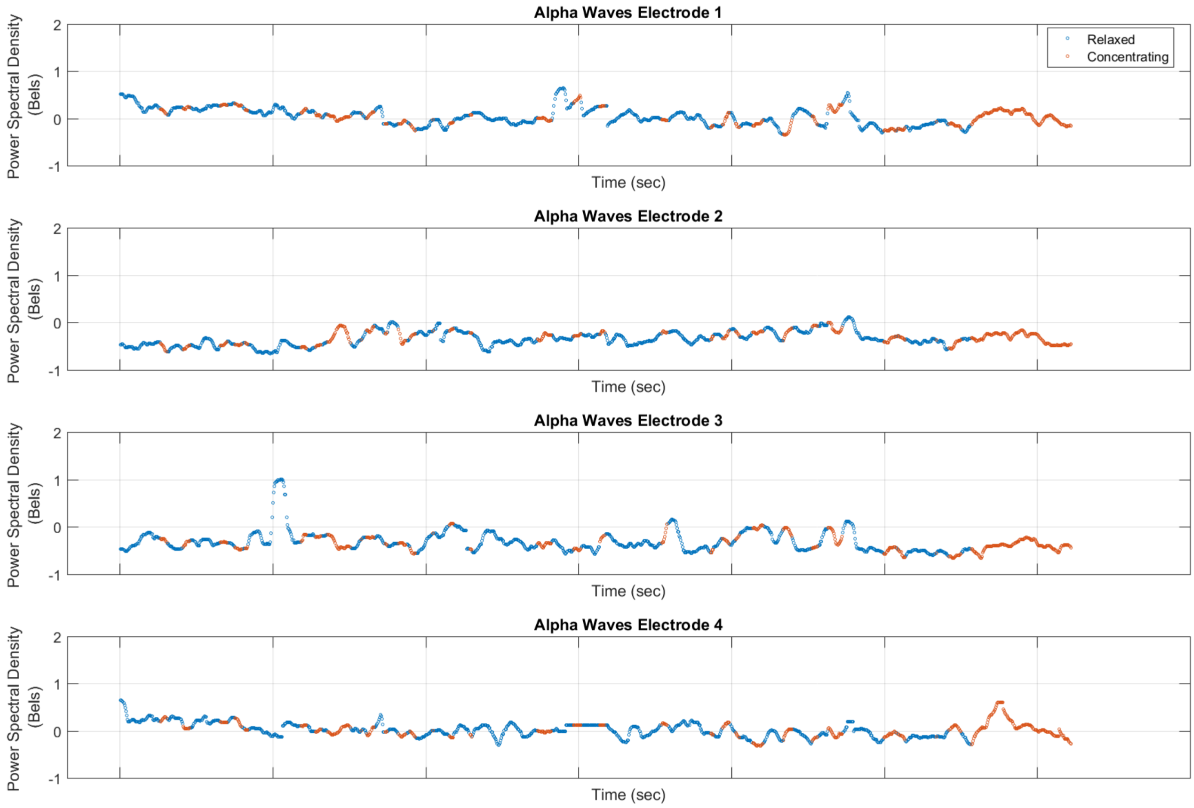


FIGURE 4.8: **All Sensors:** Yellow versus Red Male Subject No.3 Alpha Waves

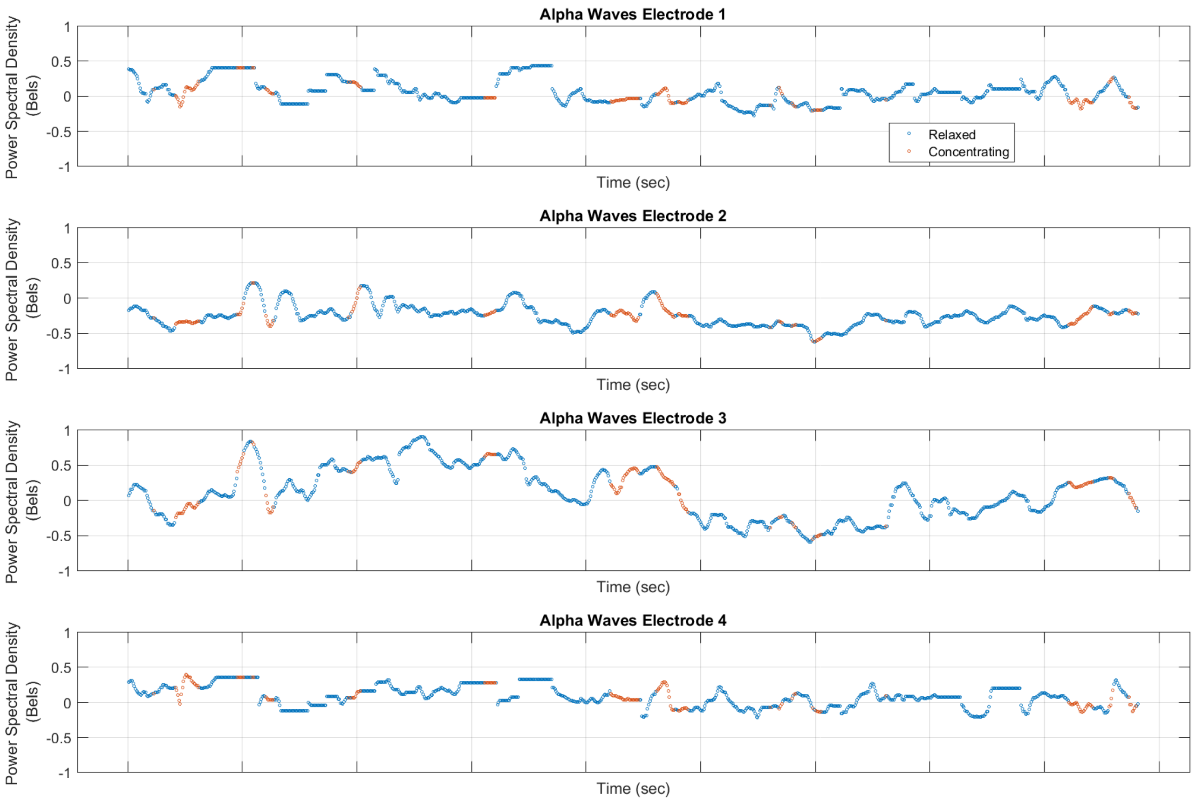


FIGURE 4.9:

[5]**All Sensors:** Yellow versus Red Female No.2 Subject Alpha Waves

4.1.3 Machine Learning Parameters

The machine learning parameter analysis demonstrated that the *Blue Vs Red* exercise was the most reliable, the *Yellow Vs Red* exercise was the most accurate, and the RBF Kernel with a complexity parameter of 10 and gamma parameter of 10 guaranteed the best results on average over the entire class of subjects (94.70%). The results can be viewed in Table 4.1 and Figure 4.10 showcasing the top 2 best performing parameters overall and the worst parameters. In each of the tests, the PolyKernel parameter manages to outperform the RBF kernel overall particularly in the *Yellow Vs Visual Distractions* as the RBF Kernel averages 69.10% whilst the PolyKernel averages 80.67%, but the consistency of the highest result of all parameters (including PolyKernel and the RBF Kernel) across each of the tests with the RBF kernel parameters (10,10) are exceptional. Over the general population of subjects, it can be seen in Table 4.2, that almost all the subjects achieve the highest accuracy when cross validated with the RBF kernel and the 10-10 complexity-gamma parameters. The statistical analysis on each of the training tests showed that the likelihood of achieving more than 80% using all the cross validation parameters was 100% for the *Yellow Vs Red* exercise at a 95% confidence interval. This meant that using any combination of parameters, if the machine was trained using the *Yellow Vs Red* training exercise, you would be 95% confident that your machine learning algorithm is functioning above an 80% success rate. Alternatively, the *Yellow Vs Vd* exercise produced the highest number of cross validation outcomes that were likely to achieve more than 85%, yet it achieved the worst overall average results for participants as outlined in 4.1.1.

Complexity Factor	Gamma Parameter	Exponential	Accuracy
10	0	3	92.90%
10	10	0	94.70%
0.001	0.001	0	70.45%

TABLE 4.1: **All Sensors:** Best and Worst Performing Cross Validation Parameters

[5] Relaxed indicates subjects are not thinking about the colour yellow. Concentrating indicates subjects are thinking about the colour yellow.

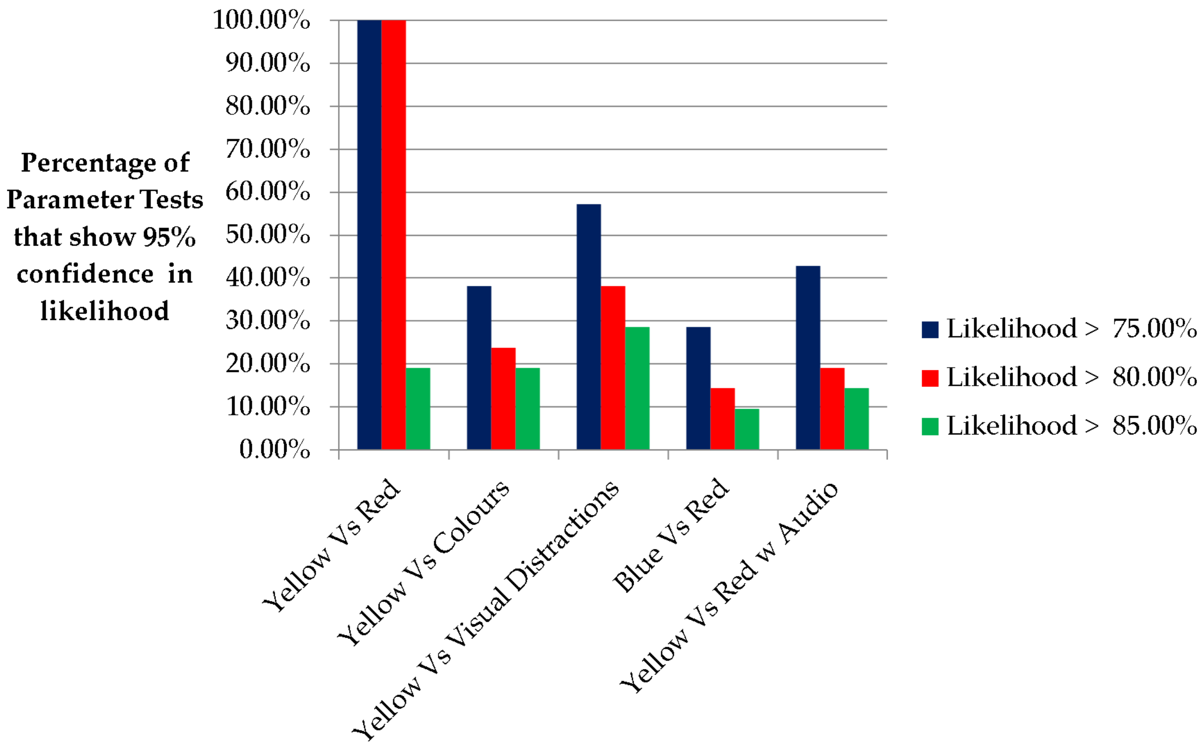


FIGURE 4.10: **All Sensors:** Likelihood of Cross Validation Results for each Test

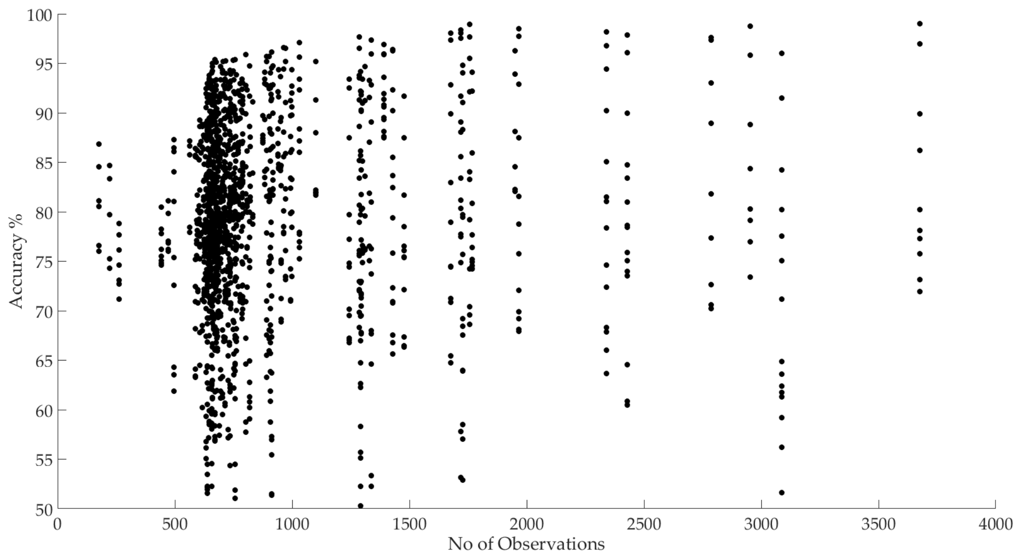


FIGURE 4.11: All Cross Validation Results

4.1.4 The relationship between instances and results

Though this question has already been answered highlighting how some males recorded a large number of observations but did not produce higher overall average results, there are some peculiar points to note. When observing Figure 4.11 I expect there to be a large spread of data as machine learning parameters will overfit or underfit the data points provided to it. I however, do expect there to be an increasing accuracy-observation trend as the general rule of thumb for machine learning is: the more instances the machine experiences the better trained it will become. However from this data it may be inferred that as the number instances increases, the additional instances required to generate improved performance rises exponentially. Inspecting the general trend of most accurate results best showcases this because individuals' best performance rises as the number of observations rise. This rise however occurs more-so from the 500-1500 observation interval and then slows until it minutely rises towards the subject with 3500 instances. These individuals additionally don't exhibit disparate characteristics from their counterparts and are a mix of adults/adolescents, male/female, and even nervous/confident subjects. Therefore the data resounds the notion that more observations will indeed produce more accurate results for machine learning but to continue increasing these results demands exponentially increasing time from the user.

Subject No	Gender	Age	Best Kernel	Complexity	Gamma	Epsilon	Max Accuracy (%)
1	'M'	50	RBFKernel	10	10	0	98.91922639
2	'F'	26	RBFKernel	10	10	0	98.5249237
3	'M'	27	RBFKernel	10	10	0	98.78090078
4	'M'	22	RBFKernel	10	10	0	98.9937449
5	'M'	20	RBFKernel	10	10	0	97.63016158
6	'M'	22	RBFKernel	10	10	0	95.19056261
7	'M'	21	RBFKernel	10	10	0	95.73033708
8	'M'	23	RBFKernel	10	10	0	93.81761978
9	'F'	21	RBFKernel	10	10	0	94.6969697
10	'F'	23	RBFKernel	10	10	0	93.02030457
11	'F'	18	RBFKernel	10	10	0	88.7987013
12	'F'	18	RBFKernel	10	10	0	89.70251716
13	'F'	19	RBFKernel	10	10	0	96.57676349
14	'F'	21	RBFKernel	10	10	0	95.88528678
15	'M'	19	RBFKernel	10	10	0	95.38690476
16	'F'	21	RBFKernel	10	10	0	94.74969475
17	'M'	24	RBFKernel	10	10	0	92.48826291
18	'M'	22	RBFKernel	10	10	0	91.77018634
19	'M'	23	RBFKernel	10	10	0	93.22289157
20	'M'	23	PolyKernel	10	0	3	94.31818182
21	'F'	22	RBFKernel	10	10	0	92.75521405
22	'M'	29	RBFKernel	10	10	0	96.50565262
23	'M'	31	RBFKernel	10	10	0	95.37648613
24	'F'	38	RBFKernel	10	10	0	94.0463065
25	'M'	55	RBFKernel	10	10	0	94.6685879
26	'M'	51	RBFKernel	10	10	0	95.11111111
27	'M'	22	RBFKernel	10	10	0	95.3271028
28	'M'	22	RBFKernel	10	10	0	92.40121581
29	'F'	50	RBFKernel	10	10	0	98.03220036
30	'M'	22	RBFKernel	10	10	0	92.48120301
31	'M'	22	RBFKernel	10	10	0	92.89026275
32	'M'	22	RBFKernel	10	10	0	96.90869878
33	'F'	22	RBFKernel	10	10	0	97.08737864
34	'F'	20	RBFKernel	10	10	0	91.23867069
35	'F'	50	RBFKernel	10	10	0	94.77124183
36	'M'	26	RBFKernel	10	10	0	93.30323552
37	'F'	52	RBFKernel	10	10	0	93.91069012
39	'F'	52	PolyKernel	10	0	3	92.93059126
40	'F'	52	PolyKernel	10	0	3	89.27392739

TABLE 4.2: **All Sensors:** Subjects' Best Cross Validation Results

4.2 Frontal Lobes Only

4.2.1 Age

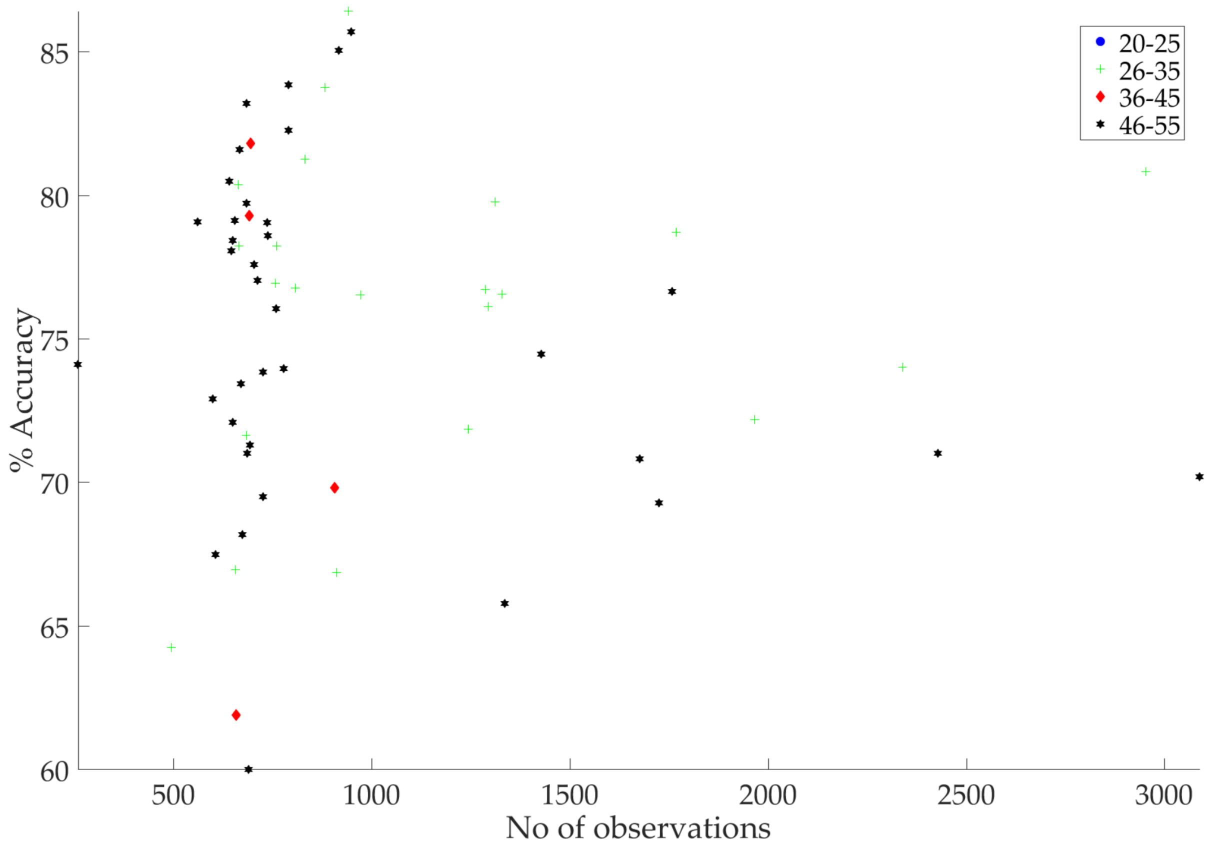


FIGURE 4.12: **Frontal Lobe Only:** Combined Cross Validation Results - sorted by Age

The results of the frontal lobe independent data analysis for adults show improved results (higher median and reduced variance) in all the visual tests and a weakened set of average results in the *Yellow Vs Red w Audio* exercise compared to 'All Sensors', while the adolescents' results were more or less similar. The adults' tests that exhibited the highest improvement were the *Yellow Vs Vd* and *Yellow Vs Red* exercises as the subjects whose data attributed to poor results around the 55% mark in both tests would achieve at least 65% in the *Yellow Vs Red* exercise and 60% in the *Yellow Vs Vd* exercise. The bounds in *Blue Vs Red* tightened producing results within 75% and 83%, which indicates mixed results but further inspection in Figure 4.16 below portrays males as the subjects who attribute poorer results for the originally better performing subjects in section 4.1. Conversely, the adolescents' results with independent frontal lobe data revealed no significant change in results, implying that the usage of only two electrodes could have been used to generate similar results.

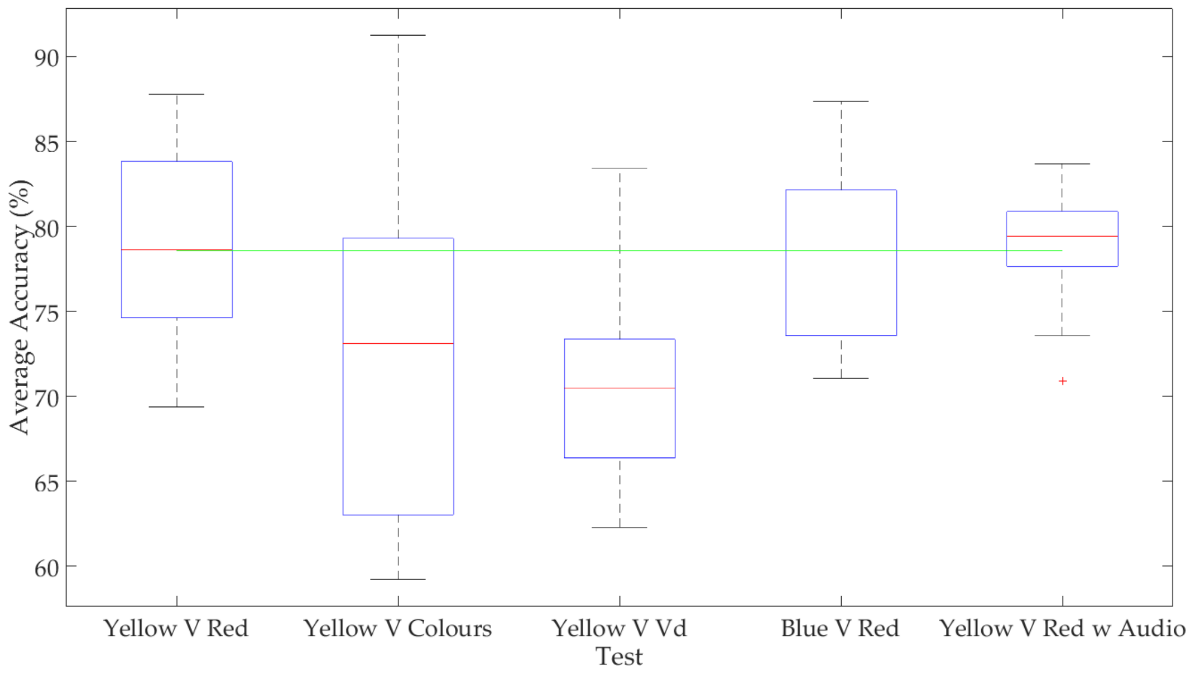


FIGURE 4.13: **Frontal Lobe:** Average Cross Validation Results for Subjects under 25 years of Age

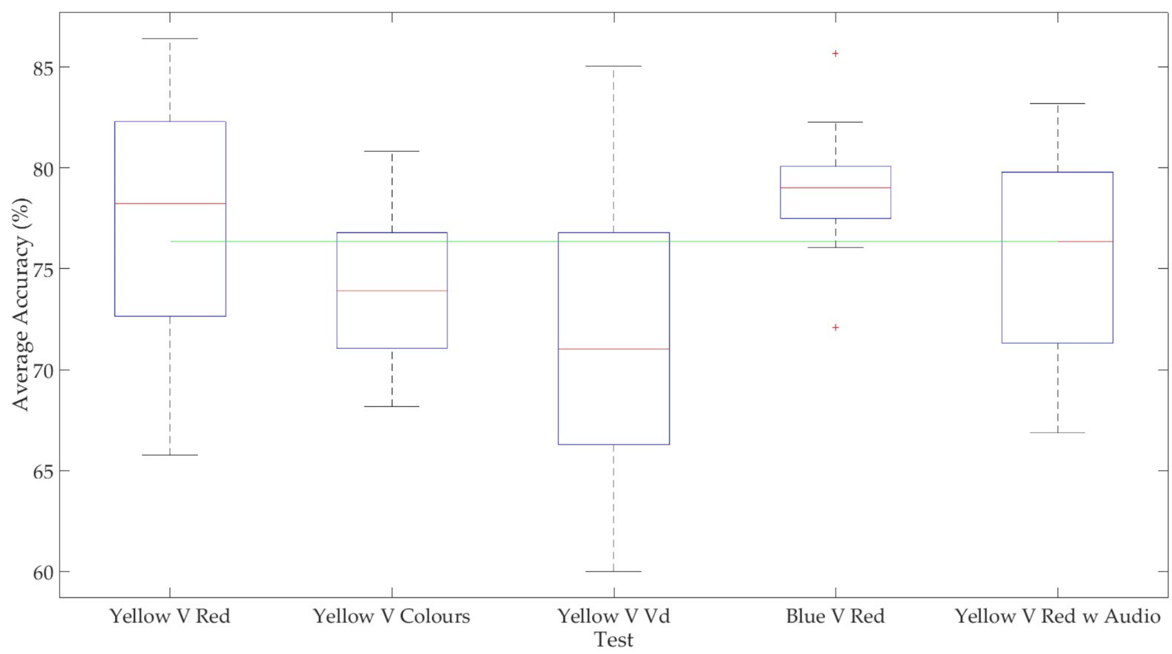


FIGURE 4.14: **Frontal Lobe:** Average Cross Validation Results for Subjects over 25 years of Age

4.2.2 Gender

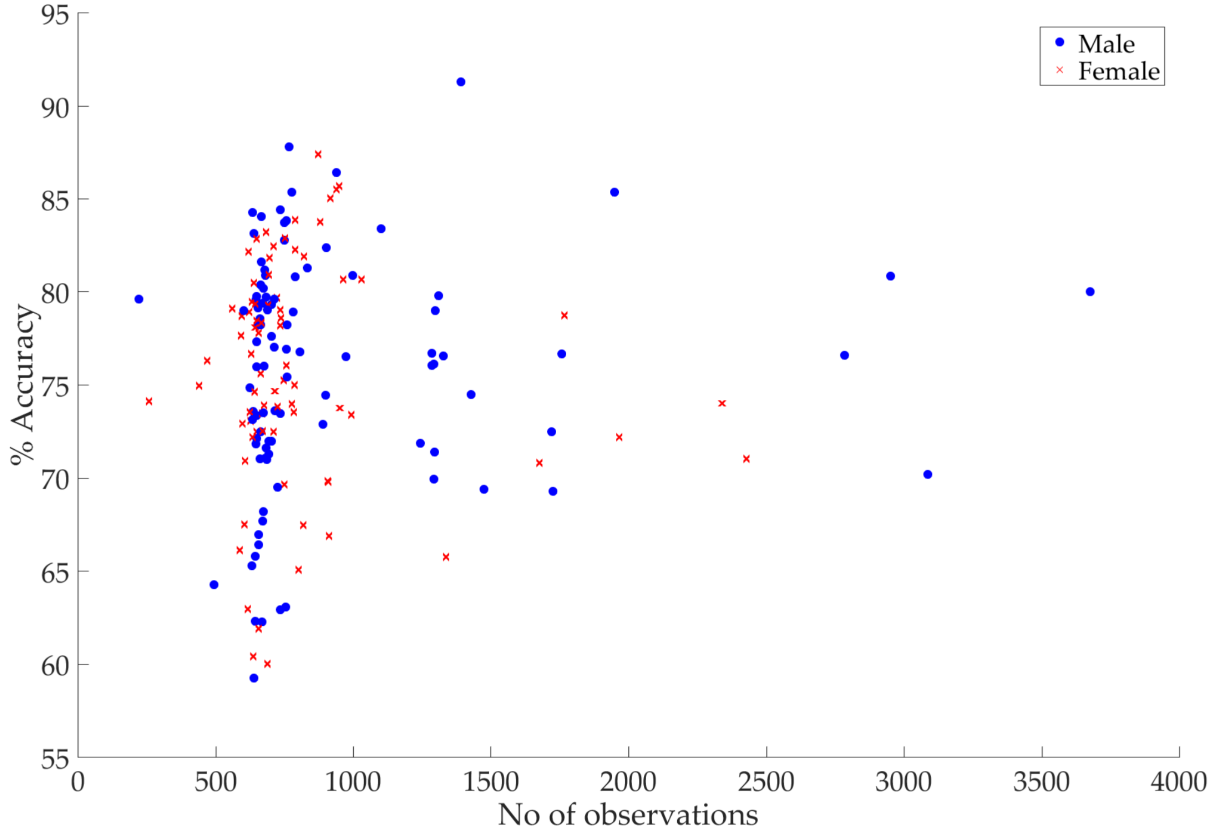


FIGURE 4.15: **Frontal Lobe:** Combined Cross Validation Results - sorted by Gender

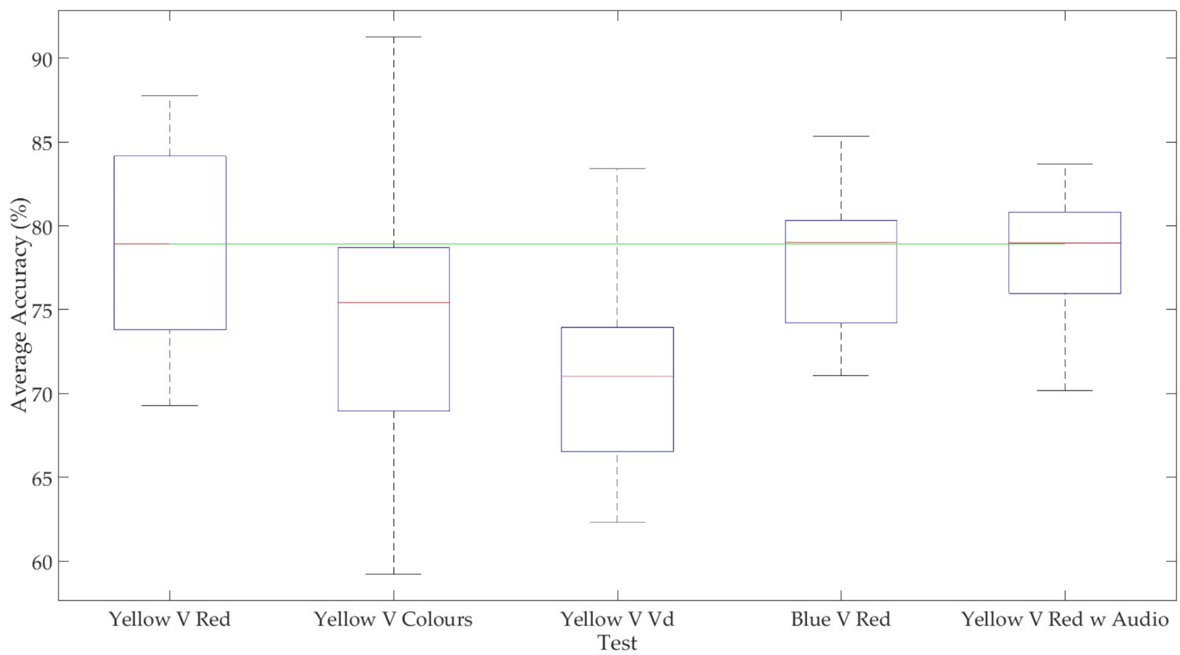


FIGURE 4.16: **Frontal Lobe:** Average Cross Validation Results for Male Subjects

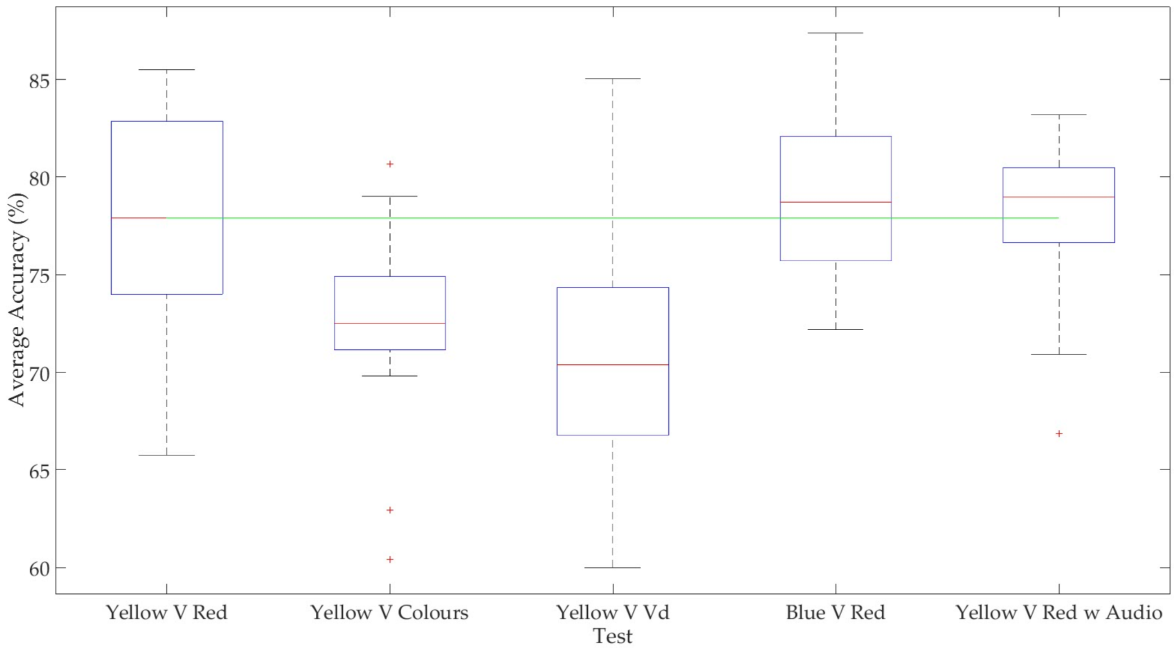


FIGURE 4.17: **Frontal Lobe:** Average Cross Validation Results for Female Subjects

Overall Participant Test Trends

Figures 4.16 and 4.17 fail to show any significant differences between their results however do demonstrate that the median result is improved when using frontal lobe data. This illustration of an absence in gender bias is therefore coherent with the theory that adults show improved performance when relying on the frontal lobe electrodes as adolescents' accuracy outcomes remain unchanged. Figure 4.15 reiterates overall improvement as generally the lower bound of scores demonstrated in 'All Sensors' cross validation results (55%) have increased to a lower bound of 60% accuracy when using frontal lobe data only.

4.2.3 Machine Learning Parameters

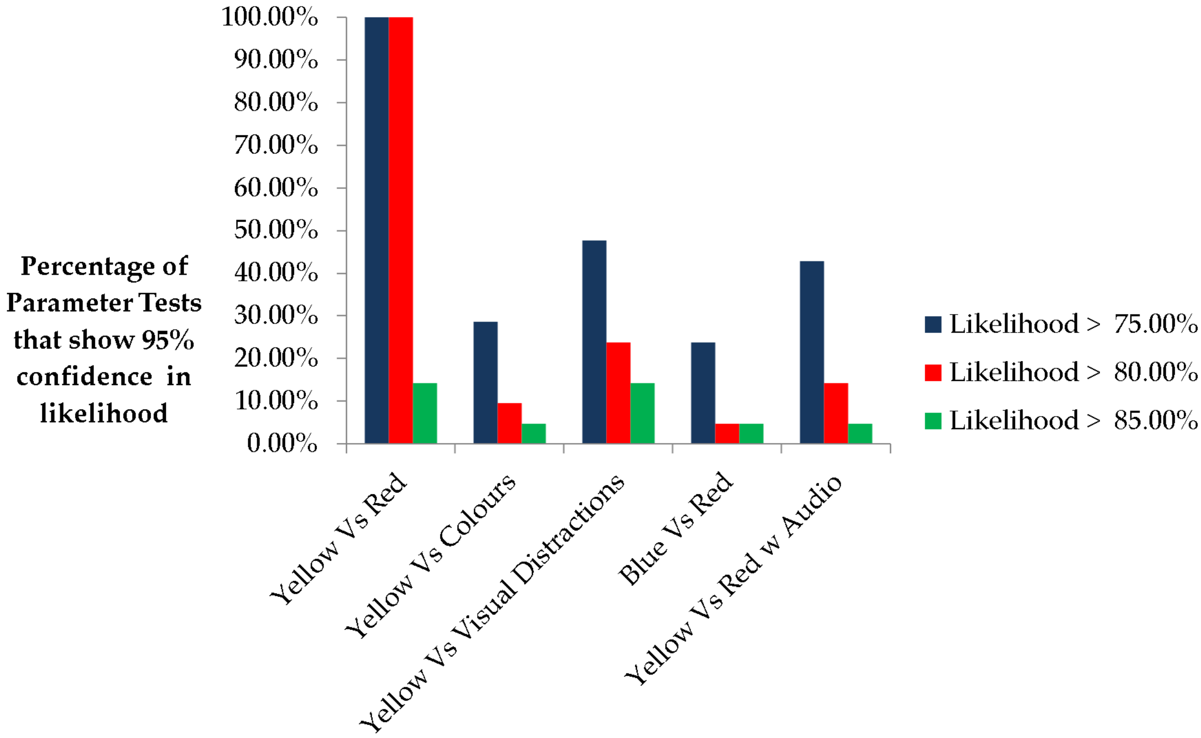


FIGURE 4.18: **Frontal Lobe:** Likelihood of Cross Validation Results for each Test

Complexity Factor	Gamma Parameter	Exponential	Accuracy
10	0	3	82.48%
10	10	0	93.60%
0.001	0.001	0	70.45%

FIGURE 4.19: **Frontal Lobe:** Best and Worst performing parameters

The results of the machine learning parameter analysis in Figures 4.18 and 4.19 indicate that the frontal lobes are the driver of the machine learning results displayed in the "All Sensors" data as the trends in likelihood and best/worst performing results are very similar, with machine learning performance only slightly weaker as a whole. Upon initial visual inspection, the likelihood results in Figure 4.18 show that the *Yellow Vs Red* exercise with or without audio showed no change in "Likelihood > 75%" and "Likelihood > 80%" remaining at their 100% values. Further the *Yellow Vs Vd* exercise resulted in extremely large falls, consistent of 10% across all the classes with the largest drop in the "Likelihood > 85%" bucket when using "All Sensor" data. It seems that if frontal lobe data is independently used, the likelihood of attaining better results is reduced but the improvement between each subject outcome is increased supported by the lifting of the lower bounds in Figure 4.15.

The best and worst parameter combinations remained identical to the "All Sensors" data with decreased average accuracies for the best results (93.60% and 82.46%) but with the same worst result of 70.45%. This insinuates that the use of frontal lobe data doesn't make the machine learning algorithm aggregately worse but prevents the algorithm from achieving higher accuracies. Moreover, the variance between cross validation outcomes was significantly different. The variance between "All Sensors" and the "Frontal Lobe" cross-validation outcomes was also quite distinguishable with Frontal Lobe data generating higher total variances between outcomes overall as a population sample than the "All Sensors data" as recognised in Figure 4.20. This figure implies that since there is more variation in the outcomes, the "All Sensors" data is more centered around the mean of the cross validation results whilst the Frontal Lobe data is more dispersed between the lower and upper bounds.

Sum of Variance of Results	All Sensors	Frontal Lobe	Temporal Lobe
Yellow Vs Red	45852	45936	48528
Yellow Vs Colours	33733	38444	36128
Yellow Vs Vd	56399	57522	42699
Blue Vs Red	27838	29036	29022
Yellow Vs Red w Audio	40211	44471	34840

FIGURE 4.20: Sum of Cross Validation Variances of each training exercise

4.3 Temporal Lobes Only

4.3.1 Age

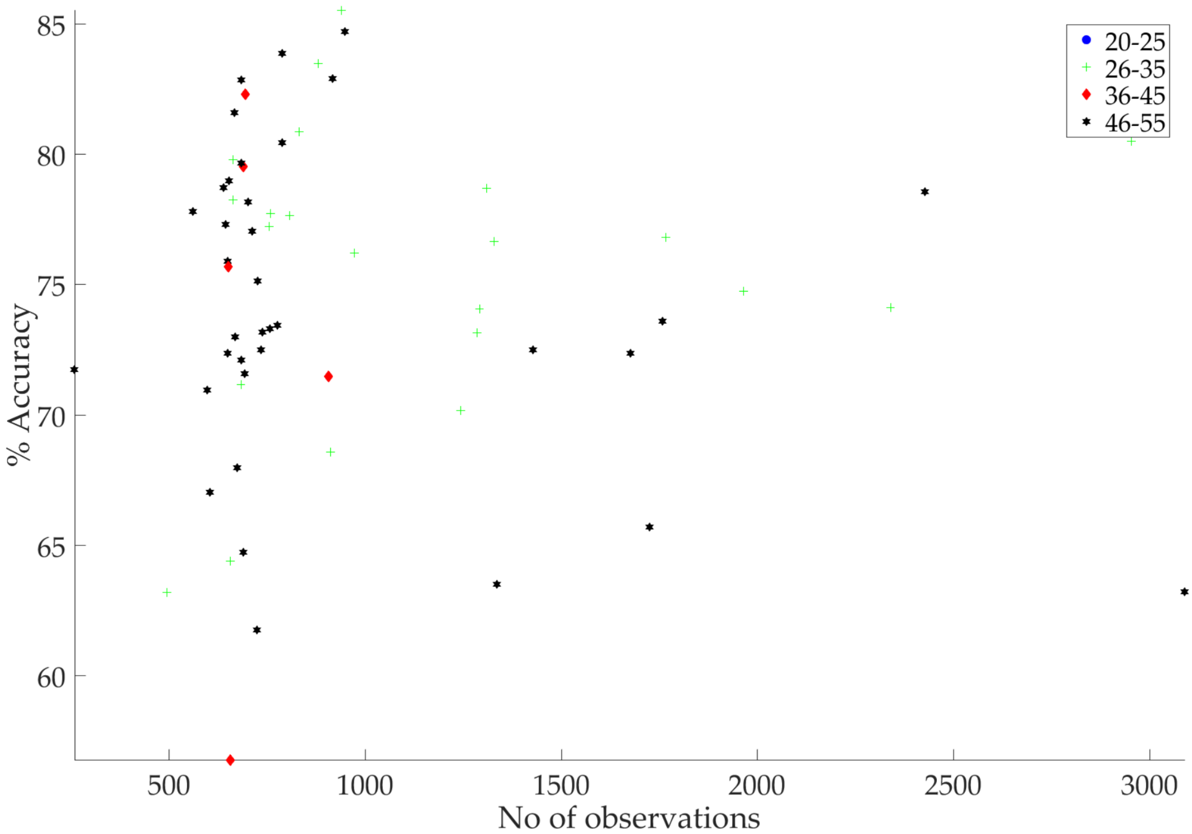


FIGURE 4.21: **Temporal Lobe:** Combined Cross Validation Results - sorted by Age

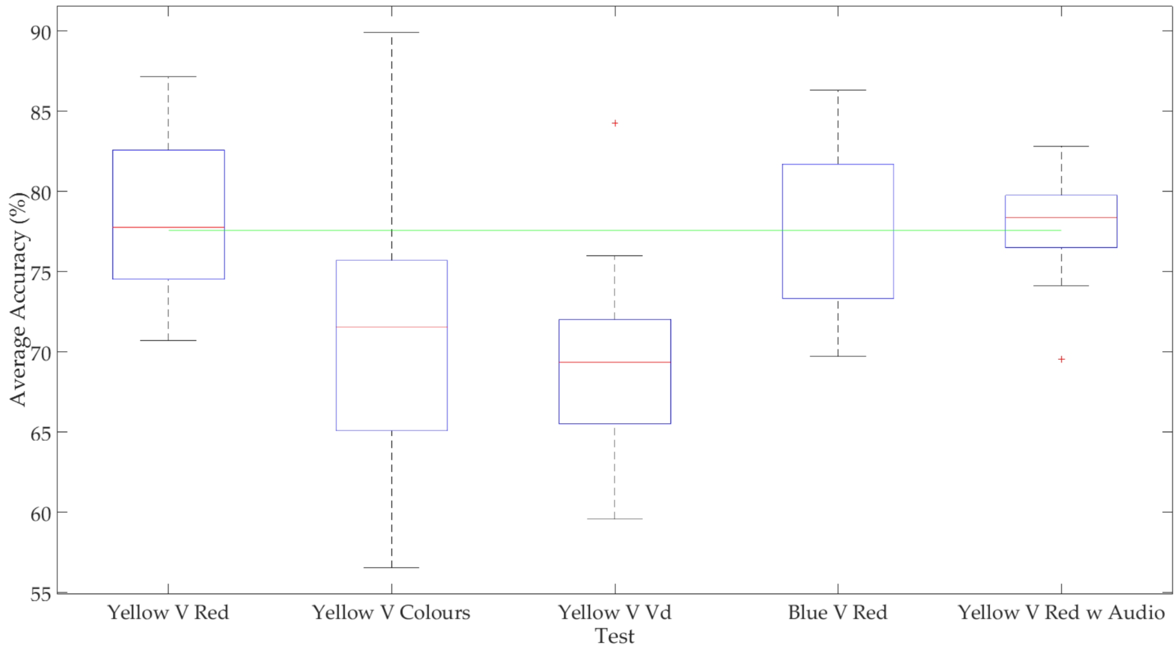


FIGURE 4.22: **Temporal Lobe:** Average Cross Validation Results for Subjects under 25 years of Age

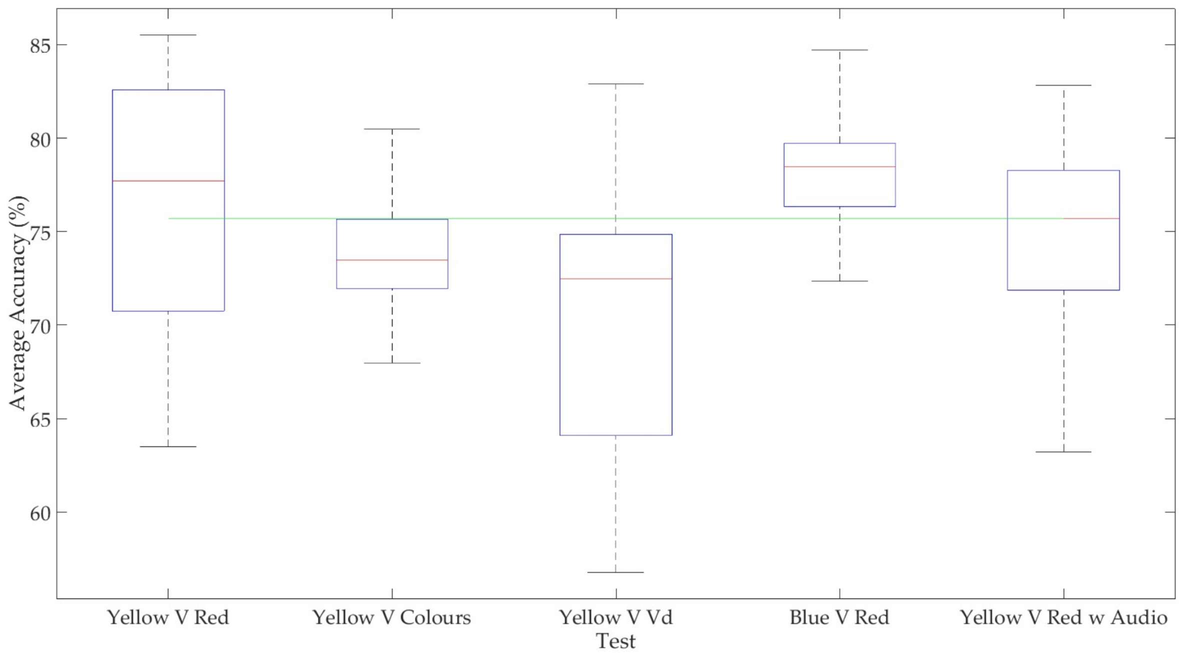


FIGURE 4.23: **Temporal Lobe:** Average Cross Validation Results for Subjects over 25 years of Age

Overall Participant Test Trends

Using individual temporal lobe data impacted adolescents' and adults' cross validation outcomes very subtly. In Figures 4.21, 4.22 and 4.23 it can be observed that there is a the tightening in the variance of results with the impact reversed between adolescents and adults, explicitly in the *Yellow Vs Vd* exercise. The spread of accuracy outcomes in the adults' tests has tightened with lower bounds increased in a similar manner to the *Yellow Vs Colours* exercise. Here the shift appears from previously the recorded figure of 63.5%, moving to 67.5%, with the median remaining more or less the same. The *Yellow Vs Red w Audio* exercise was the exception where the lower bound of data remained at 55% but the median improved 7%, from 66% to 73%, contrary to my personal speculation. My hypothesis that the aural distraction of music would require the subject to concentrate more, generating more frontal lobe activity and in turn better results for independent frontal lobe data, did not occur. Rather it verifies the theory that adolescent temporal lobe data produces superior results in light of aural distractions in an analogous manner to frontal lobe data producing better results in the *Yellow Vs Vd* exercise. An outcome driven by the elevated electrical activity in various regions of the brain attempting to suppress the distractions created by the stimulus. Here, the results reveal a clear contrast between adolescents' and adults' brain function, giving rise to theories consistent with previous studies. These studies, such as Dumontheil (2015)'s investigation into the behaviour of adolescents correlated to their respective stages of brain maturation, elucidate that the temporal lobe is the last section of the brain to completely develop up to the age of 25.

4.3.2 Gender

Overall Participant Test Trends

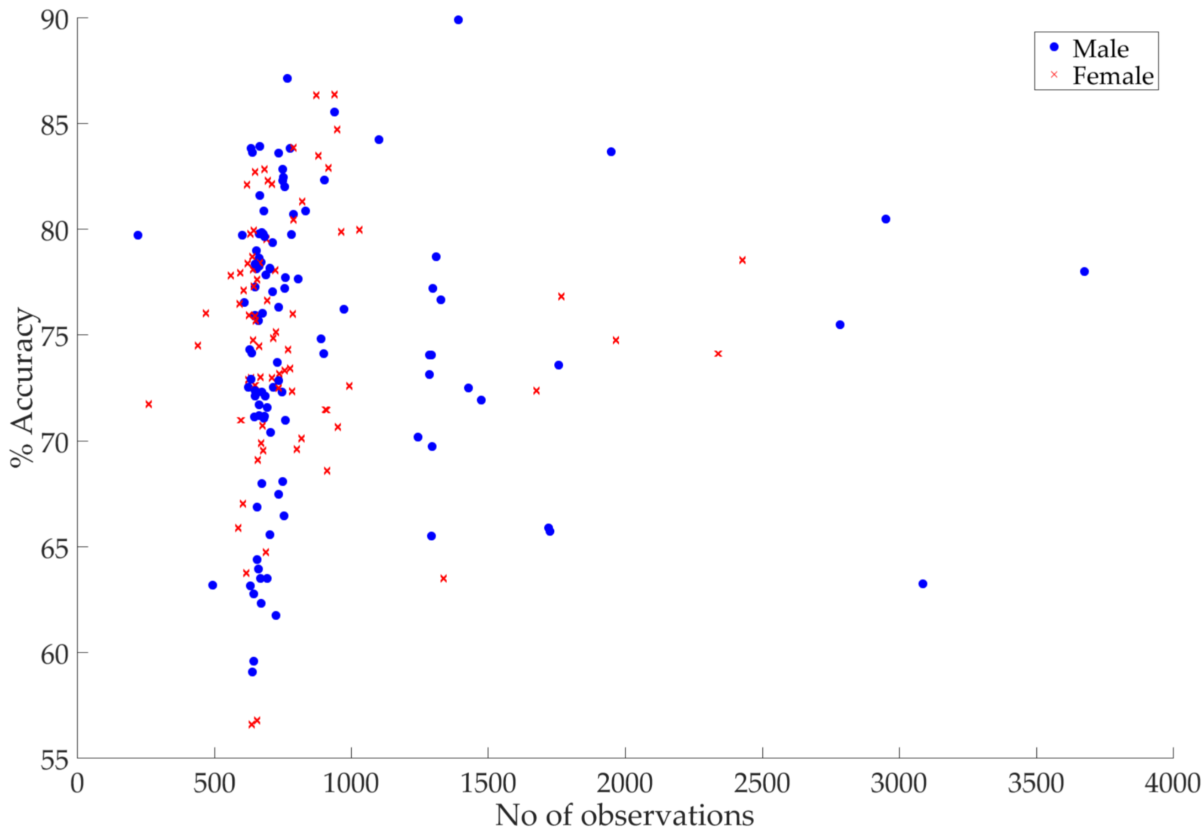


FIGURE 4.24: **Temporal Lobe:** Combined Cross Validation Results - sorted by Gender

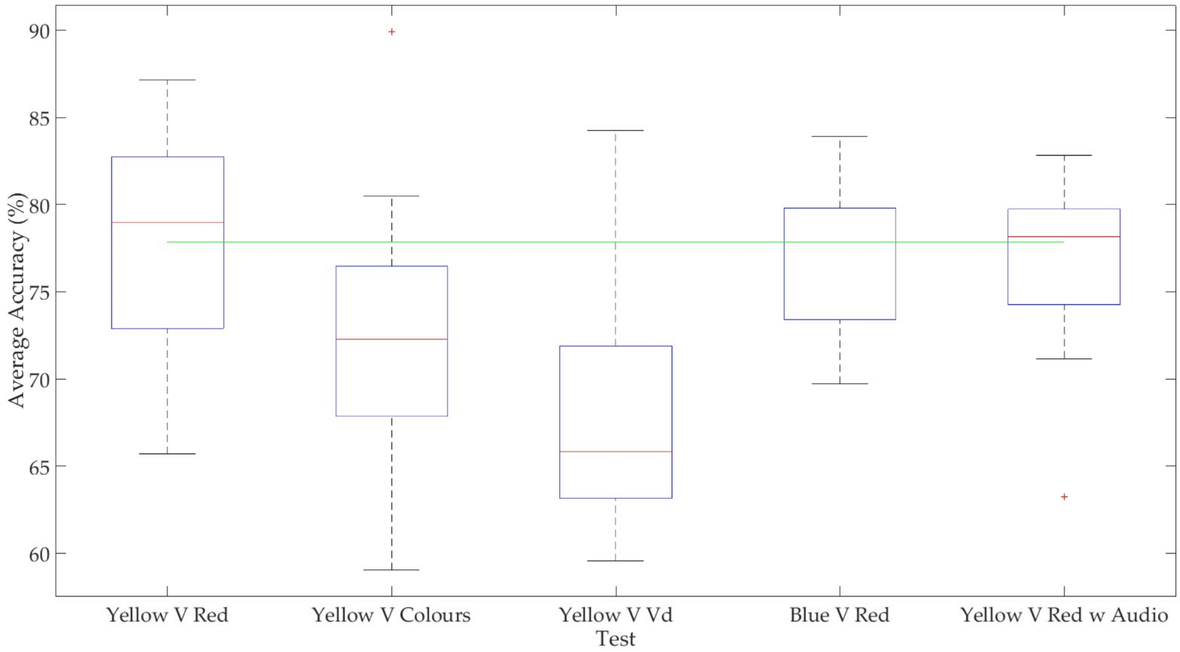


FIGURE 4.25: **Temporal Lobe:** Average Cross Validation Results for Male Subjects

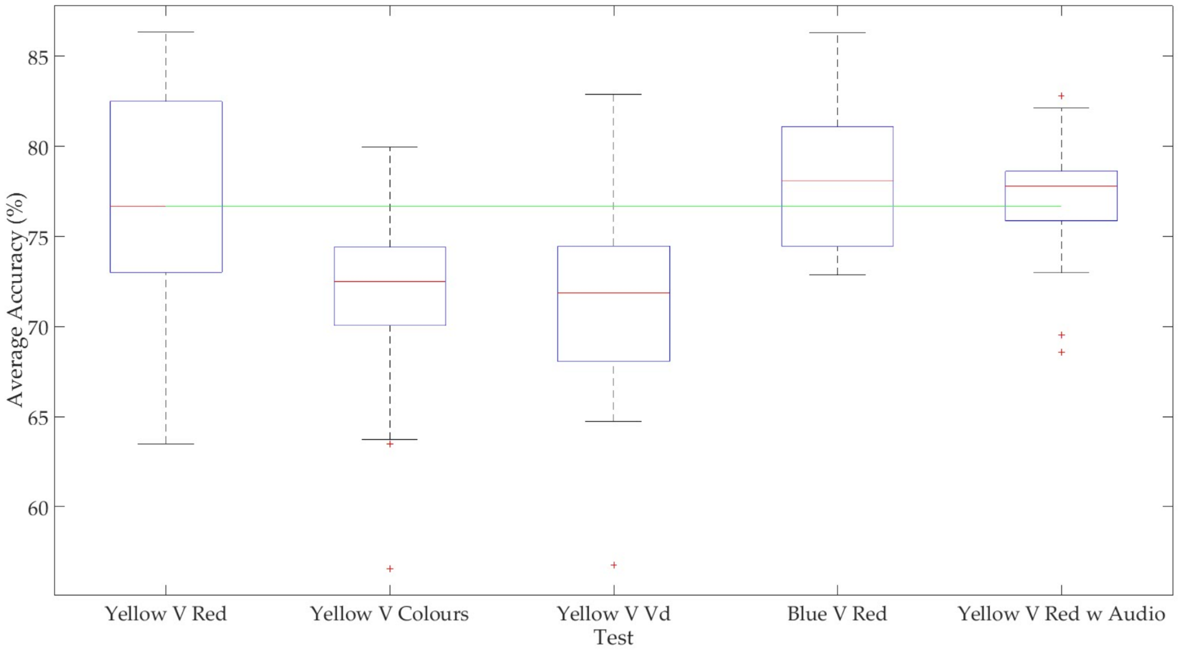


FIGURE 4.26: **Frontal Lobe:** Average Cross Validation Results for Female Subjects

In Figures 4.24, 4.25 and 4.26 the independent analysis of the temporal lobe for gender specifics highlighted that all the changes viewed in the *Age* section, such as subtle changes in *Yellow Vs Red*, could be attributed to specifically changes in *females'* cross validation data. Males' cross validation results did not differ from the "All sensors" data however the females' *Yellow Vs Colours* and *Yellow Vs Vd* exercises showed significant improvements moving from median accuracies of 70% and 68% to 73% and 72% respectively. Their range of score accuracies also were reduced by a similar amount for these two tests as the *Age* data supported for adults shifting from 55-78% and 55-83% bounds to 64-80% bounds and 65-83% bounds. This iterated that females were the dominant contributor of any variation in cross validation analysis when using temporal lobe independent data.

4.3.3 Machine Learning Parameters

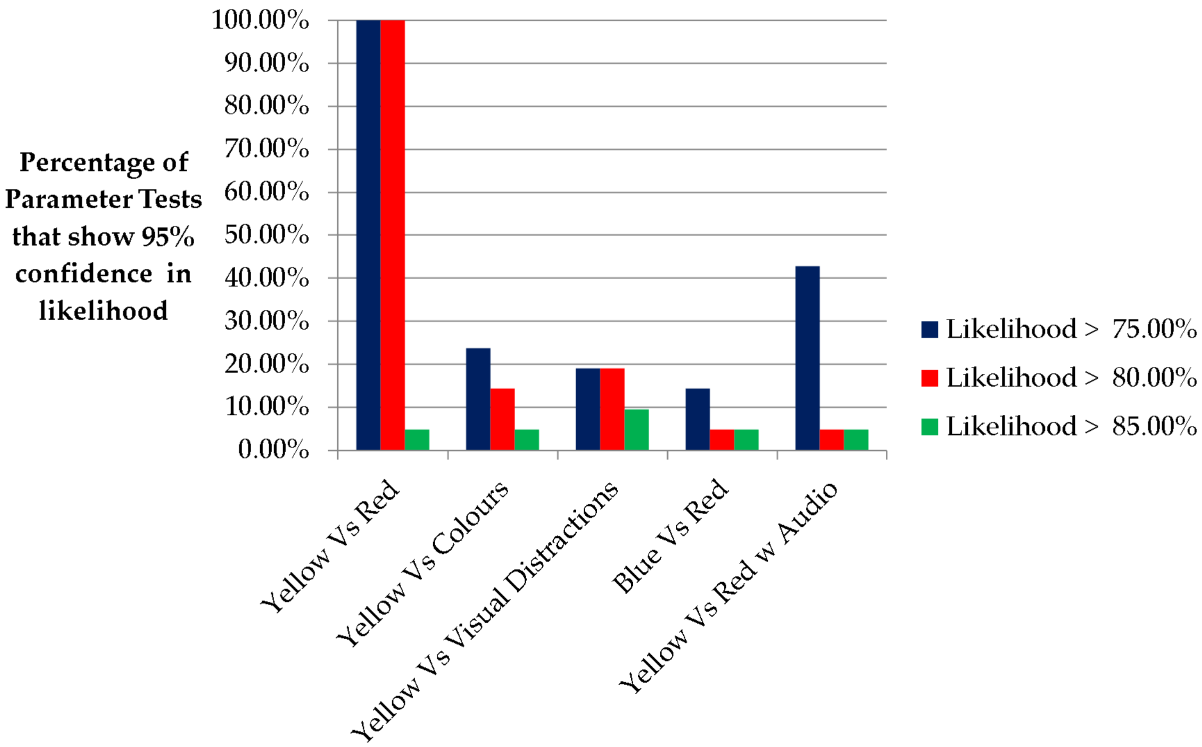


FIGURE 4.27: **Temporal Lobe:** Likelihood of Cross Validation Results for each Test

Complexity Factor	Gamma Parameter	Exponential	Accuracy
10	0	3	81.30%
10	10	0	92.05%
0.001	0.001	0	70.23%

FIGURE 4.28: **Temporal Lobe:** Best and Worst performing parameters

From Figure 4.27 it is immediately noticeable that the temporal lobe sensors independently are unable to generate accurate results based on the likelihoods presented, though using the complex and gamma parameters 10 and 10 respectively, average cross validation results of 92.05% can still be achieved. The *Yellow Vs Red* exercise still achieves a 100% probability (at 95% confidence) that the cross validation results will achieve a resulting score over 80%, a pattern that has been consistent with all independent electrode data. However the remaining tests do not nearly achieve the same magnitude of results demonstrated by the "Frontal Lobe" or "All Sensor" data as only 4 of the 15 tests managed to illustrate probabilities over 20% for their respective tests compared to 7 in "Frontal" and 9 in "All". The *Yellow Vs Red w Audio* exercise with the "Likelihood > 75%" bucket performed consistently among the 3 sets of data (Frontal, All, Temporal), recording a probability value of 42%. However the probabilities in the "Likelihood > 85%" indicated that any improvement in accuracy or reliability is very slim (4.76%). The machine learning parameters indicate that the best results and worst results are similar to the those of the frontal lobe data but with the high result of 92.05% achieved with the 10-10 complexity-gamma factors and the minimum score with a 10-10 combination scoring 88% (*Yellow Vs Vd)*, it would be redundant to claim that high averages cannot be achieved using only temporal lobe data.

4.4 Longitudinal Results

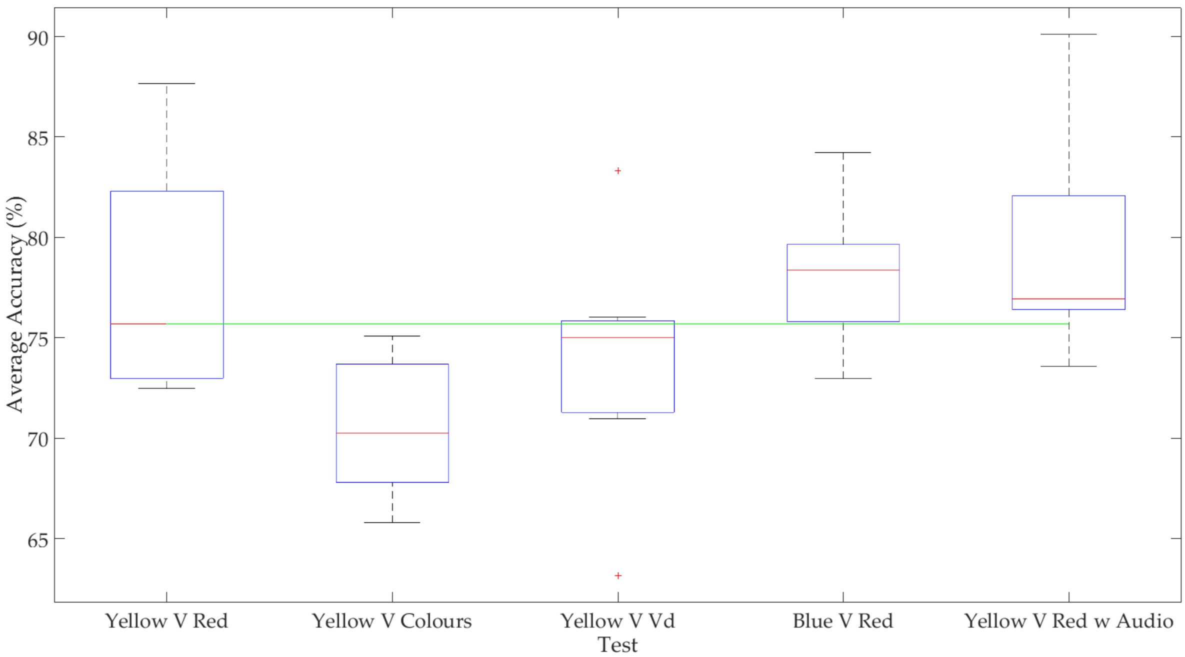


FIGURE 4.29: Average Results of Longitudinal Tests - Female

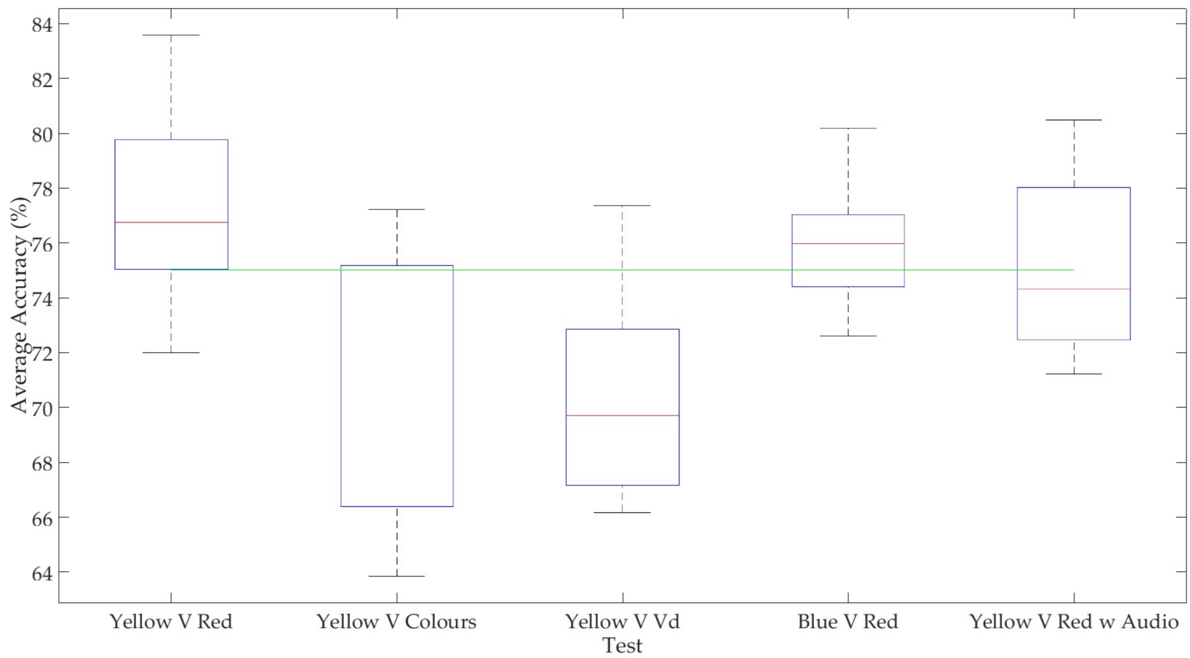


FIGURE 4.30: Average Results of Longitudinal Tests - Male

Subject and Test	Kernel	Complexity Parameter	Gamma (RBF)	Exponent (Poly)	Classification Accuracy(%)
Female Yellow Vs Red Combined	RBFKernel	10	10	0	89.88
Female Yellow Vs Colours Combined	RBFKernel	10	10	0	93.29
Female Yellow Vs Vd Combined	RBFKernel	10	10	0	91.81
Female Blue Vs Red Combined	RBFKernel	10	10	0	87.42
Female Yellow Vs Red w Audio Combined	RBFKernel	10	10	0	89.31
Male Yellow Vs Red Combined	RBFKernel	10	10	0	89.12
Male Yellow Vs Colours Combined	RBFKernel	10	10	0	92.34
Male Yellow Vs Vd Combined	RBFKernel	10	10	0	92.66
Male Blue Vs Red Combined	RBFKernel	10	10	0	87.71
Male Yellow Vs Red w Audio Combined	RBFKernel	10	10	0	87.71

TABLE 4.3: Longitudinal Cross Validation Results

Figures 4.29 and 4.30 display the average results of (all cross validation) training exercise sets performed for each of the 7 days indicating that the data results are not consistent when performing exercises and differ quite significantly between training periods. This is inferred from the large interquartile ranges presented for each of the male and female test subjects. The reliability of the machine learning algorithm is also showcased as attempting tests at different times can result in fluctuations in accuracies of more than 15% (see *Yellow Vs Colours* in Figure 4.30).

The cross validation of both participants results was performed with the RBF Kernel using complexity-gamma parameters 10-10 to portray that the SMO algorithm is able to produce classification accuracies above 85% for all tests regardless of external factors in the participants' environment. The cross validation results are presented in Table 4.3. The seven participant tests were all combined into one data file incorporating the different times at which tests were performed and different mindsets/moods of the participants. It is interesting to note that despite these unknown variables, the SMO algorithm was still able to achieve accuracy levels above 85%. While this cannot be stated confidently with only 2 subject participants, it provides insight into the scepticism that the algorithms will not be able to classify binary outputs under alternative environments accurately.

4.5 Summary

The results indicate that high classification accuracies can be achieved using the RBF Kernel with complexity-gamma parameters 10-10 and that gender bias is less prevalent in the determination of classification accuracy as results displayed between genders are very similar. The independent data analysis of frontal and temporal lobe sensors reveals that temporal lobe sensors are not ideal to use independently due to weak performance accuracies but frontal lobe sensors are, as they can achieve comparable accuracies to the "All sensors" analysis. Lastly, the results from the longitudinal test demonstrate machine learning algorithm's versatility and resilience to outliers by illustrating strong performance amidst the incorporation of external factors faced by subjects in their environments.

5 Discussion

5.1 EEG Headset Functionality

The Muse (2016) headset exhibited high levels of functionality, vindicating why it is the preferred brain recording modality used today. Consistent with Lin et al. (2010), the dry electrodes have been able to filter data efficiently and produce results with minimal effort by the investigator (myself) and the participants using the headset. Participants were able to move around freely and relax when wearing the headset as the live observations of their brain activity did not respond to minute subject movements such as jaw movement, blinking, head tilting or a combination of head movements as described by Bashivan, Rish, and Heisig (2016). Further, upon observation of the EEG raw signals acquired, it was seemingly obvious that wet electrodes were not required to attain any signal data. Not only did this enhance functionality, but improved practicality and ease of use for participants whilst still being able to produce very clear discrete brain signals that allowed me to construct a BCI with accuracies above 85% in any environment.

The idea of using the headset in any environment is the key focal point of the discussion in this section as highly accurate results (namely Table 4.3) can be achieved whilst subjects are in any state of mind. This fills the gap of previously conducted research as it has been usually very scientific in nature and tried to broaden the number of controlled variables in their experiments, such as indicated in studies conducted by Bashivan, Rish, and Heisig (2016), Bhushan, Wei, and Haddad (2005) or Sannelli et al. (2008). Authors outcomes are usually dependent on the presence of noise generated within their electrode signal. Particularly as those who utilised higher quality electrodes, conductive gel, an effective data filter and asked subjects to rest between exercises tended to discover better results. While this is not surprising, it does not echo the concept of practicality nor present the essence of functionality in the real world. The purpose of monitoring brain activity has been to aid physically challenged people through the means of communicating with a BCI and thus the results displayed highlight the extent to which this can be done. Achieving accuracies of 90% in each of the tests with complexity and gamma parameters of 10 and 10 therefore expresses a more meaningful result in the progression of EEG BCIs.

Moreover, the wireless nature of the EEG headset illustrates certain advantages and disadvantages. The advantages entailed the comparability of the EEG headset with smart devices which allows the headset to be used in conjunction with other services such as GPS which would aid a physically challenged subject to move around an unknown environment - a practical possibility explored by Rebsamen et al. (2006). Though alternatively, the disadvantages are presented in Figure 4.10 where the number of observations for each subject in training tests is vastly different and thus can deduce highly varying results as demonstrated by the box plots in Chapter 4.

Lastly, variations to the final results have been less impacted by issues surrounding the comfort of test participants that usually arise among several different studies. These issues included tiredness from doing repetitive tests, discomforts with a headset and elongated testing times. Tests were conducted within time-frames of 10mins and participants felt very comfortable with the experimental set-up.

5.2 Universality

The concept of BCI illiteracy was discussed in Chapter 2, specifically with Brunner et al. (2010) emphasising that 20% of subjects usually exhibit some form of BCI illiteracy for a multitude of reasons. The outcomes of all cross validation tests within this thesis study also showed weaknesses in certain tests for some subjects. Some examples include Subject 29: A 50 year old female who presented cross validation averages of less than 70% for 3 tests, reinforcing this concept. However, the optimised 10-10 complexity-gamma parameter combination showcased that this subject was able to achieve a 98% accuracy with the RBF Kernel using the *Yellow Vs Colours* test. Nonetheless, along with many other elder female participants, this subject was unable to produce any useful[1] results for *Blue Vs Red* and *Yellow Vs Red w Audio* test as the signal was not strong enough despite sufficient contact on her forehead. Thus the existence of BCI illiteracy is resounded within my EEG experiments, consistent with the findings of Blankertz et al. (2008) who also established that improved training could reduce the number of participants who exhibit illiteracy.

Universality has consequently not been achieved but similarities determined between males and females, adults and adolescents in all three data sets encompasses a progression in universality. The ability for all test subjects to use the EEG headset and acquire results that are alike to peer test subjects is an important development in the usage of the machine learning algorithm with EEG headset. Although the headset's shape was more biased towards males and certain head shapes, the versatility of its electrodes and minimalist design delineates that EEGs are becoming more universal with time and technology.

[1] *useful* defined as not having any corrupt lines of raw data

Another point of interest lies in the feasibility of using adolescents for EEG testing as their cross validation performances generally did not differ greatly from their adult counterparts. Rather, they appeared to be stronger in some tests and weaker in others, more likely attributed to their behavioural characteristics such as fidgeting and lack of attention span than their brain activity specifically.

5.3 Machine Learning Algorithms

The analysis of various machine learning parameters used and utilisation of using differing features namely, Temporal Lobe and Frontal Lobe data sets independently has elucidated the importance of using more features from other sources to improve accuracy as all likelihood classes at the highest performance in 'All Sensors' when compared to 'Frontal Lobes' and 'Temporal Lobes' only. However to improve practicality of EEG headsets, authors are attempting to achieve better results with only 2 electrodes such as demonstration of a frontal lobe only BCI by McCrimmon et al. (2017). And while they are succeeding on limited measures, Hinterberger et al. (2004) give the reason as to why additional brain features will provide an improvement of results as showcased in Chapter 4 - they postulate that in training tests where frontal lobes do not exhibit desirable features, a hybrid use of temporal lobe features contributes to an improved machine learning output (and vice versa). My findings reinforce this idea as Figures 4.14 iterated that there was a significant improvement in results observed for all tests except for *Yellow Vs Red w Audio*, and the overall median of test results rose above the 75% mark. Additionally, frontal lobe data had also increased the cross validation scores of all test subjects who had poor results ($\leq$60%) with 'All Sensors' data.

The SMO machine learning algorithms showed that they modelled high speeds but still performed admirably despite many literature surveys recommending the use of ANN as the preferred machine learning algorithm because neural networks are more powerful than the SMO algorithms. In constructing my methodology I used ANN to perform a single cross validation test which took 6.4 seconds [2] whilst the SMO algorithm performed the same cross validation for me in 0.65 seconds. This concept of speed here becomes important in *real time* BCI applications as subjects are not expected to wait for machines to perform calculations. Further, the ability to adjust machine learning parameters also plays a significant role in the SMO algorithm selection process. The option to use 21 different combinations of parameter selections (9 PolyKernel and 12 RBF Kernel) was only available with the SMO algorithm with more additional options to utilise other kernels (though this was not pursued). Machine learning models face the challenge of over-fitting data where there is too much discrimination between data sets provided, usually resulting from a complexity parameter that is too high or

[2] On a 1000 instance data set

having too many instances (observations). This clarifies the outlier in Figure 4.7 where a data point with over 3000 observations resulted in the worst outcome (57% accuracy). Figure 5.1 demonstrates the balance between have too little data and too much data, in particular as too much noise in features can detract machines from learning the underlying objective (such as brain activity induced from fidgeting rather than recognition of Yellow in an adolescent to switch on a light bulb).

The employment of multiple brainwave bands as features was a subject of concern as too much data precedes slow computation time. Li, Xu, and Zhu (2015) used their study to demonstrate that alpha and beta waves alone were not suitable in achieving satisfactory results when subjects were concentrating and, literature had not attempted to discriminate between the data processing of electrodes at the temporal and frontal lobes of the brain. Thus, it was ideal to investigate the concept of using multiple bands. There was though, consideration in only using specific wave bands as features, as previously done in studies, but due to our limited understanding of brain activity in performing explicit tasks, I decided against doing so. Thinking of the colour yellow can induce magnitude changes in any type of waveband and as significant analysis on brain function has not been done for this activity it was best to attempt using all incoming data. There is potential to explore condensing the number of features used for more efficient machine learning but this was not examined in my study.

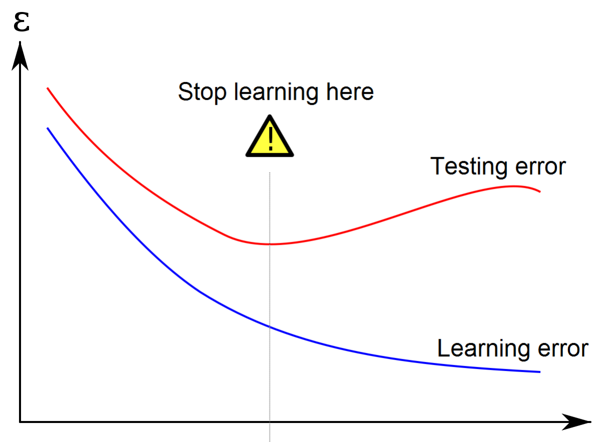


FIGURE 5.1:
[3]Over-fitting leads to Testing error - Kacmajor (2016)

[3]x-axis is not labelled in this source but should be defined as Learning Volume (does not have units)

5.4 Limitations and Weaknesses

While this thesis highlights a progression in our understanding of the capabilities of technology surrounding EEG and machine learning, it also indicates a number of limitations and weaknesses that should be considered:

1. **Binary Output Only**: Subjects' thoughts are measured in binary output only which is problematic even for switching lights on/off as there are 3 states that must be considered: *switch on*, *switch off*, *do nothing*. To apply the results of this thesis into a live environment means that if a subject is not thinking about the colour yellow then the light will switch off because the program needs an indicator to switch a light on *and* off.

2. **Delay in Stimuli Recognition**: McCrimmon et al. (2017) justified removing the first second of data features in their machine learning analysis to eliminate noise at the expense of delays. This methodology was not adopted as the aim of the experiments was to create the most practical system yet by not doing so exposes the results to a time delay risk. This time delay is elicited by the processing time required for a participant's brain to recognise a colour change on the screen. It may seem negligible at first glance, but further analysis reveals that consideration must be taken as the CPU's computing time is in the order of microseconds while the human brain processes in the order of milliseconds (Trafton, 2014). This means although the computer assumes the subject has recognised a colour change, in reality, it is may have classified the subject's state too early.

3. **Longitudinal Data Reliability**: The number of subjects used for longitudinal data was minimal (two) and isn't enough to reliably attest a proof of concept. The strength of the longitudinal data results lie in subject selection as an adolescent male's results were contrasted in parallel to an adult female with the machine learning able to highlight equally positive results.

This discussion explores that despite weaknesses in thesis conduction, the investigation has yielded positive results. It has displayed superior use of a commercial EEG headset in flexible environments, advanced upon exhibiting increased universality and applied machine learning techniques to multiple features whilst keeping computing time low and accuracy high.

6 Conclusion

6.1 Uncovering Objectives

The objectives of this study seek to explore the capabilities of a consumer grade EEG system in a more technologically adapt society with the more specific focus on exploring their ability to interact with a BCI in one's natural surroundings. The outcomes of this thesis demonstrate that there has been a huge leap in BCI development with machine learning algorithms running in real-time at accuracies surpassing 85%, with EEG headsets that cost less than \$300. Results of this study also indicate heightened reliability in the 'real world' as the performance of 39 subjects is documented under a multitude of environments and thus skewness in results cannot be attributed to the presence of controlled environmental characteristics. The following outlines how the conclusions in this investigation were reached.

The construction of a BCI system, whereby five tests were used to capture brain activity of subjects when they envisage the colour yellow, displayed a range of outcomes that cannot incite concrete conclusions. Rather, general trends showed improved cross validation performance in adolescents against adults with an absence of class superiority between males and females contrary to the findings of other studies (Lee, 2005). However my hypothesis that subjects' cross validation results would perform better among tests that required increased concentration (the idea that stronger brainwaves could force better results) was disproved in this instance as simpler tests produced superior results against more complex ones by orders of up to 10% with median results exceeding 75% accuracy.

The conduction of longitudinal tests, spanning 14 days to analyse and predict subjects' various states of minds at arbitrary times with the same machine learning algorithm suggests an adjustment for future methodologies. Methodologies have to-date been conducted using ordered procedures to delineate proof-of-concept such as those conducted by Rebsamen et al. (2006) or Siuly, Li, and Wen (2011) who achieved accuracies above 95%. However, authors should aim to find the impact of participants' various states of mind upon their research before drawing any conclusion as this thesis shows that subjects can have large performance outcomes whilst attempting an identical test on different days. Yet nonetheless, tuning algorithms to utilise more efficient machine learning parameters can reduce this variance and endorse truly exceptional results. Performing machine learning analysis on filtered electrodes, that is either on temporal or frontal lobe data revealed that the number of electrodes required to accomplish such outcomes is reducing and the efficacy of building a device that is a natural extension of our day to day lives appears to be becoming a certain reality.

6.2 Future Development and Applications

Academic research in the EEG-BCI field has been growing exponentially and thus there is a colossal volume of research still to be conducted in order to truly comprehend the capabilities of today's technology and brain activity.

Further application of *this* investigation originates from the concept of using more specific thoughts to enable an action as opposed to a general movement (motor task). Literature has demonstrated that the catalyst for capturing brain activity has always associated with a task involving a large variance in signal data - almost so, that a person can detect the existence of the catalyst, when visually observing the electrode signals. However my research strives to push the limitations of society's understandings because it has encapsulated a methodology that is naturally thought provoking for a subject in a more niche environment (ie. switching a light bulb by thinking about the colour yellow) rather than generally, such as concentrating/relaxing. Here, brain activity changes are very subtle but with machine learning becoming more powerful in modern context, it is able to detect the presence of it amidst a large noise to signal ratio. This creates an undoubtedly huge leap in the movement of BCI application as society will be able to apply more specific thoughts to other applications such as driving a car where the thought of stopping a car alone would be sufficient, to do so [1].

Extensions of this study can be performed in multiple ways to enrich society's understanding of the EEG-BCI framework. For this purpose, all raw data is included in this study and the following suggestions may be able to acknowledge many questions that some of the outcomes that this study provides:

1. Using the colour red and blue as catalysts alongside yellow and implementing them into the BCI with each colour allocated to its own action.
2. Use different machine learning algorithms with various features and EEG headsets. Comparing the performance of the 3 characteristics of a BCI (headset, algorithm and features) with their counterparts will answer the question: *'Which BCI system will be the most efficient to use today to establish a well-functioning brain controlled environment?'*

[1]See video: World Economic Forum (2017)

3. Live testing the results of this study. While analysis was adequately performed and timing delays were counted, there are unforeseen issues that can arise when using a BCI system live begging the question: *'Will it perform in the real life?'* .
   Recognising and imagining yellow may induce completely different brain activity. Li et al (2015) showed that the SVM performed poorly in the BCI. Although the cross validation testing took 0.43s to perform when tested in cases with less than 1000 instances of data, the time delay in capturing the data, performing the cross validation and then sending information back to the controller could exceed more than 1.5s. This is an issue in real-time BCI as you cannot expect such long delay times as most users expect instantaneous changes. They become hasty and give up (particularly demonstrated by the Adolescent Age Group general participant analysis in Section 4). Therefore live testing forms a vital part of future development stemming from this study.

Answering the questions originating from this study and pursuing some of the future developments that grow upon the fundamentals explored in this thesis will continue to drive innovation in the BCI field. It will change the lives of people who are handicapped and also set the grounds for a new level of human-machine interaction that the world has never seen before.

Glossary

Brain imaging modality: A technique through which brain activity and brain waves are measured.

Cerebral Related to the brain. Particularly to the cerebral cortex the part of the brain - involved in day to day functionality.

Complexity Parameters (C): The SMO complexity parameter is used to control how soft margins are between the decision boundary and the nearest training points. The C parameter is defined as the characteristic which informs the algorithm how much the user wants to avoid misclassifying data points. A larger value of C will choose very small margins, attempting to classifying all the points correctly. Think of this as a line of best fit that weaves in between points to ensure it has adequately separated the training data as opposed to drawing a just straight line with one or two data points that have been misclassified.

Gamma parameter: defined by the influence of each training point on the final decision boundary. A smaller value of gamma will only capture the influence of points very close to the initial estimated decision boundary, and thus will not capture the complexity of the data as the region of influence from each data point is limited.

Curse of dimensionality: A phenomenon that describes the issue surrounding an increased number of feature dimensions when analysing data. An example of the curse of dimensionality can be thought of through the process of a child choosing the best cookie in a cookie factory. The more cookie characteristics that a child takes into account to pick their favourite cookie, the longer they will take to make a decision. Think about selecting the best cookie by only flavour over using colour, flavour, crunchiness, temperature, price etc.

Cross validation: Partitioning data into k sets, where the testing set is a kth partition of the whole set and the rest of the data is used as a training set. The result is recorded and repeated k-times with each of the k sets used as the testing set and the remaining data used as a training set.

Features (Machine Learning): a characteristic of the data being observed. It is selected by the user for efficient machine learning and pattern recognition. A simple example entails looking at the volume of dark pixels in a photograph to classify whether the photograph was taken during the day or at night.

Real time: defined as performing an event in a constrained time usually perceived by humans as instantly. By computing standards, this may be in the order of milliseconds.

Bibliography

Allison, B. Z. et al. (2010). “Toward a hybrid brain–computer interface based on imagined movement and visual attention”. en. In: *Journal of Neural Engineering* 7.2, p. 026007. ISSN: 1741-2552. DOI: 10.1088/1741-2560/7/2/026007. URL: http://stacks.iop.org/1741-2552/7/i=2/a=026007.

Alonso-Valerdi, Luz Maria, Ricardo Antonio Salido-Ruiz, and Ricardo A. Ramirez-Mendoza (2015). “Motor imagery based brain–computer interfaces: An emerging technology to rehabilitate motor deficits”. In: *Neuropsychologia*. Special Issue: Sensory Motor Integration 79, Part B, pp. 354–363. ISSN: 0028-3932. DOI: 10.1016/j.neuropsychologia.2015.09.012. URL: http://www.sciencedirect.com/science/article/pii/S002839321530155X.

AlZoubi, Omar, Rafael A. Calvo, and Ronald H. Stevens (2009). “Classification of EEG for Affect Recognition: An Adaptive Approach”. en. In: *AI 2009: Advances in Artificial Intelligence*. Springer, Berlin, Heidelberg, pp. 52–61. DOI: 10.1007/978-3-642-10439-8_6. URL: https://link-springer-com.ezproxy1.library.usyd.edu.au/chapter/10.1007/978-3-642-10439-8_6 (visited on 06/11/2017).

Bashivan, Pouya, Irina Rish, and Steve Heisig (2016). “Mental State Recognition via Wearable EEG”. In: *arXiv:1602.00985 [cs]*. arXiv: 1602.00985. URL: http://arxiv.org/abs/1602.00985 (visited on 06/05/2017).

Berg, Daniel (2012). *NeuroSky Mindwave Mobile*. URL: https://www.laptopmag.com/reviews/accessories/neurosky-mindwave-mobile (visited on 09/25/2017).

Bhushan, Bharat, Guohua Wei, and Paul Haddad (2005). “Friction and wear studies of human hair and skin”. In: *Wear*. 15th International Conference on Wear of Materials 259.7, pp. 1012–1021. ISSN: 0043-1648. DOI: 10.1016/j.wear.2004.12.026. URL: http://www.sciencedirect.com/science/article/pii/S0043164804004533.

Blankertz, B. et al. (2006). “The Berlin brain-computer interface: EEG-based communication without subject training”. In: *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 14.2, pp. 147–152. ISSN: 1534-4320. DOI: 10.1109/TNSRE.2006.875557.

Blankertz, Benjamin et al. (2008). “The Berlin Brain–Computer Interface: accurate performance from first-session in BCI-naïve subjects”. eng. In: *IEEE transactions on* *bio-medical engineering* 55.10, pp. 2452–2462. ISSN: 1558-2531. DOI: 10.1109/TBME.2008.923152.

Bostanov, V. (2004). “BCI competition 2003-data sets Ib and IIb: feature extraction from event-related brain potentials with the continuous wavelet transform and the t-value scalogram”. In: *IEEE Transactions on Biomedical Engineering* 51.6, pp. 1057–1061. ISSN: 0018-9294. DOI: 10.1109/TBME.2004.826702.

Brunner, Clemens et al. (2010). “Improved signal processing approaches in an offline simulation of a hybrid brain–computer interface”. In: *Journal of neuroscience methods* 188.1, pp. 165–173. ISSN: 0165-0270. DOI: 10.1016/j.jneumeth.2010.02.002. URL: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3422070/.

Chiappa, Silvia and Samy Bengio (2003). *HMM and IOHMM modeling of EEG rhythms for asynchronous BCI systems*. Tech. rep. IDIAP.

Classifier Parameter Optimisation (2017). URL: http://www.coxdocs.org/doku.php?id=perseus:user:activities:matrixprocessing:learning:classificationparameteroptimization (visited on 11/01/2017).

Copeland, Michael (2016). *The Difference Between AI, Machine Learning, and Deep Learning? | NVIDIA Blog*. URL: https://blogs.nvidia.com/blog/2016/07/29/whats-difference-artificial-intelligence-machine-learning-deep-learning-ai/ (visited on 06/06/2017).

Curtin, A. et al. (2012). “A P300-based EEG-BCI for spatial navigation control”. In: *2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 3841–3844. DOI: 10.1109/EMBC.2012.6346805.

David, Olivier, James M. Kilner, and Karl J. Friston (2006). “Mechanisms of evoked and induced responses in MEG/EEG”. eng. In: *NeuroImage* 31.4, pp. 1580–1591. ISSN: 1053-8119. DOI: 10.1016/j.neuroimage.2006.02.034.

Developers, Muse (2017). *Research Tools Bluetooth Connectivity*. URL: http://developer.choosemuse.com/hardware-firmware/bluetooth-connectivity/research-tools-bluetooth-connectivity (visited on 09/24/2017).

Dumontheil, Iroise (2015). “Development of the social brain during adolescence”. In: *Psicología Educativa*. Neuroscience and education: We already reached the tipping point [Neurociencia y educación: ya hemos alcanzado el punto crítico] 21.2, pp. 117–124. ISSN: 1135-755X. DOI: 10.1016/j.pse.2015.08.001. URL: http://www.sciencedirect.com/science/article/pii/S1135755X15000172.

Farwell, L. A. and E. Donchin (1988). “Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials”. In: *Electroencephalography and Clinical Neurophysiology* 70.6, pp. 510–523. ISSN: 0013-4694. DOI: 10.1016/0013-4694(88)

90149-6. URL: http://www.sciencedirect.com/science/article/pii/0013469488901496.

Fazel-Rezai, Reza and Waqas Ahmad (2011). “P300-based Brain-Computer Interface Paradigm Design”. en. In: DOI: 10.5772/14858. URL: http://www.intechopen.com/books/recent-advances-in-brain-computer-interface-systems/p300-based-brain-computer-interface-paradigm-design (visited on 10/30/2017).

Ferree, T. C. et al. (2001). “Scalp electrode impedance, infection risk, and EEG data quality”. eng. In: *Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology* 112.3, pp. 536–544. ISSN: 1388-2457.

Ge, S. et al. (2017). “A Brain-Computer Interface Based on a Few-Channel EEG-fNIRS Bimodal System”. In: *IEEE Access* 5, pp. 208–218. ISSN: 2169-3536. DOI: 10.1109/ACCESS.2016.2637409.

Gerven, Marcel van et al. (2009). “The brain-computer interface cycle”. eng. In: *Journal of Neural Engineering* 6.4, p. 041001. ISSN: 1741-2552. DOI: 10.1088/1741-2560/6/4/041001.

Guger, Christoph et al. (2012). “Comparison of Dry and Gel Based Electrodes for P300 Brain–Computer Interfaces”. In: *Frontiers in Neuroscience* 6. ISSN: 1662-4548. DOI: 10.3389/fnins.2012.00060. URL: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3345570/.

Hari, Riitta and Aina Puce (2017). *MEG-EEG Primer*. en. Oxford University Press. ISBN: 978-0-19-049777-4.

Hinterberger, Thilo et al. (2004). “A multimodal brain-based feedback and communication system.” English. In: *Experimental brain research* 154.4, pp. 521–526. ISSN: 0014-4819. URL: http://search.proquest.com.ezproxy1.library.usyd.edu.au/docview/80126853/F8AACD15143543C1PQ/1.

Hiraiwa, A., K. Shimohara, and Y. Tokunaga (1989). “EEG topography recognition by neural network”. In: *Conference Proceedings., IEEE International Conference on Systems, Man and Cybernetics*, 1116–1117 vol.3. DOI: 10.1109/ICSMC.1989.71473.

Hochberg, Leigh R. et al. (2006). “Neuronal ensemble control of prosthetic devices by a human with tetraplegia”. en. In: *Nature* 442.7099, pp. 164–171. ISSN: 0028-0836. DOI: 10.1038/nature04970. URL: https://www.nature.com/nature/journal/v442/n7099/abs/nature04970.html (visited on 06/08/2017).

Hoffmann, U. et al. (2005). “A Boosting Approach to P300 Detection with Application to Brain-Computer Interfaces”. In: *Conference Proceedings. 2nd International IEEE EMBS Conference on Neural Engineering, 2005.* Pp. 97–100. DOI: 10.1109/CNE.2005.1419562.

Hwang, Han-Jeong et al. (2013). "EEG-Based Brain-Computer Interfaces: A Thorough Literature Survey". In: *International Journal of Human–Computer Interaction* 29.12, pp. 814–826. ISSN: 1044-7318. DOI: 10.1080/10447318.2013.780869. URL: http://dx.doi.org/10.1080/10447318.2013.780869 (visited on 05/29/2017).

Ilyas, M. Z. et al. (2016). "Classification of EEG signals for brain-computer interface applications: Performance comparison". In: *2016 International Conference on Robotics, Automation and Sciences (ICORAS)*, pp. 1–4. DOI: 10.1109/ICORAS.2016.7872610.

James, Manyika et al. (2011). "Big data: the next frontier for innovation, competition, and productivity". In: *The McKinsey Global Institute*.

Kaper, Matthias et al. (2004). "BCI Competition 2003–Data set IIb: support vector machines for the P300 speller paradigm". eng. In: *IEEE transactions on bio-medical engineering* 51.6, pp. 1073–1076. ISSN: 0018-9294. DOI: 10.1109/TBME.2004.826698.

Kaur, J. and A. Kaur (2015). "A review on analysis of EEG signals". In: *2015 International Conference on Advances in Computer Engineering and Applications*, pp. 957–960. DOI: 10.1109/ICACEA.2015.7164844.

Kropotov, Juri D. (2010). *Quantitative EEG, Event-Related Potentials and Neurotherapy*. en. Google-Books-ID: szECZPz7FvcC. Academic Press. ISBN: 978-0-08-092297-3.

Kübler, A. et al. (2006). "BCI Meeting 2005–workshop on clinical issues and applications". eng. In: *IEEE transactions on neural systems and rehabilitation engineering: a publication of the IEEE Engineering in Medicine and Biology Society* 14.2, pp. 131–134. ISSN: 1534-4320.

Lee, Janice W. (2005). *Focus on Gender Identity*. en. Google-Books-ID: jvdGniM1eZ4C. Nova Publishers. ISBN: 978-1-59454-212-1.

Leuthardt, Eric C. et al. (2004). "A brain-computer interface using electrocorticographic signals in humans". eng. In: *Journal of Neural Engineering* 1.2, pp. 63–71. ISSN: 1741-2560. DOI: 10.1088/1741-2560/1/2/001.

Li, Y. et al. (2010). "An EEG-Based BCI System for 2-D Cursor Control by Combining Mu/Beta Rhythm and P300 Potential". In: *IEEE Transactions on Biomedical Engineering* 57.10, pp. 2495–2505. ISSN: 0018-9294. DOI: 10.1109/TBME.2010.2055564.

Li, Zhen, Jianjun Xu, and Tingshao Zhu (2015). "Prediction of Brain States of Concentration and Relaxation in Real Time with Portable Electroencephalographs". In: *arXiv:1509.07642 [cs]*. arXiv: 1509.07642. URL: http://arxiv.org/abs/1509.07642.

Lin, C. T. et al. (2008). "Noninvasive Neural Prostheses Using Mobile and Wireless EEG". In: *Proceedings of the IEEE* 96.7, pp. 1167–1183. ISSN: 0018-9219. DOI: 10.1109/JPROC.2008.922561.

Lin, C. T. et al. (2011). "Novel Dry Polymer Foam Electrodes for Long-Term EEG Measurement". In: *IEEE Transactions on Biomedical Engineering* 58.5, pp. 1200–1207. ISSN: 0018-9294. DOI: 10.1109/TBME.2010.2102353.

Lin, Chin-Teng et al. (2010). "Review of wireless and wearable electroencephalogram systems and brain-computer interfaces–a mini-review". eng. In: *Gerontology* 56.1, pp. 112–119. ISSN: 1423-0003. DOI: 10.1159/000230807.

Lopez-Gordo, M. A., D. Sanchez-Morillo, and F. Pelayo Valle (2014). "Dry EEG Electrodes". en. In: *Sensors* 14.7, pp. 12847–12870. DOI: 10.3390/s140712847. URL: http://www.mdpi.com/1424-8220/14/7/12847 (visited on 06/11/2017).

Lotte, F. et al. (2007). "A review of classification algorithms for EEG-based brain-computer interfaces". eng. In: *Journal of Neural Engineering* 4.2, R1–R13. ISSN: 1741-2560. DOI: 10.1088/1741-2560/4/2/R01.

Lotte, Fabien et al. (2015). "Electroencephalography (EEG)-Based Brain–Computer Interfaces". en. In: *Wiley Encyclopedia of Electrical and Electronics Engineering*. DOI: 10.1002/047134608X.W8278. John Wiley & Sons, Inc. ISBN: 978-0-471-34608-1. URL: http://onlinelibrary.wiley.com/doi/10.1002/047134608X.W8278/abstract.

L'Heureux, A. et al. (2017). "Machine Learning With Big Data: Challenges and Approaches". In: *IEEE Access* 5, pp. 7776–7797. ISSN: 2169-3536. DOI: 10.1109/ACCESS.2017.2696365.

Mak, Joseph N. and Jonathan R. Wolpaw (2009). "Clinical Applications of Brain-Computer Interfaces: Current State and Future Prospects". In: *IEEE reviews in biomedical engineering* 2, pp. 187–199. ISSN: 1937-3333. DOI: 10.1109/RBME.2009.2035356. URL: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2862632/.

Manyakov, Nikolay V. et al. (2011). "Comparison of Classification Methods for P300 Brain-Computer Interface on Disabled Subjects". en. In: *Computational Intelligence and Neuroscience* 2011, e519868. ISSN: 1687-5265. DOI: 10.1155/2011/519868. URL: https://www.hindawi.com/journals/cin/2011/519868/abs/ (visited on 06/11/2017).

McCrimmon, C. M. et al. (2017). "Performance Assessment of a Custom, Portable, and Low-Cost Brain #8211;Computer Interface Platform". In: *IEEE Transactions on Biomedical Engineering* 64.10, pp. 2313–2320. ISSN: 0018-9294. DOI: 10.1109/TBME.2017.2667579.

McCrimmon, Colin M. et al. (2015). "Brain-controlled functional electrical stimulation therapy for gait rehabilitation after stroke: a safety study". In: *Journal of NeuroEngineering and Rehabilitation* 12, p. 57. ISSN: 1743-0003. DOI: 10.1186/s12984-015-0050-4. URL: https://doi.org/10.1186/s12984-015-0050-4.

Murphy, Kevin P. (2012). *Machine Learning: A Probabilistic Perspective*. The MIT Press. ISBN: 978-0-262-01802-9.

Myrden, A. and T. Chau (2017). "A Passive EEG-BCI for Single-Trial Detection of Changes in Mental State". In: *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 25.4, pp. 345–356. ISSN: 1534-4320. DOI: `10.1109/TNSRE.2016.2641956`.

Palaniappan, Ramaswamy (2005). "Brain computer interface design using band powers extracted during mental tasks". In: *Neural Engineering, 2005. Conference Proceedings. 2nd International IEEE EMBS Conference on*. IEEE, pp. 321–324.

Parini, Sergio et al. (2009). "A Robust and Self-Paced BCI System Based on a Four Class SSVEP Paradigm: Algorithms and Protocols for a High-Transfer-Rate Direct Brain Communication". en. In: *Computational Intelligence and Neuroscience* 2009, e864564. ISSN: 1687-5265. DOI: `10.1155/2009/864564`. URL: `https://www.hindawi.com/journals/cin/2009/864564/abs/` (visited on 06/09/2017).

Perelmouter, J. and N. Birbaumer (2000). "A binary spelling interface with random errors". eng. In: *IEEE transactions on rehabilitation engineering: a publication of the IEEE Engineering in Medicine and Biology Society* 8.2, pp. 227–232. ISSN: 1063-6528.

Rakotomamonjy, A. et al. (2005). "Ensemble of SVMs for Improving Brain Computer Interface P300 Speller Performances". en. In: *Artificial Neural Networks: Biological Inspirations – ICANN 2005*. Springer, Berlin, Heidelberg, pp. 45–50. DOI: `10.1007/11550822_8`. URL: `https://link-springer-com.ezproxy1.library.usyd.edu.au/chapter/10.1007/11550822_8` (visited on 06/11/2017).

Raudys, S. J. and A. K. Jain (1991). "Small sample size effects in statistical pattern recognition: recommendations for practitioners". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 13.3, pp. 252–264. ISSN: 0162-8828. DOI: `10.1109/34.75512`.

Rebsamen, B. et al. (2006). "A Brain-Controlled Wheelchair Based on P300 and Path Guidance". In: *The First IEEE/RAS-EMBS International Conference on Biomedical Robotics and Biomechatronics, 2006. BioRob 2006.* Pp. 1101–1106. DOI: `10.1109/BIOROB.2006.1639239`.

Sannelli, C. et al. (2008). "Estimating noise and dimensionality in BCI data sets: Towards BCI illiteracy comprehension". In: *4th International Brain-Computer Interface Workshop and Training Course 2008. Proceedings*, pp. 26–31. ISBN: 978-3-85125-004-6.

Sato, Hikaru and Yoshikazu Washizawa (2016). "An N100-P300 Spelling Brain-Computer Interface with Detection of Intentional Control". en. In: *Computers* 5.4, p. 31. DOI: `10.3390/computers5040031`. URL: `http://www.mdpi.com/2073-431X/5/4/31` (visited on 06/09/2017).

Siuly, Yan Li, and Peng (Paul) Wen (2011). "Clustering technique-based least square support vector machine for EEG signal classification". English. In: *Computer Methods and Programs in Biomedicine* 104.3, pp. 358–372. ISSN: 0169-2607, 1872-7565. DOI: `10.1016/j.`

cmpb.2010.11.014. URL: http://www.cmpbjournal.com/article/S0169-2607(10)00290-7/fulltext (visited on 06/03/2017).

Trafton, Anne (2014). "Brain can 'see' in the blink of an eye: Study". In: *The Economic Times*. URL: https://economictimes.indiatimes.com/brain-can-see-in-the-blink-of-an-eye-study/articleshow/28953796.cms.

Vourvopoulos, Athanasios and Fotis Liarokapis (2014). "Evaluation of Commercial Brain-computer Interfaces in Real and Virtual World Environment". In: *Comput. Electr. Eng.* 40.2, pp. 714–729. ISSN: 0045-7906. DOI: 10.1016/j.compeleceng.2013.10.009. URL: http://dx.doi.org/10.1016/j.compeleceng.2013.10.009.

Waikato, University of (2017). *Attribute-Relation File Format (ARFF)*. URL: https://www.cs.waikato.ac.nz/ml/weka/arff.html (visited on 10/04/2017).

World Economic Forum (2017). *YGL Moments | Young Global Leaders and Alumni Annual Summit 2017*. URL: https://www.youtube.com/watch?v=LaNdkBSOFKo.

Yoto, Ai et al. (2007). "Effects of object color stimuli on human brain activities in perception and attention referred to EEG alpha band response". eng. In: *Journal of Physiological Anthropology* 26.3, pp. 373–379. ISSN: 1880-6791.

Appendix Figures

.1 All Sensors

.1.1 Age

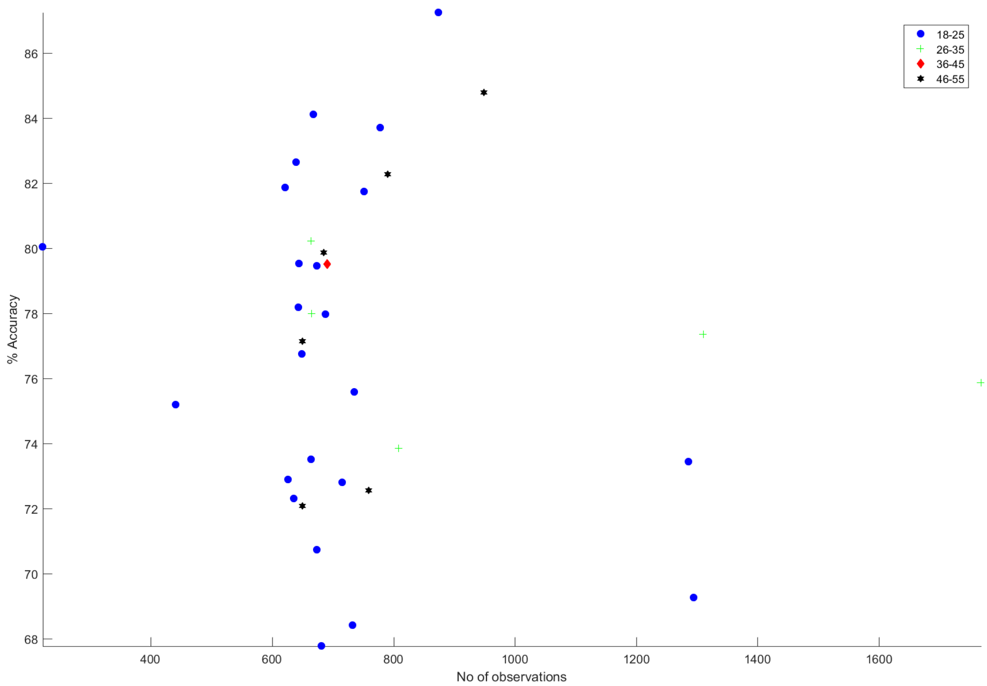


FIGURE 1: Blue versus Red Age Results - All Sensors

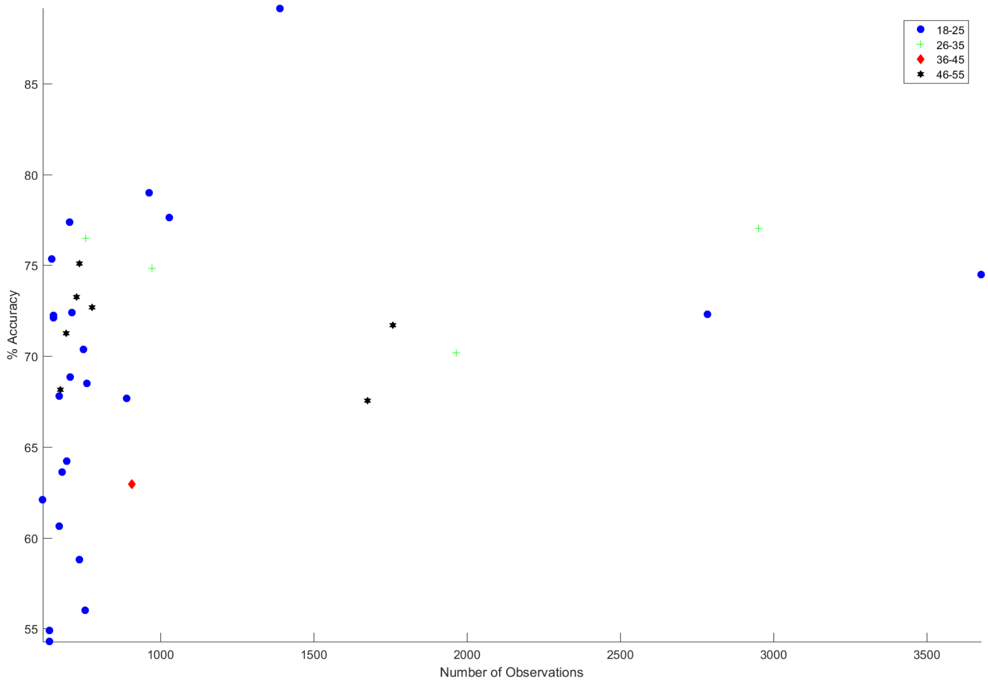


FIGURE 2: Yellow versus Colours Age Results - All Sensors

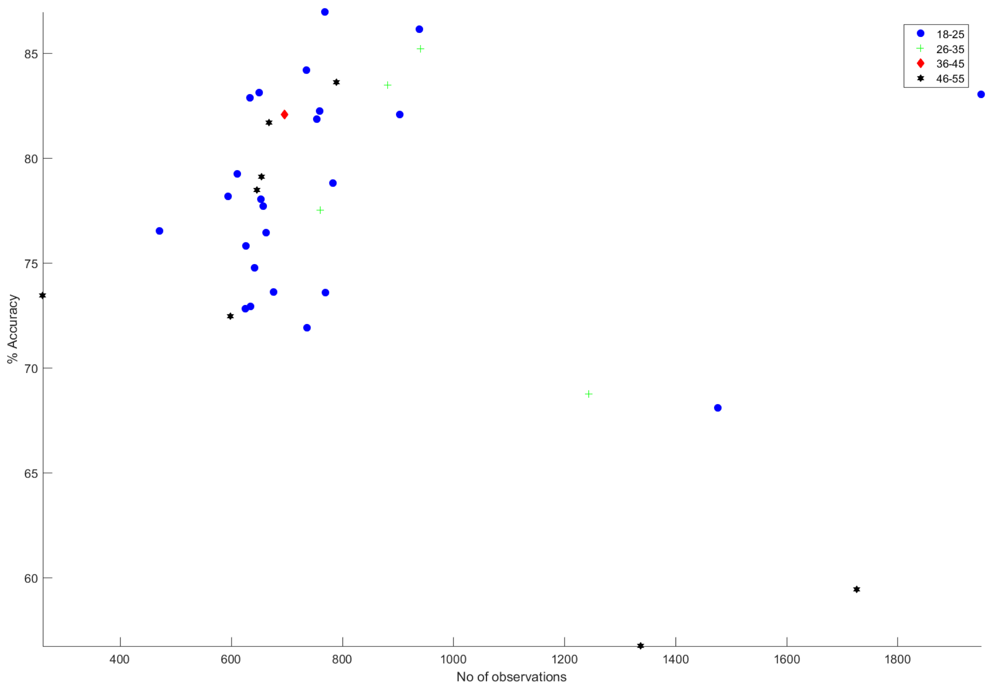


FIGURE 3: Yellow versus Red Age Results - All Sensors

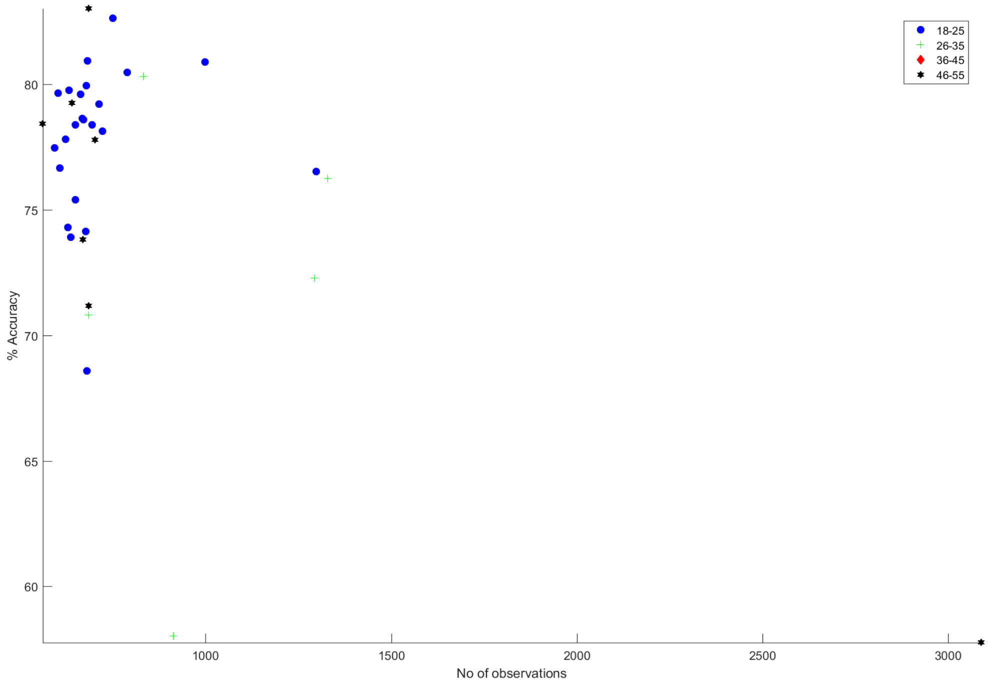


FIGURE 4: Yellow versus Red Audio Age Results - All Sensors

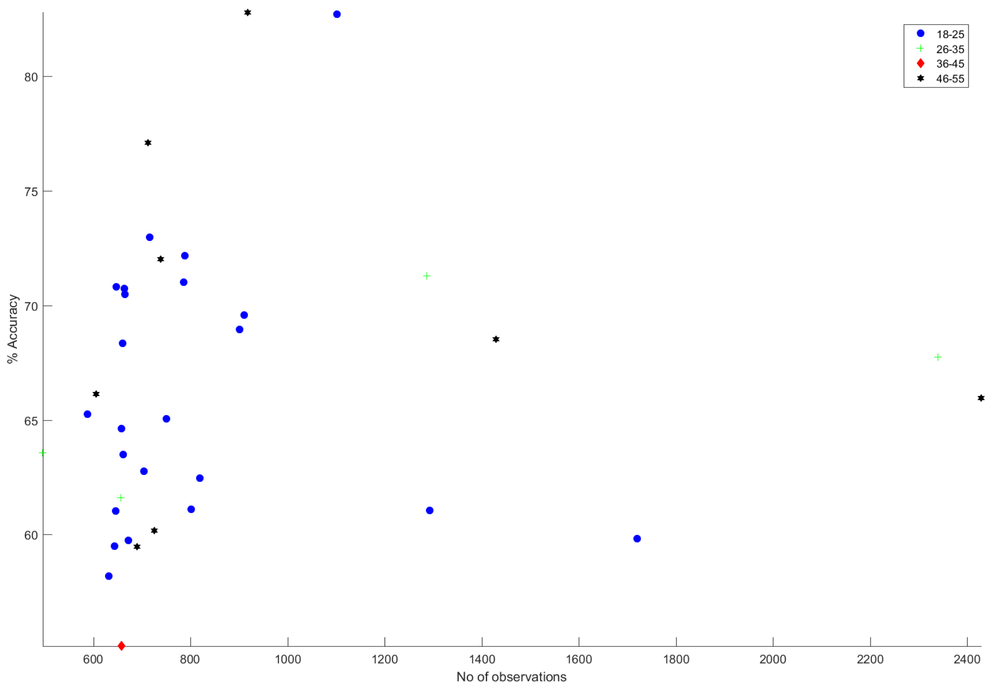


FIGURE 5: Yellow versus Visual Distractions Age Results - All Sensors

.1.2 Gender

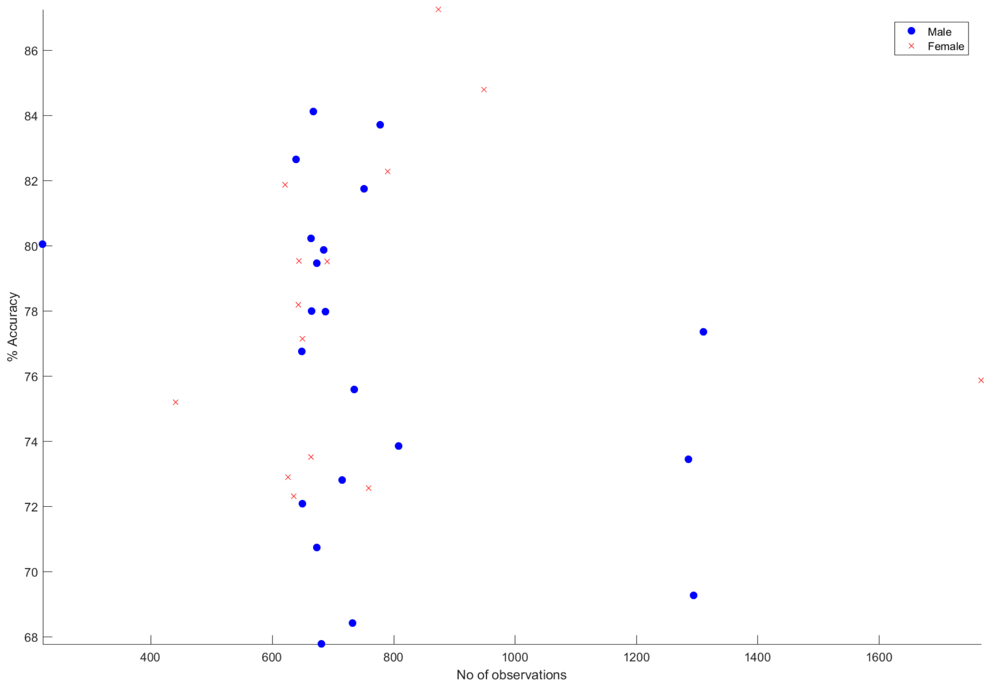


FIGURE 6: Blue versus Red Gender Results - All Sensors

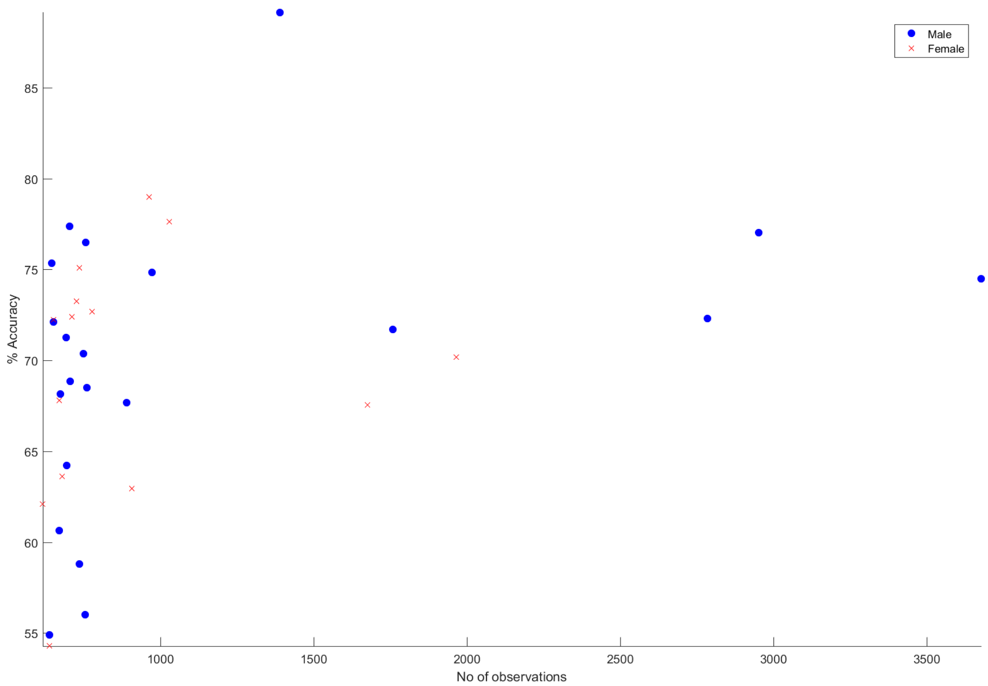


FIGURE 7: Yellow versus Colours Gender Results - All Sensors

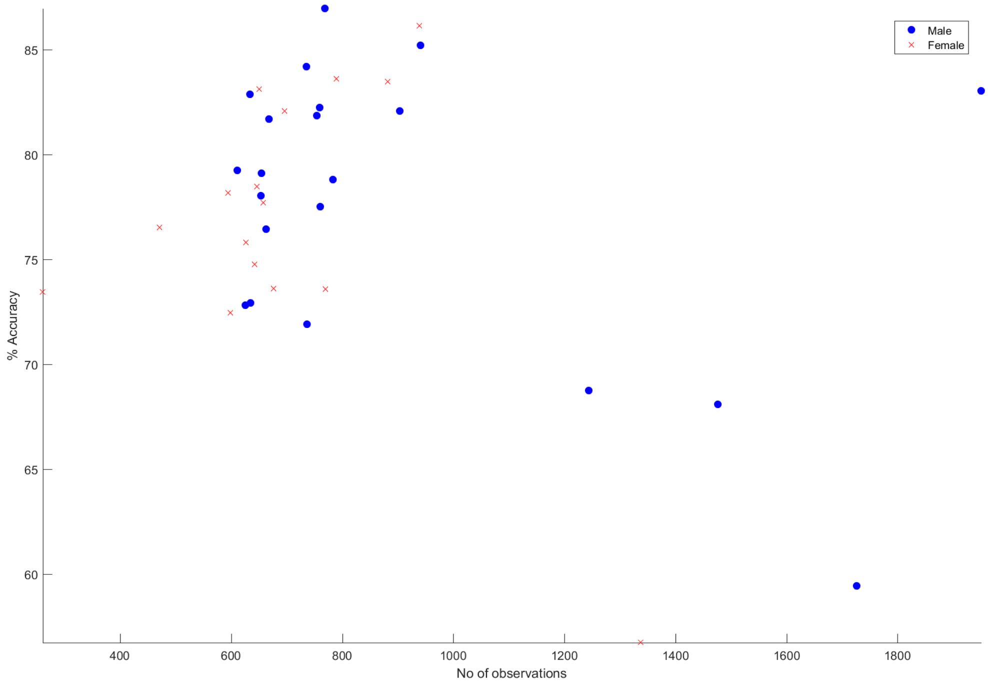


FIGURE 8: Yellow versus Red Gender Results - All Sensors

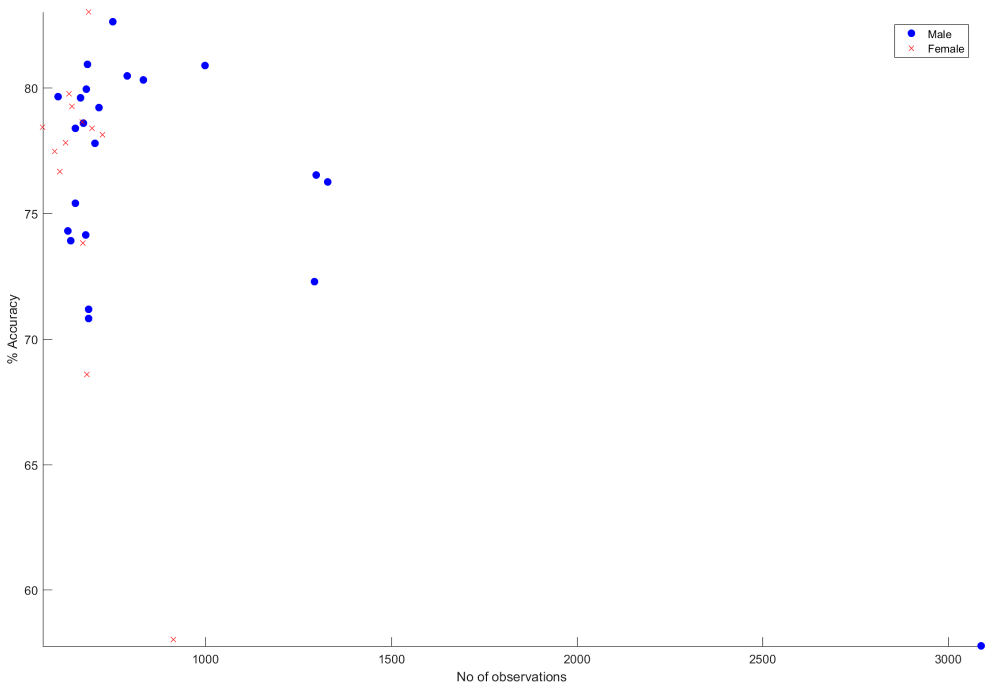


FIGURE 9: Yellow versus Red Audio Gender Results - All Sensors

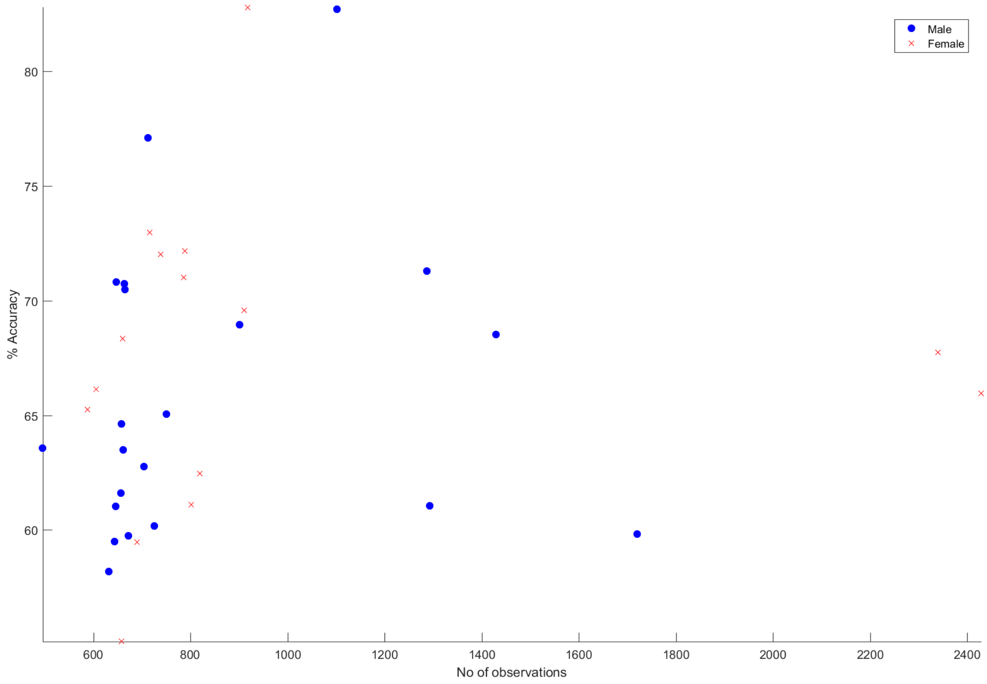


FIGURE 10: Yellow versus Visual Distractions Gender Results - All Sensors

.2 Frontal Lobe Sensor Data

.2.1 Age

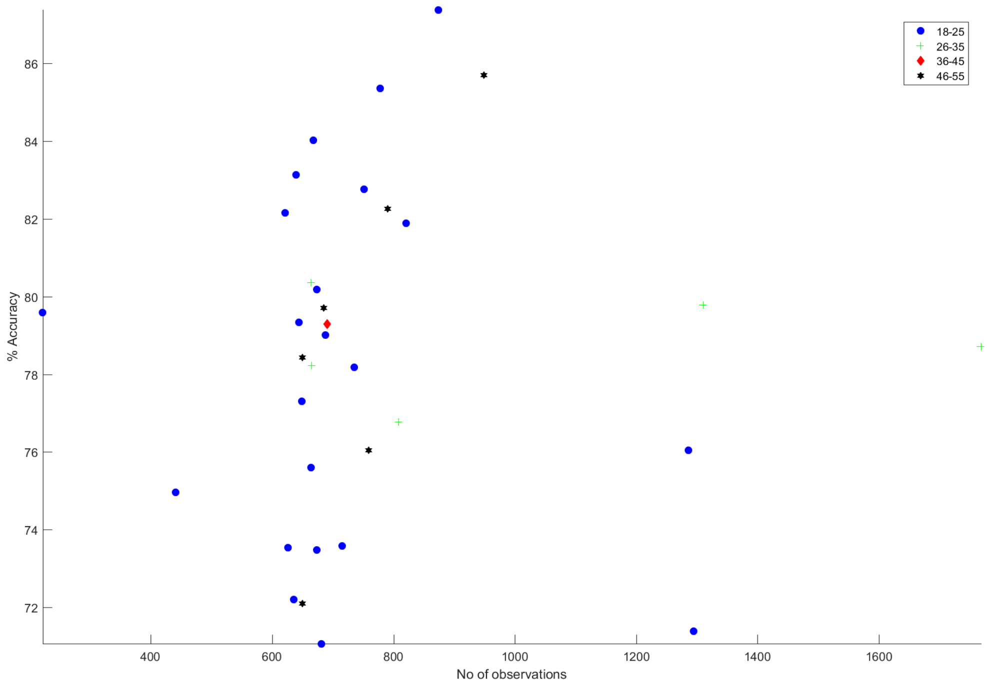


FIGURE 11: Blue versus Red Age Results - Frontal Lobes

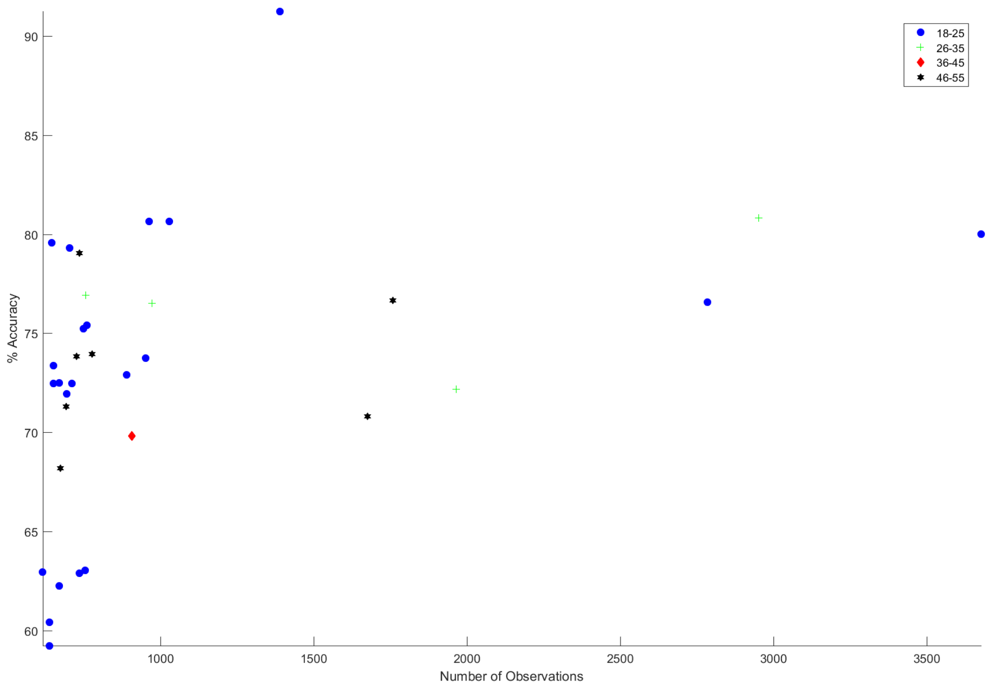


FIGURE 12: Yellow versus Colours Age Results - Frontal Lobes

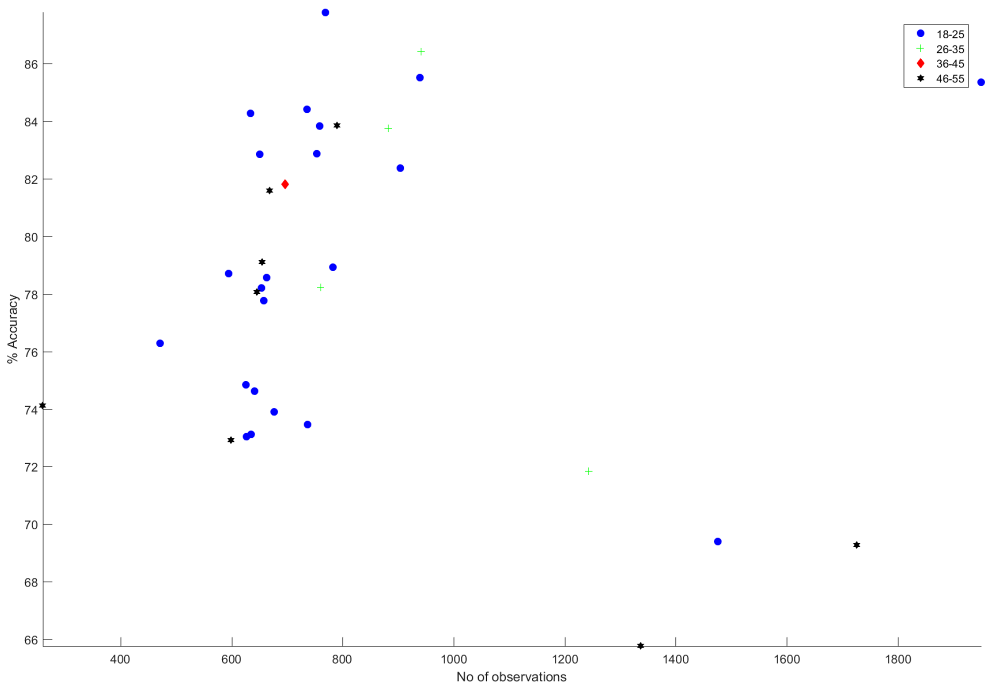


FIGURE 13: Yellow versus Red Age Results - Frontal Lobes

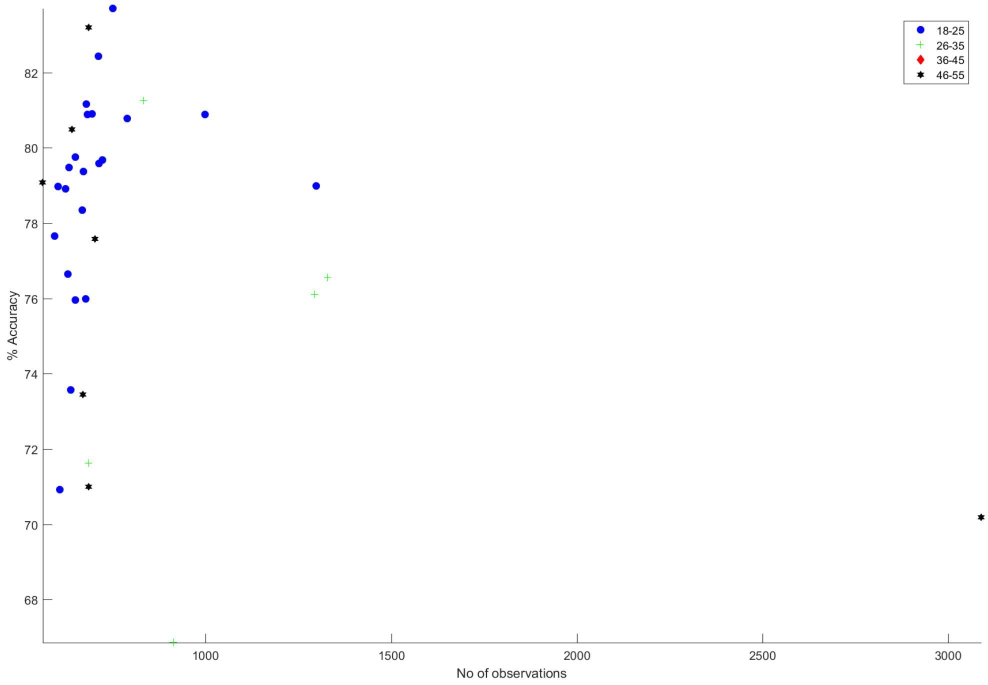


FIGURE 14: Yellow versus Red Audio Age Results - Frontal Lobes

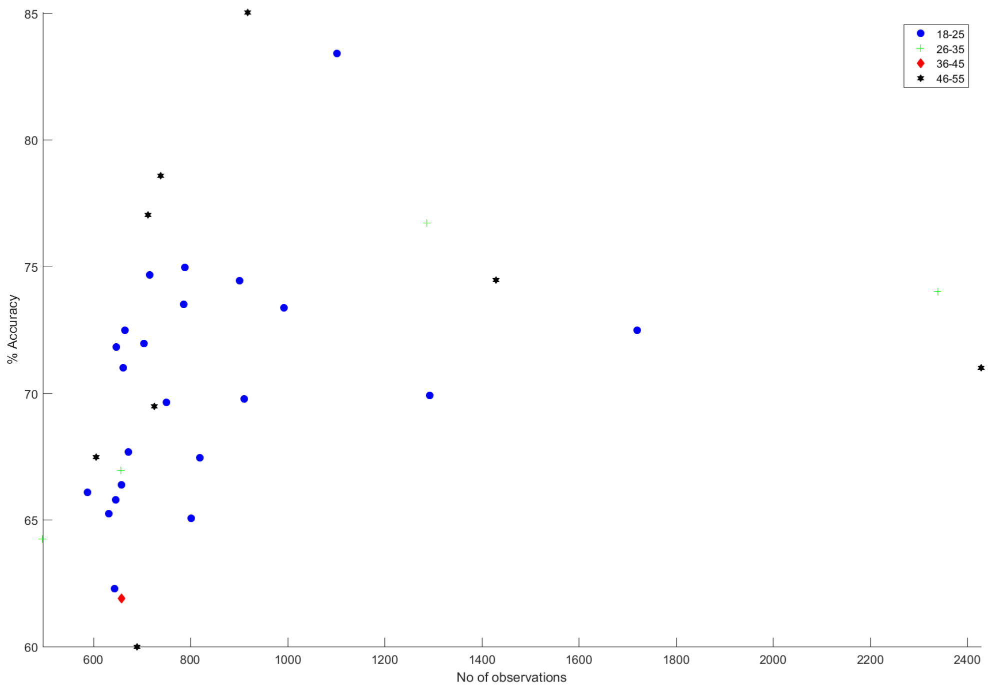


FIGURE 15: Yellow versus Visual Distractions Age Results - Frontal Lobes

.2.2 Gender

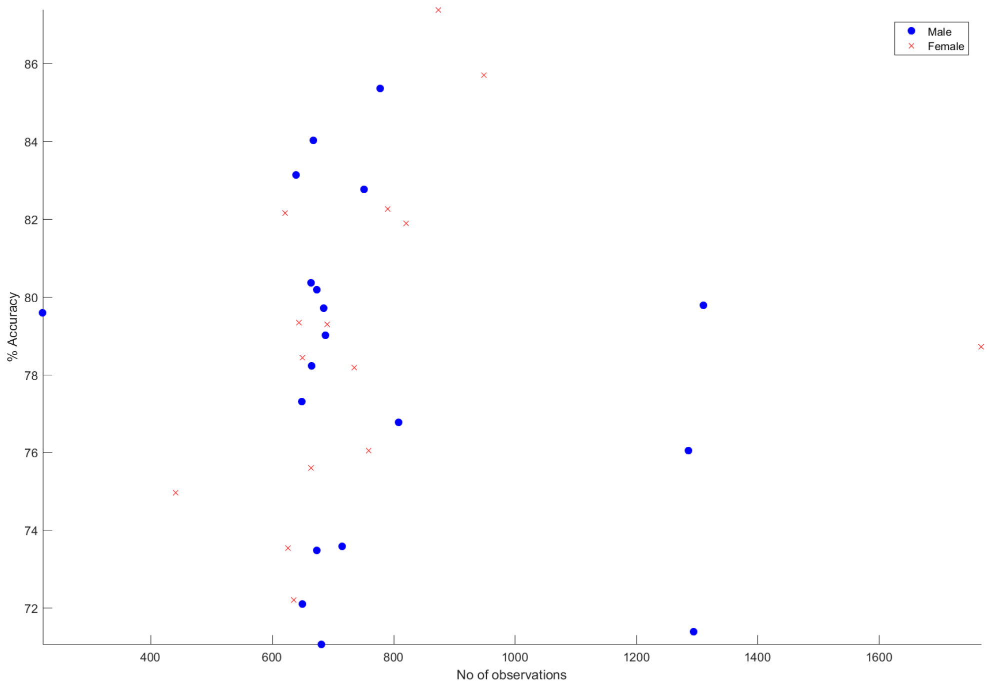


FIGURE 16: Blue versus Red Gender Results - Frontal Lobes

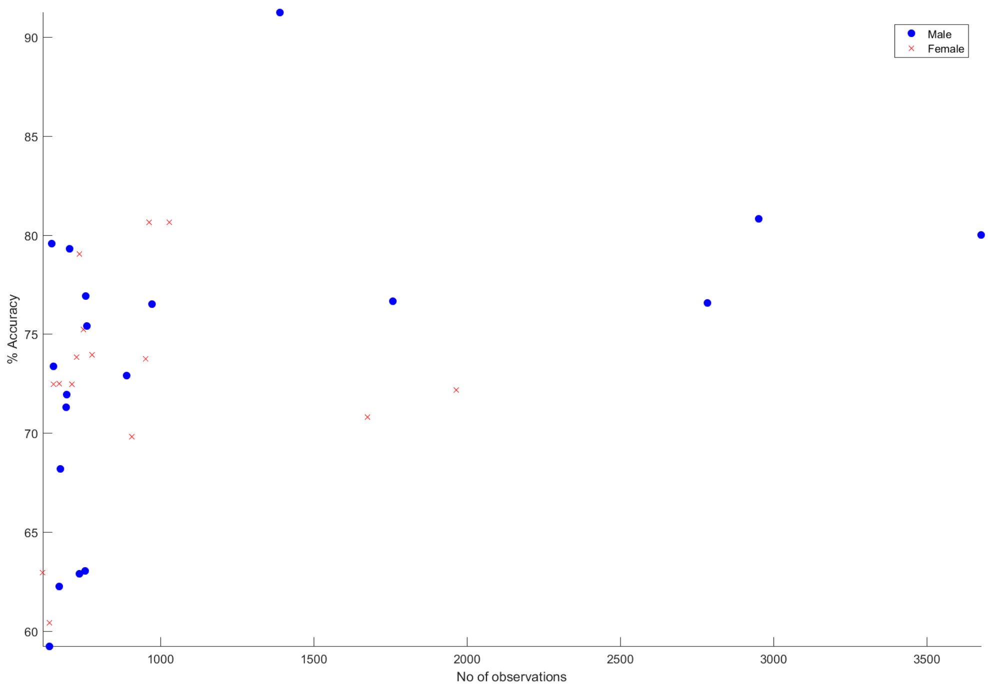


FIGURE 17: Yellow versus Colours Gender Results - Frontal Lobes

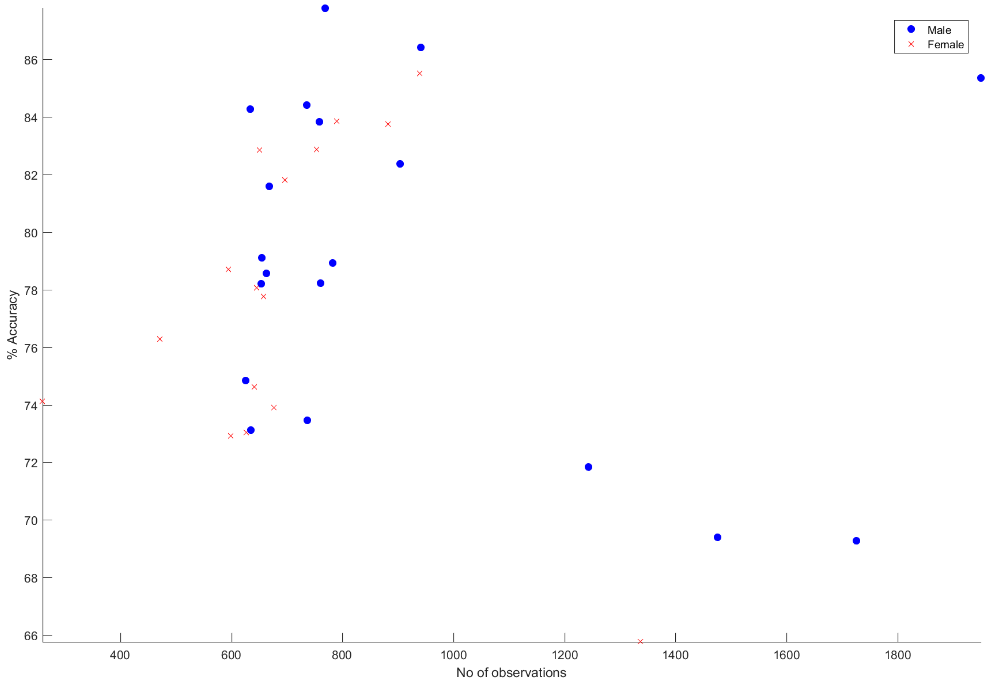


FIGURE 18: Yellow versus Red Gender Results - Frontal Lobes

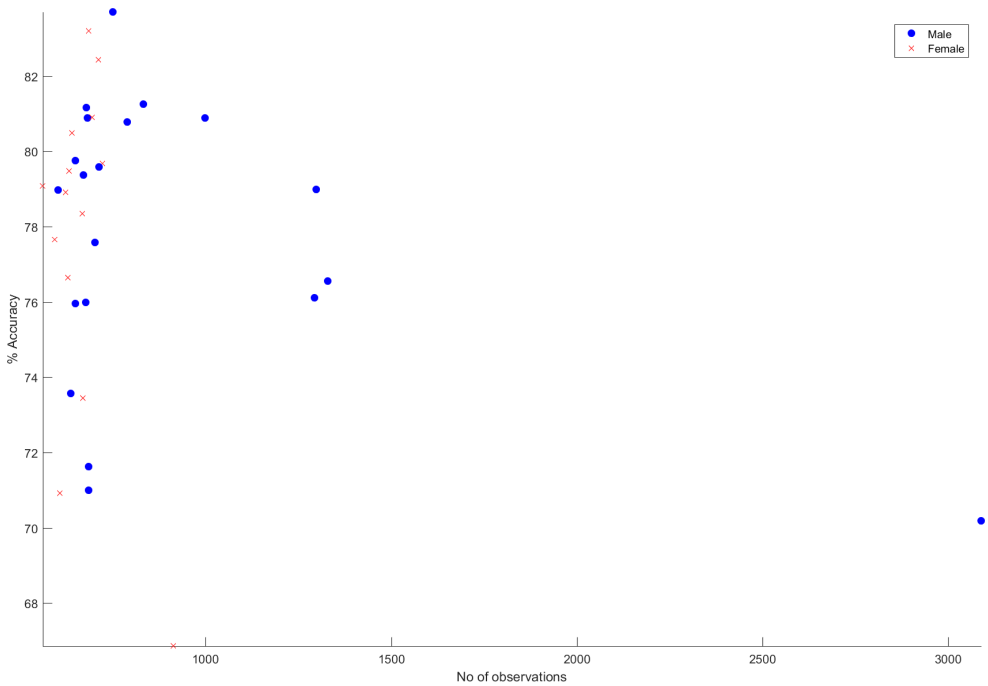


FIGURE 19: Yellow versus Red Audio Gender Results - Frontal Lobes

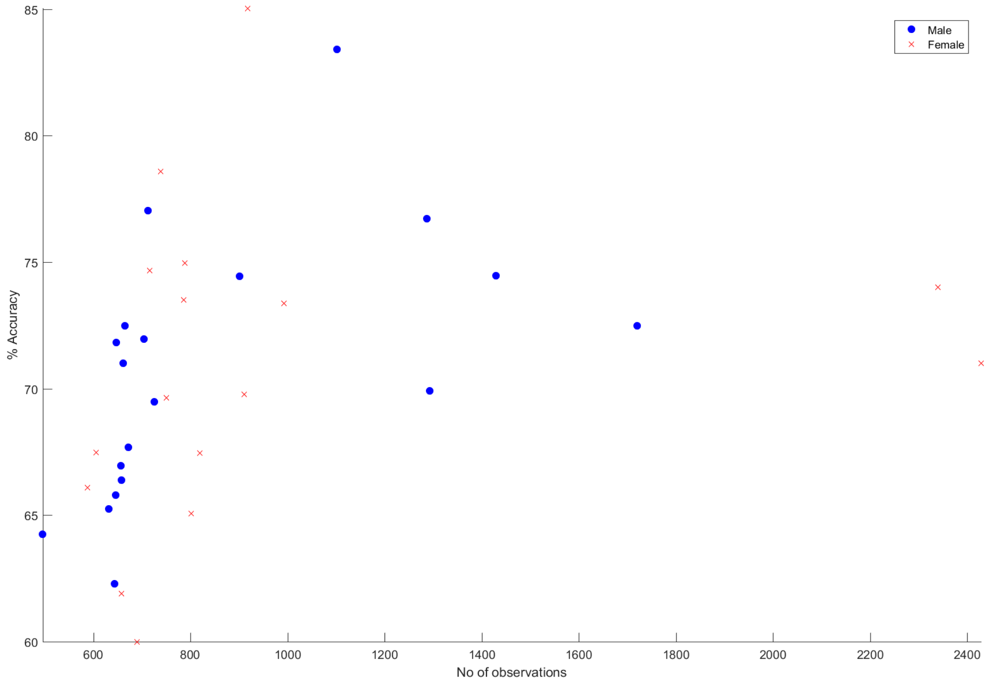


FIGURE 20: Yellow versus Visual Distractions Gender Results - Frontal Lobes

.3 Temporal Lobe Data

.3.1 Age

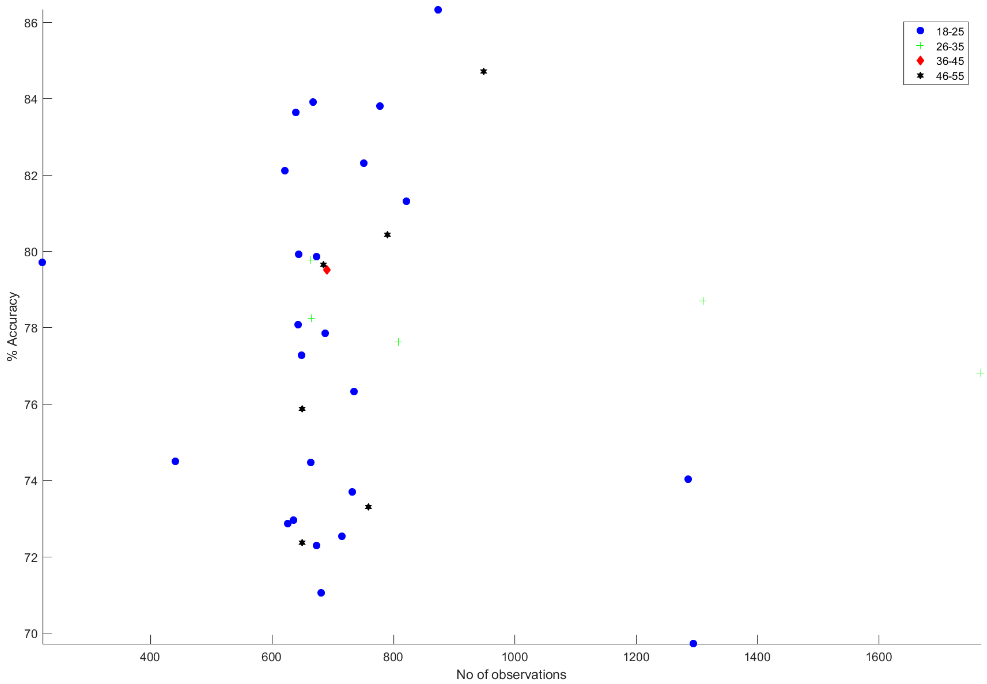


FIGURE 21: Blue versus Red Age Results - Temporal Lobes

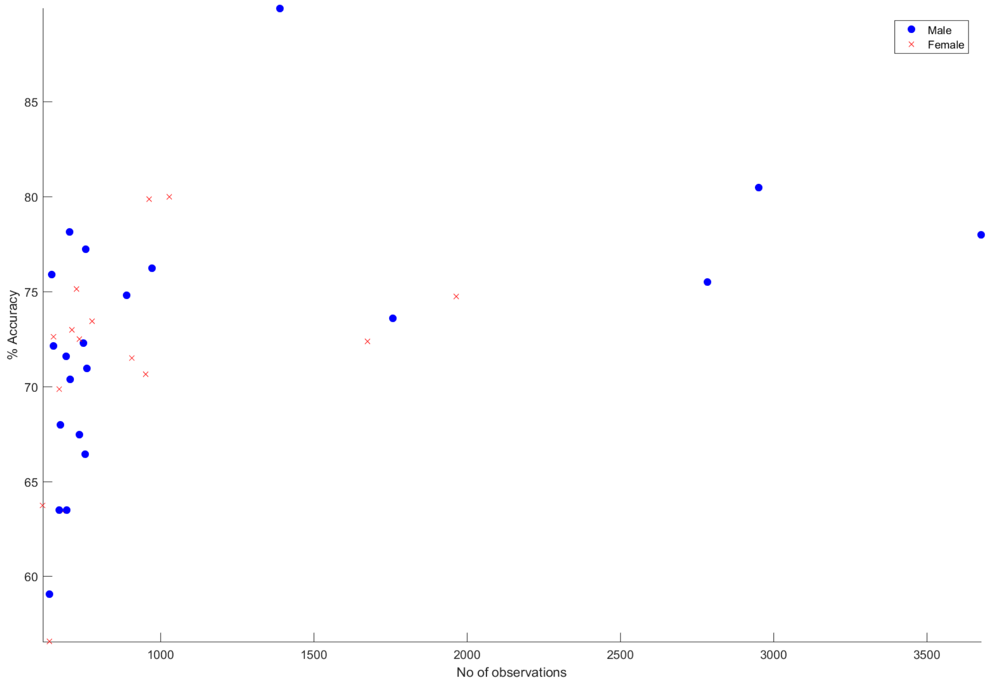


FIGURE 22: Yellow versus Colours Age Results - Temporal Lobes

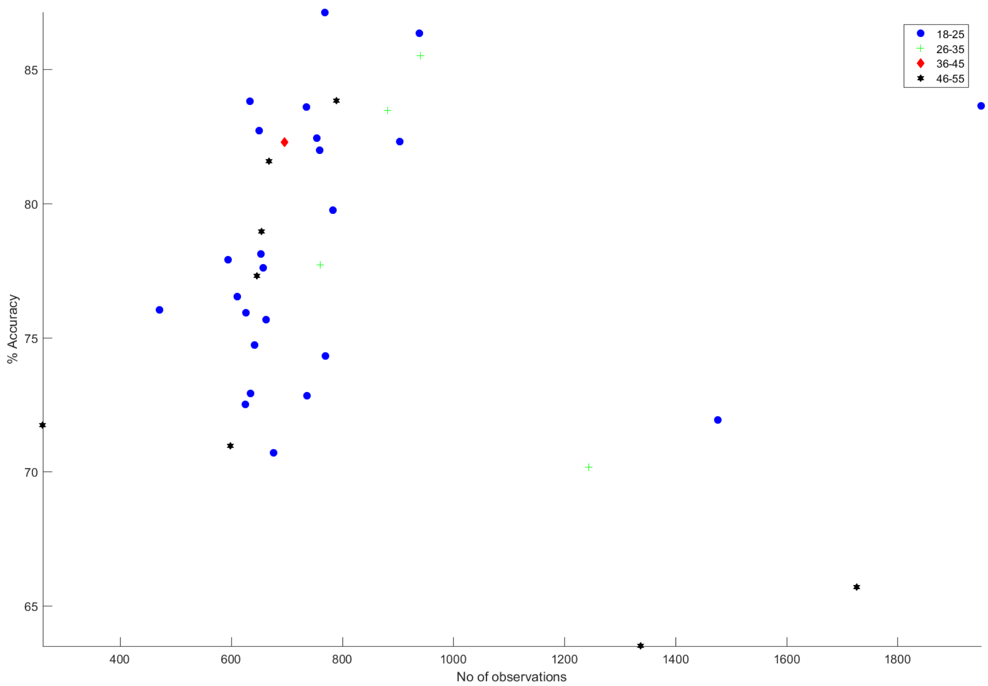


FIGURE 23: Yellow versus Red Age Results - Temporal Lobes

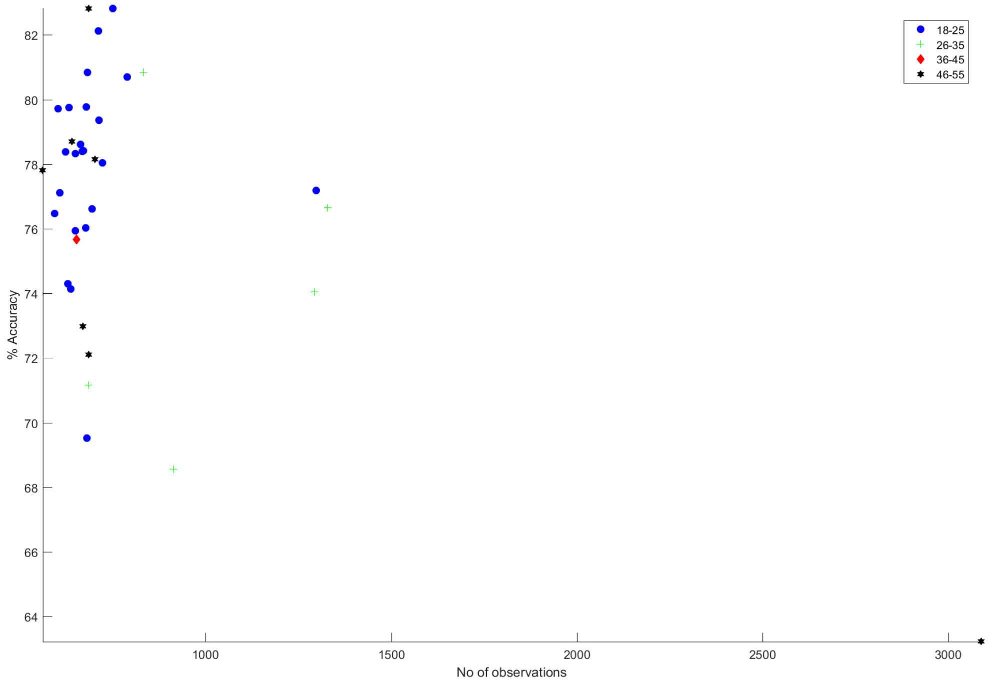


FIGURE 24: Yellow versus Red Audio Age Results - Temporal Lobes

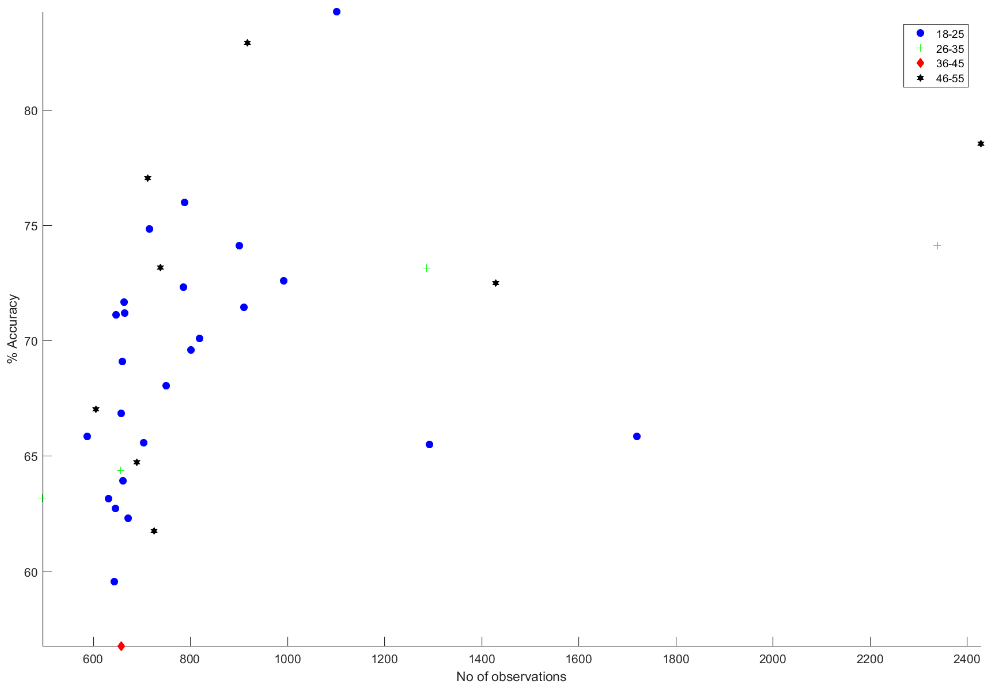


FIGURE 25: Yellow versus Visual Distractions Age Results - Temporal Lobes

.3.2 Gender

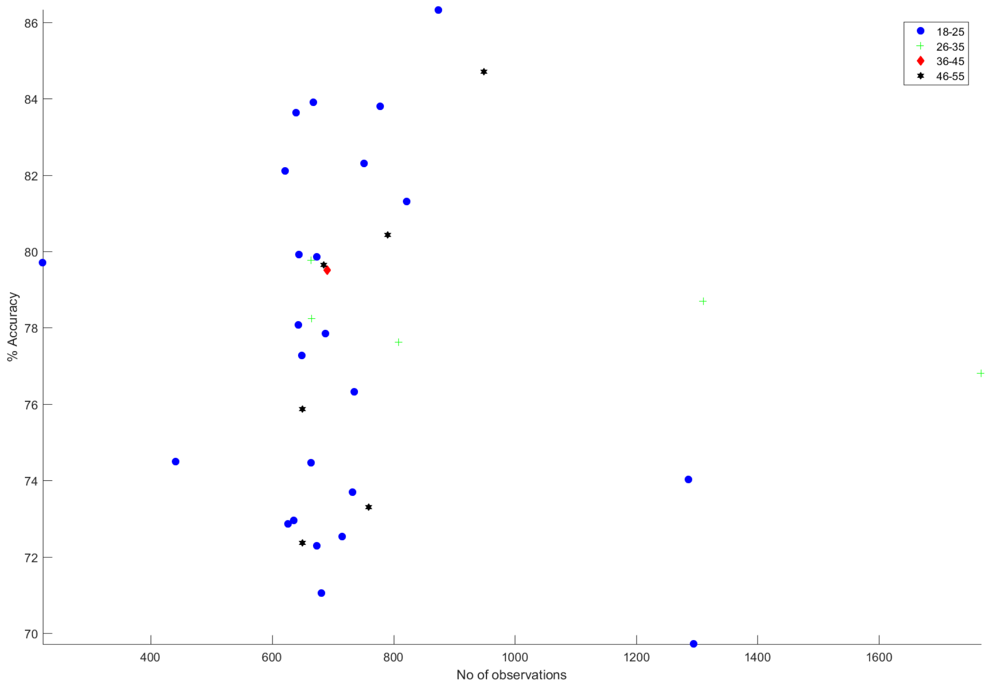


FIGURE 26: Blue versus Red Gender Results - Temporal Lobes

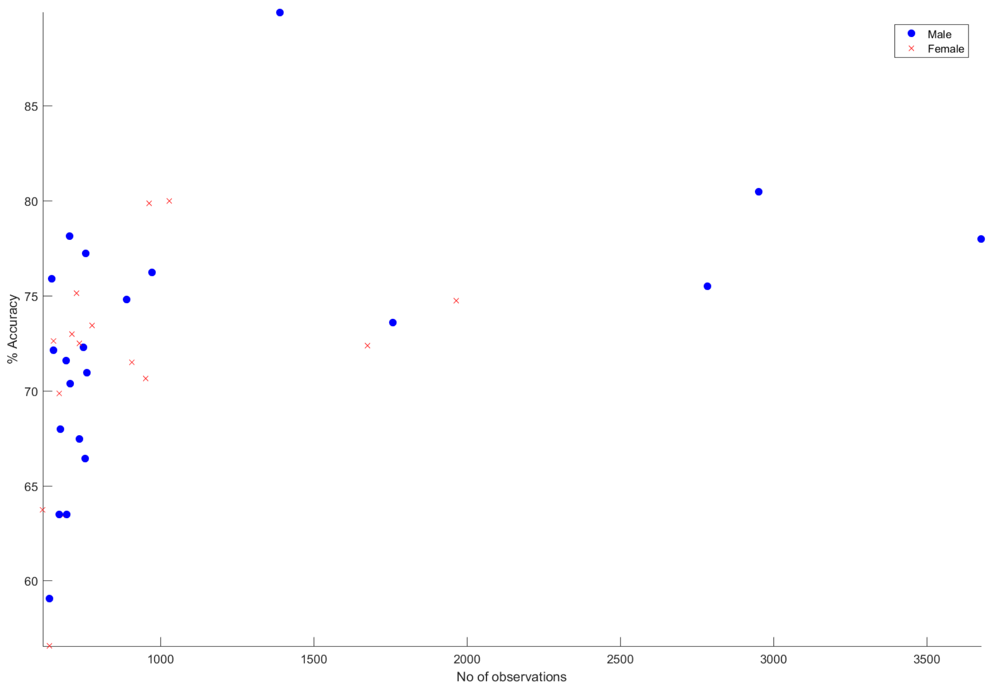


FIGURE 27: Yellow versus Colours Gender Results - Temporal Lobes

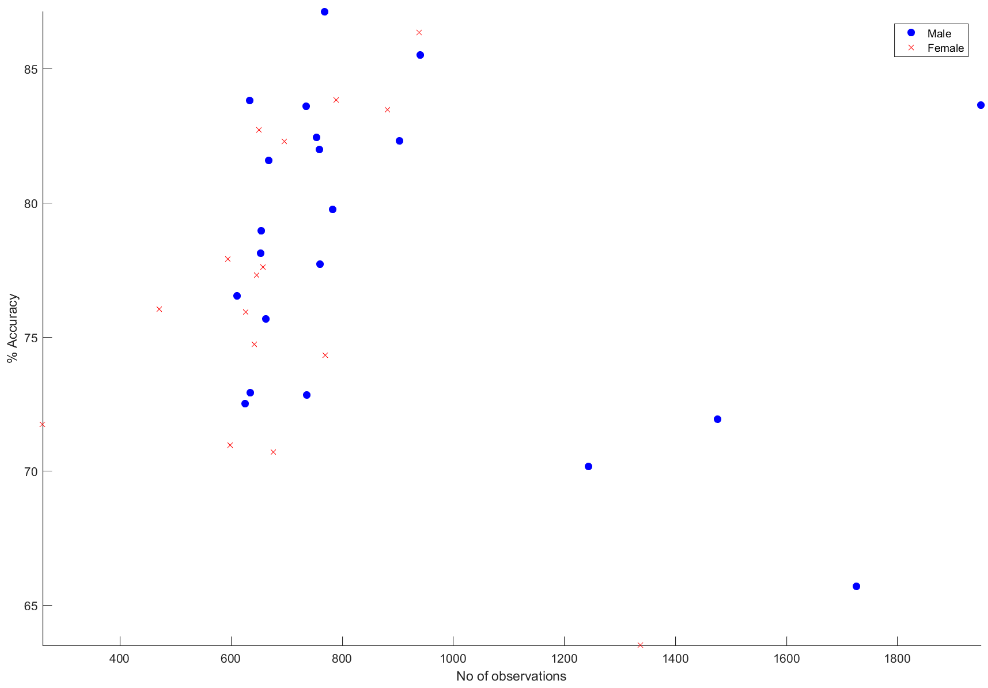


FIGURE 28: Yellow versus Red Gender Results - Temporal Lobes

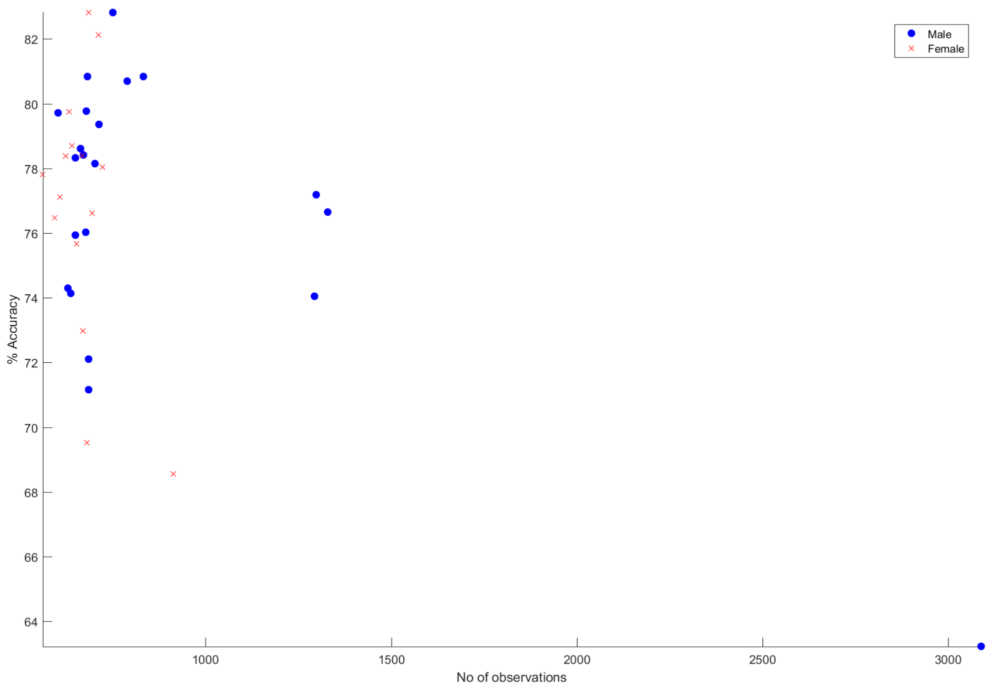


FIGURE 29: Yellow versus Red Audio Gender Results - Temporal Lobes

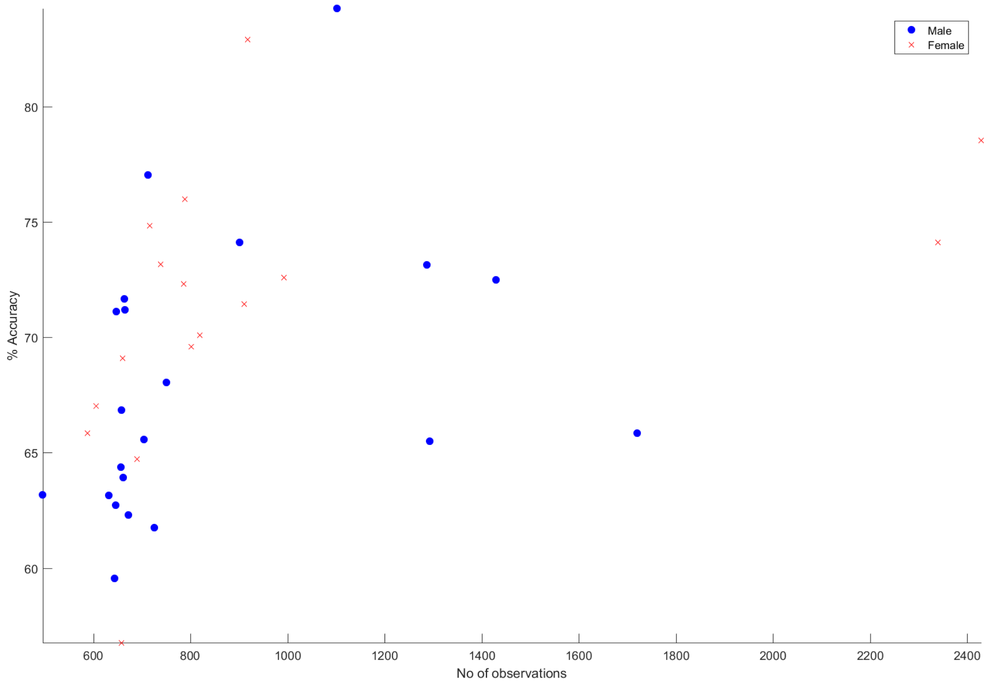


FIGURE 30: Yellow versus Visual Distractions Gender Results - Temporal Lobes

Machine Learning Parameters and Likelihood Tables

.4 Machine Parameters

.4.1 All Sensors

	Complexity Factor	Gamma Parameter	Exponential	Mean Accuracy	Average of Kernels	Variance * No of Subjects	S.D
PolyKernel	0.001	0	1	**81.19**	83.52	**2964.445**	8.718
	0.1	0	1	**81.29**		**1543.128**	6.290
	10	0	1	**81.32**		**1227.725**	5.611
	0.001	0	2	**81.19**		**2745.801**	8.391
	0.1	0	2	**81.54**		**1066.474**	5.229
	10	0	2	**87.46**		**1990.892**	7.145
	0.001	0	3	**81.31**		**1391.602**	5.973
	0.1	0	3	**84.46**		**1500.452**	6.203
	10	0	3	**91.92**		**2313.275**	7.702
RBFKernel	0.001	0.001	0	**81.19**	82.33	**2964.445**	8.718
	0.1	0.001	0	**81.19**		**2964.445**	8.718
	10	0.001	0	**81.19**		**2462.724**	7.946
	0.001	0.1	0	**81.19**		**2964.445**	8.718
	0.1	0.1	0	**81.19**		**2460.725**	7.943
	10	0.1	0	**81.71**		**843.269**	4.650
	0.001	10	0	**81.19**		**2964.445**	8.718
	0.1	10	0	**81.41**		**1285.199**	5.741
	10	10	0	**94.01**		**1620.847**	6.447
	0.001	1000	0	**81.19**		**2964.445**	8.718
	0.1	1000	0	**81.19**		**2964.445**	8.718
	10	1000	0	**81.30**		**2648.719**	8.241

FIGURE 31: Machine Parameters - Yellow Vs Red - All Sensors

	Complexity Factor	Gamma Parameter	Exponential	Mean Accuracy	Average of Kernels	Variance * No of Subjects	S.D
PolyKernel	0.001	0	1	68.82	78.20	2267.815	7.626
	0.1	0	1	70.32		1148.609	5.427
	10	0	1	75.94		887.719	4.771
	0.001	0	2	68.89		1948.807	7.069
	0.1	0	2	77.29		883.567	4.760
	10	0	2	90.66		529.773	3.686
	0.001	0	3	70.32		1200.707	5.549
	0.1	0	3	85.64		712.194	4.273
	10	0	3	95.95		839.680	4.640
RBFKernel	0.001	0.001	0	68.82	72.68	2267.815	7.626
	0.1	0.001	0	68.82		2267.815	7.626
	10	0.001	0	69.08		1883.260	6.949
	0.001	0.1	0	68.82		2267.815	7.626
	0.1	0.1	0	68.98		1904.563	6.988
	10	0.1	0	80.45		616.919	3.977
	0.001	10	0	68.82		2267.815	7.626
	0.1	10	0	73.41		2344.725	7.754
	10	10	0	97.64		815.273	4.572
	0.001	1000	0	68.82		2267.815	7.626
	0.1	1000	0	68.82		2240.255	7.579
	10	1000	0	69.65		2170.348	7.460

FIGURE 32: Machine Parameters - Yellow Vs Colours - All Sensors

	Complexity Factor	Gamma Parameter	Exponential	Mean Accuracy	Average of Kernels	Variance * No of Subjects	S.D
PolyKernel	0.001	0	1	59.39	80.67	2609.541	8.180
	0.1	0	1	75.84		2628.898	8.210
	10	0	1	80.99		1731.634	6.663
	0.001	0	2	70.49		2805.265	8.481
	0.1	0	2	82.52		1957.127	7.084
	10	0	2	94.45		1298.152	5.769
	0.001	0	3	77.19		3522.790	9.504
	0.1	0	3	89.09		1635.979	6.477
	10	0	3	96.11		864.214	4.707
RBFKernel	0.001	0.001	0	59.39	69.10	2609.541	8.180
	0.1	0.001	0	59.39		2609.541	8.180
	10	0.001	0	73.02		3838.814	9.921
	0.001	0.1	0	59.39		2609.541	8.180
	0.1	0.1	0	73.03		3951.699	10.066
	10	0.1	0	85.29		1933.599	7.041
	0.001	10	0	59.39		2609.541	8.180
	0.1	10	0	82.23		9422.147	15.543
	10	10	0	96.90		712.152	4.273
	0.001	1000	0	59.39		2609.541	8.180
	0.1	1000	0	59.39		2576.905	8.129
	10	1000	0	62.41		1862.575	6.911

FIGURE 33: Machine Parameters - Yellow Vs Visual Distractions - All Sensors

	Complexity Factor	Gamma Parameter	Exponential	Mean Accuracy	Average of Kernels	Variance * No of Subjects	S.D
PolyKernel	0.001	0	1	72.15	76.06	1518.405	6.240
	0.1	0	1	72.15		1541.685	6.287
	10	0	1	72.19		1597.582	6.400
	0.001	0	2	72.15		1489.884	6.181
	0.1	0	2	73.62		1150.812	5.432
	10	0	2	82.18		714.643	4.281
	0.001	0	3	72.38		1482.272	6.165
	0.1	0	3	78.39		532.085	3.694
	10	0	3	89.37		1256.669	5.676
RBFKernel	0.001	0.001	0	72.15	74.01	1518.405	6.240
	0.1	0.001	0	72.15		1518.405	6.240
	10	0.001	0	72.15		1489.884	6.181
	0.001	0.1	0	72.15		1518.405	6.240
	0.1	0.1	0	72.15		1489.884	6.181
	10	0.1	0	75.02		836.630	4.632
	0.001	10	0	72.15		1518.405	6.240
	0.1	10	0	72.15		1499.872	6.201
	10	10	0	91.41		650.293	4.083
	0.001	1000	0	72.15		1518.405	6.240
	0.1	1000	0	72.15		1518.405	6.240
	10	1000	0	72.40		1477.045	6.154

FIGURE 34: Machine Parameters - Blue Vs Red - All Sensors

	Complexity Factor	Gamma Parameter	Exponential	Mean Accuracy	Average of Kernels	Variance * No of Subjects	S.D
PolyKernel	0.001	0	1	70.68	77.45	3145.118	8.980
	0.1	0	1	73.01		1292.266	5.756
	10	0	1	74.46		753.485	4.395
	0.001	0	2	70.71		2666.392	8.269
	0.1	0	2	75.86		393.283	3.176
	10	0	2	85.88		1663.325	6.531
	0.001	0	3	73.68		1017.279	5.107
	0.1	0	3	81.61		658.397	4.109
	10	0	3	91.17		1833.503	6.857
RBFKernel	0.001	0.001	0	70.68	73.82	3145.118	8.980
	0.1	0.001	0	70.68		3145.118	8.980
	10	0.001	0	71.79		1901.024	6.982
	0.001	0.1	0	70.68		3145.118	8.980
	0.1	0.1	0	71.44		2082.928	7.308
	10	0.1	0	76.88		244.169	2.502
	0.001	10	0	70.68		3145.118	8.980
	0.1	10	0	75.61		529.780	3.686
	10	10	0	93.55		1045.410	5.177
	0.001	1000	0	70.68		3145.118	8.980
	0.1	1000	0	70.68		3145.118	8.980
	10	1000	0	72.44		2114.331	7.363

FIGURE 35: Machine Parameters -Yellow Vs Red w Audio - All Sensors

.4.2 Frontal Lobes

	Complexity Factor	Gamma Parameter	Exponential	Mean Accuracy	Average of Kernels	Variance * No of Subjects	S.D
PolyKernel	0.001	0	1	**81.19**	81.55	2896.276	8.618
	0.1	0	1	**81.23**		2145.694	7.417
	10	0	1	**81.26**		1678.826	6.561
	0.001	0	2	**81.19**		2896.276	8.618
	0.1	0	2	**81.27**		1591.047	6.387
	10	0	2	**81.47**		970.511	4.988
	0.001	0	3	**81.19**		2387.723	7.825
	0.1	0	3	**81.33**		1277.593	5.724
	10	0	3	**83.81**		1432.478	6.061
RBFKernel	0.001	0.001	0	**81.19**	82.42	2896.276	8.618
	0.1	0.001	0	**81.19**		2896.276	8.618
	10	0.001	0	**81.19**		2411.017	7.863
	0.001	0.1	0	**81.19**		2896.276	8.618
	0.1	0.1	0	**81.19**		2494.889	7.998
	10	0.1	0	**81.29**		1431.253	6.058
	0.001	10	0	**81.19**		2896.276	8.618
	0.1	10	0	**81.38**		1068.302	5.234
	10	10	0	**92.85**		1967.492	7.103
	0.001	1000	0	**81.19**		2896.276	8.618
	0.1	1000	0	**81.19**		2893.083	8.613
	10	1000	0	**84.03**		1912.412	7.003

FIGURE 36: Machine Parameters -Yellow Vs Red - Frontal Lobes

	Complexity Factor	Gamma Parameter	Exponential	Mean Accuracy	Average of Kernels	Variance * No of Subjects	S.D
PolyKernel	0.001	0	1	**68.82**	72.51	2215.227	7.537
	0.1	0	1	**69.23**		1789.333	6.774
	10	0	1	**72.15**		1072.494	5.244
	0.001	0	2	**68.82**		2215.227	7.537
	0.1	0	2	**70.91**		1349.598	5.883
	10	0	2	**77.80**		1130.344	5.384
	0.001	0	3	**68.82**		1975.999	7.118
	0.1	0	3	**73.35**		1406.283	6.005
	10	0	3	**82.74**		1393.795	5.978
RBFKernel	0.001	0.001	0	**68.82**	72.93	2215.227	7.537
	0.1	0.001	0	**68.82**		2215.227	7.537
	10	0.001	0	**68.82**		1915.400	7.008
	0.001	0.1	0	**68.82**		2215.227	7.537
	0.1	0.1	0	**68.82**		1916.143	7.009
	10	0.1	0	**73.58**		1219.965	5.593
	0.001	10	0	**68.82**		2215.227	7.537
	0.1	10	0	**77.50**		1819.383	6.830
	10	10	0	**96.18**		1541.314	6.287
	0.001	1000	0	**68.82**		2215.227	7.537
	0.1	1000	0	**68.82**		2002.774	7.166
	10	1000	0	**77.31**		2404.134	7.851

FIGURE 37: Machine Parameters -Yellow Vs Colours - Frontal Lobes

	Complexity Factor	Gamma Parameter	Exponential	Mean Accuracy	Average of Kernels	Variance * No of Subjects	S.D
PolyKernel	0.001	0	1	59.39	72.64	2644.311	8.234
	0.1	0	1	73.02		3530.448	9.514
	10	0	1	74.99		2212.173	7.531
	0.001	0	2	61.08		2154.642	7.433
	0.1	0	2	75.08		2675.212	8.282
	10	0	2	80.93		2022.527	7.201
	0.001	0	3	68.02		2053.793	7.257
	0.1	0	3	76.53		2449.751	7.926
	10	0	3	84.71		1750.507	6.700
RBFKernel	0.001	0.001	0	59.39	69.46	2644.311	8.234
	0.1	0.001	0	59.39		2644.311	8.234
	10	0.001	0	71.61		3670.424	9.701
	0.001	0.1	0	59.39		2644.311	8.234
	0.1	0.1	0	71.30		3453.606	9.410
	10	0.1	0	75.89		1865.893	6.917
	0.001	10	0	59.39		2644.311	8.234
	0.1	10	0	82.77		6212.391	12.621
	10	10	0	95.71		1714.036	6.629
	0.001	1000	0	59.39		2644.311	8.234
	0.1	1000	0	59.40		2644.771	8.235
	10	1000	0	79.95		3246.424	9.124

FIGURE 38: Machine Parameters -Yellow Vs Visual Distractions - Frontal Lobes

	Complexity Factor	Gamma Parameter	Exponential	Mean Accuracy	Average of Kernels	Variance * No of Subjects	S.D
PolyKernel	0.001	0	1	72.15	73.56	1515.675	6.234
	0.1	0	1	72.15		1526.320	6.256
	10	0	1	72.15		1530.536	6.265
	0.001	0	2	72.15		1515.675	6.234
	0.1	0	2	72.15		1561.878	6.328
	10	0	2	76.62		610.168	3.955
	0.001	0	3	72.15		1537.593	6.279
	0.1	0	3	73.99		1053.165	5.197
	10	0	3	78.57		662.291	4.121
RBFKernel	0.001	0.001	0	72.15	74.03	1515.675	6.234
	0.1	0.001	0	72.15		1515.675	6.234
	10	0.001	0	72.15		1478.932	6.158
	0.001	0.1	0	72.15		1515.675	6.234
	0.1	0.1	0	72.15		1478.231	6.157
	10	0.1	0	72.17		1568.540	6.342
	0.001	10	0	72.15		1515.675	6.234
	0.1	10	0	72.38		1441.697	6.080
	10	10	0	90.50		1495.992	6.193
	0.001	1000	0	72.15		1515.675	6.234
	0.1	1000	0	72.15		1515.675	6.234
	10	1000	0	76.19		965.752	4.976

FIGURE 39: Machine Parameters Frontal Lobes

	Complexity Factor	Gamma Parameter	Exponential	Mean Accuracy	Average of Kernels	Variance * No of Subjects	S.D
PolyKernel	0.001	0	1	**70.68**	74.77	**3139.386**	8.972
	0.1	0	1	**72.38**		**1792.434**	6.779
	10	0	1	**73.82**		**1227.558**	5.610
	0.001	0	2	**70.69**		**2908.763**	8.636
	0.1	0	2	**73.90**		**1052.881**	5.196
	10	0	2	**80.05**		**948.541**	4.932
	0.001	0	3	**71.88**		**2054.988**	7.259
	0.1	0	3	**76.99**		**796.826**	4.520
	10	0	3	**82.55**		**1513.074**	6.229
RBFKernel	0.001	0.001	0	**70.68**	73.93	**3139.386**	8.972
	0.1	0.001	0	**70.68**		**3139.386**	8.972
	10	0.001	0	**71.22**		**2475.368**	7.967
	0.001	0.1	0	**70.68**		**3139.386**	8.972
	0.1	0.1	0	**70.94**		**2649.906**	8.243
	10	0.1	0	**75.27**		**629.090**	4.016
	0.001	10	0	**70.68**		**3139.386**	8.972
	0.1	10	0	**75.64**		**463.275**	3.447
	10	10	0	**92.75**		**2414.764**	7.869
	0.001	1000	0	**70.68**		**3139.386**	8.972
	0.1	1000	0	**70.68**		**3139.386**	8.972
	10	1000	0	**77.25**		**1568.286**	6.341

FIGURE 40: Machine Parameters Frontal Lobes

.4.3 Temporal Lobes

	Complexity Factor	Gamma Parameter	Exponential	Mean Accuracy	Average of Kernels	Variance * No of Subjects	S.D
PolyKernel	0.001	0	1	**81.19**	81.58	**2964.45**	8.718453
	0.1	0	1	**81.23**		**2304.44**	7.686878
	10	0	1	**81.27**		**1780.82**	6.757376
	0.001	0	2	**81.19**		**2964.45**	8.718453
	0.1	0	2	**81.28**		**1877.84**	6.939011
	10	0	2	**81.76**		**937.76**	4.903596
	0.001	0	3	**81.20**		**2916.56**	8.647743
	0.1	0	3	**81.64**		**1506.40**	6.214962
	10	0	3	**83.42**		**1060.47**	5.214539
RBFKernel	0.001	0.001	0	**81.19**	82.17	**2964.45**	8.718453
	0.1	0.001	0	**81.19**		**2964.45**	8.718453
	10	0.001	0	**81.19**		**2964.45**	8.718453
	0.001	0.1	0	**81.19**		**2964.45**	8.718453
	0.1	0.1	0	**81.19**		**2964.45**	8.718453
	10	0.1	0	**81.31**		**1414.58**	6.02257
	0.001	10	0	**81.19**		**2964.45**	8.718453
	0.1	10	0	**81.42**		**968.22**	4.982572
	10	10	0	**91.42**		**2208.60**	7.525347
	0.001	1000	0	**81.19**		**2964.45**	8.718453
	0.1	1000	0	**81.19**		**2940.94**	8.68382
	10	1000	0	**82.32**		**1931.13**	7.036776

FIGURE 41: Machine Parameters -Yellow Vs Red - Temporal Lobes

	Complexity Factor	Gamma Parameter	Exponential	Mean Accuracy	Average of Kernels	Variance * No of Subjects	S.D
PolyKernel	0.001	0	1	**68.82**	72.27	**2267.82**	7.625556
	0.1	0	1	**69.69**		**1297.43**	5.76779
	10	0	1	**70.11**		**1128.70**	5.379696
	0.001	0	2	**68.82**		**2067.95**	7.281778
	0.1	0	2	**69.97**		**1020.32**	5.114898
	10	0	2	**78.18**		**658.63**	4.109499
	0.001	0	3	**69.08**		**1757.58**	6.713142
	0.1	0	3	**73.03**		**1056.04**	5.203646
	10	0	3	**82.76**		**625.93**	4.006179
RBFKernel	0.001	0.001	0	**68.82**	72.88	**2267.82**	7.625556
	0.1	0.001	0	**68.82**		**2267.82**	7.625556
	10	0.001	0	**68.86**		**1963.66**	7.095786
	0.001	0.1	0	**68.82**		**2267.82**	7.625556
	0.1	0.1	0	**68.81**		**1981.03**	7.127102
	10	0.1	0	**72.38**		**981.65**	5.017024
	0.001	10	0	**68.82**		**2267.82**	7.625556
	0.1	10	0	**79.02**		**2264.43**	7.61986
	10	10	0	**95.72**		**1254.22**	5.670925
	0.001	1000	0	**68.82**		**2267.82**	7.625556
	0.1	1000	0	**68.92**		**2237.74**	7.574825
	10	1000	0	**76.74**		**2226.23**	7.555321

FIGURE 42: Machine Parameters -Yellow Vs Colours - Temporal Lobes

	Complexity Factor	Gamma Parameter	Exponential	Mean Accuracy	Average of Kernels	Variance * No of Subjects	S.D
PolyKernel	0.001	0	1	**59.39**	68.77	**2609.54**	8.179934
	0.1	0	1	**67.49**		**1519.25**	6.241402
	10	0	1	**68.22**		**1284.58**	5.739153
	0.001	0	2	**59.42**		**2570.27**	8.118146
	0.1	0	2	**67.41**		**1329.41**	5.838441
	10	0	2	**79.07**		**527.15**	3.676496
	0.001	0	3	**63.21**		**1729.28**	6.658863
	0.1	0	3	**70.88**		**936.49**	4.900272
	10	0	3	**83.87**		**520.19**	3.652154
RBFKernel	0.001	0.001	0	**59.39**	67.22	**2609.54**	8.179934
	0.1	0.001	0	**59.39**		**2609.54**	8.179934
	10	0.001	0	**63.25**		**1786.99**	6.769068
	0.001	0.1	0	**59.39**		**2609.54**	8.179934
	0.1	0.1	0	**63.16**		**1825.19**	6.841029
	10	0.1	0	**71.88**		**784.61**	4.48534
	0.001	10	0	**59.39**		**2609.54**	8.179934
	0.1	10	0	**83.84**		**6194.56**	12.60297
	10	10	0	**95.10**		**1332.91**	5.846119
	0.001	1000	0	**59.39**		**2609.54**	8.179934
	0.1	1000	0	**59.45**		**2571.04**	8.119373
	10	1000	0	**72.99**		**2129.78**	7.389843

FIGURE 43: Machine Parameters -Yellow Vs Visual Distractions - Temporal Lobes

	Complexity Factor	Gamma Parameter	Exponential	Mean Accuracy	Average of Kernels	Variance * No of Subjects	S.D
PolyKernel	0.001	0	1	72.15	72.97	1516.72	6.236203
	0.1	0	1	72.15		1524.16	6.251475
	10	0	1	72.15		1496.72	6.194942
	0.001	0	2	72.15		1516.72	6.236203
	0.1	0	2	72.31		1446.24	6.089592
	10	0	2	73.38		1272.83	5.71286
	0.001	0	3	72.15		1515.67	6.234038
	0.1	0	3	72.76		1364.17	5.914277
	10	0	3	77.52		486.77	3.532871
RBFKernel	0.001	0.001	0	72.15	73.68	1516.72	6.236203
	0.1	0.001	0	72.15		1516.72	6.236203
	10	0.001	0	72.15		1480.38	6.16104
	0.001	0.1	0	72.15		1516.72	6.236203
	0.1	0.1	0	72.15		1490.84	6.182779
	10	0.1	0	72.70		1369.96	5.926826
	0.001	10	0	72.15		1516.72	6.236203
	0.1	10	0	72.44		1476.15	6.152242
	10	10	0	88.17		843.04	4.649357
	0.001	1000	0	72.15		1516.72	6.236203
	0.1	1000	0	72.15		1490.63	6.182332
	10	1000	0	73.65		1147.00	5.423134

FIGURE 44: Machine Parameters Temporal Lobes

	Complexity Factor	Gamma Parameter	Exponential	Mean Accuracy	Average of Kernels	Variance * No of Subjects	S.D
PolyKernel	0.001	0	1	69.62	73.43	2700.42	8.321153
	0.1	0	1	73.07		1010.41	5.089991
	10	0	1	74.01		712.05	4.272894
	0.001	0	2	69.62		2654.46	8.250032
	0.1	0	2	74.09		708.79	4.263116
	10	0	2	77.21		265.29	2.608108
	0.001	0	3	69.95		2319.10	7.711299
	0.1	0	3	74.39		598.54	3.917552
	10	0	3	78.94		246.40	2.513558
RBFKernel	0.001	0.001	0	69.62	72.89	2700.42	8.321153
	0.1	0.001	0	69.62		2700.42	8.321153
	10	0.001	0	69.62		2455.55	7.934915
	0.001	0.1	0	69.62		2700.42	8.321153
	0.1	0.1	0	69.62		2459.03	7.940525
	10	0.1	0	74.20		615.53	3.972753
	0.001	10	0	69.62		2700.42	8.321153
	0.1	10	0	75.70		376.51	3.107103
	10	10	0	89.82		1049.96	5.188649
	0.001	1000	0	69.62		2700.42	8.321153
	0.1	1000	0	69.62		2700.42	8.321153
	10	1000	0	77.97		465.83	3.45606

FIGURE 45: Machine Parameters Temporal Lobes

.5 Likelihood Tables

.5.1 All Sensors

Likelihood > 75.00	Likelihood > 80.00	Likelihood > 85.00
4.490	0.863	-2.764
6.320	1.293	-3.734
7.129	1.493	-4.143
4.665	0.897	-2.872
7.905	1.858	-4.189
11.034	6.608	2.182
6.683	1.390	-3.904
9.649	4.551	-0.548
13.892	9.786	5.680
4.490	0.863	-2.764
4.490	0.863	-2.764
4.926	0.947	-3.033
4.490	0.863	-2.764
4.928	0.947	-3.034
9.124	2.323	-4.477
4.490	0.863	-2.764
7.067	1.558	-3.951
18.646	13.741	8.835
4.490	0.863	-2.764
4.490	0.863	-2.764
4.834	0.996	-2.841
21	**21**	**4**

FIGURE 46: Probability - Yellow V Red - All Sensors (at 95% Confidence)

Likelihood > 75.00	Likelihood > 80.00	Likelihood > 85.00
-5.125	-9.272	-13.419
-5.459	-11.286	-17.113
1.241	-5.387	-12.016
-5.469	-9.943	-14.416
3.041	-3.603	-10.246
26.880	18.300	9.720
-5.334	-11.033	-16.733
15.747	8.347	0.947
28.549	21.734	14.918
-5.125	-9.272	-13.419
-5.125	-9.272	-13.419
-5.385	-9.936	-14.487
-5.125	-9.272	-13.419
-5.448	-9.973	-14.499
8.668	0.717	-7.234
-5.125	-9.272	-13.419
-1.300	-5.379	-9.457
31.312	24.395	17.479
-5.125	-9.272	-13.419
-5.156	-9.329	-13.501
-4.539	-8.778	-13.017
8	**5**	**4**

FIGURE 47: Probability - Yellow V Colours - All Sensors (at 95% Confidence)

Likelihood > 75.00	Likelihood > 80.00	Likelihood > 85.00
-12.068	-15.934	-19.800
0.643	-3.208	-7.060
5.689	0.943	-3.803
-3.360	-7.088	-10.817
6.716	2.252	-2.212
21.322	15.841	10.360
1.457	-1.871	-5.198
13.754	8.872	3.989
28.358	21.640	14.922
-12.068	-15.934	-19.800
-12.068	-15.934	-19.800
-1.264	-4.452	-7.639
-12.068	-15.934	-19.800
-1.238	-4.379	-7.521
9.240	4.749	0.258
-12.068	-15.934	-19.800
2.940	0.905	-1.129
32.419	25.019	17.618
-12.068	-15.934	-19.800
-12.145	-16.035	-19.925
-11.518	-16.094	-20.670
12	**8**	**6**

FIGURE 48: Probability - Yellow V Visual Distractions - All Sensors (at 95% Confidence)

Likelihood > 75.00	Likelihood > 80.00	Likelihood > 85.00
-2.891	-7.959	-13.027
-2.867	-7.897	-12.927
-2.780	-7.721	-12.662
-2.918	-8.034	-13.151
-1.612	-7.434	-13.255
10.604	3.217	-4.170
-2.691	-7.821	-12.950
5.805	-2.756	-11.318
16.016	10.445	4.874
-2.891	-7.959	-13.027
-2.891	-7.959	-13.027
-2.918	-8.034	-13.151
-2.891	-7.959	-13.027
-2.918	-8.034	-13.151
0.026	-6.802	-13.629
-2.891	-7.959	-13.027
-2.908	-8.008	-13.107
25.423	17.679	9.935
-2.891	-7.959	-13.027
-2.891	-7.959	-13.027
-2.677	-7.815	-12.954
6	**3**	**2**

FIGURE 49: Probability - Blue v Red - All Sensors (at 95% Confidence)

Likelihood > 75.00	Likelihood > 80.00	Likelihood > 85.00
-3.039	-6.561	-10.082
-2.192	-7.685	-13.179
-0.772	-7.966	-15.161
-3.285	-7.110	-10.934
1.711	-8.247	-18.206
10.533	5.691	0.849
-1.631	-7.823	-14.014
10.174	2.477	-5.219
14.916	10.303	5.691
-3.039	-6.561	-10.082
-3.039	-6.561	-10.082
-2.905	-7.434	-11.963
-3.039	-6.561	-10.082
-3.077	-7.404	-11.731
4.755	-7.883	-20.522
-3.039	-6.561	-10.082
1.053	-7.527	-16.107
22.654	16.546	10.439
-3.039	-6.561	-10.082
-3.039	-6.561	-10.082
-2.201	-6.496	-10.791
9	**4**	**3**

FIGURE 50: Probability - Yellow V Red w Audio - All Sensors (at 95% Confidence)

.5.2 Frontal Lobes

Likelihood > 75.00	Likelihood > 80.00	Likelihood > 85.00
4.543	0.873	-2.796
5.308	1.045	-3.219
6.033	1.213	-3.606
4.543	0.873	-2.796
6.208	1.257	-3.694
8.199	1.860	-4.479
5.005	0.963	-3.078
6.994	1.469	-4.056
9.190	3.973	-1.245
4.543	0.873	-2.796
4.543	0.873	-2.796
4.979	0.957	-3.065
4.543	0.873	-2.796
4.894	0.941	-3.013
6.570	1.350	-3.870
4.543	0.873	-2.796
7.713	1.671	-4.371
15.891	11.439	6.987
4.543	0.873	-2.796
4.545	0.874	-2.798
8.155	3.640	-0.876
21	**21**	**3**

FIGURE 51: Probability - Yellow V Red - Frontal Lobes (at 95% Confidence)

Likelihood > 75.00	Likelihood > 80.00	Likelihood > 85.00
-5.185	-9.381	-13.577
-5.392	-10.060	-14.729
-3.439	-9.470	-15.500
-5.185	-9.381	-13.577
-4.400	-9.776	-15.152
3.286	-2.588	-8.462
-5.489	-9.931	-14.374
-1.738	-7.004	-12.270
8.185	2.896	-2.394
-5.185	-9.381	-13.577
-5.185	-9.381	-13.577
-5.576	-10.088	-14.600
-5.185	-9.381	-13.577
-5.575	-10.086	-14.598
-1.610	-7.264	-12.918
-5.185	-9.381	-13.577
2.317	-2.313	-6.942
21.311	16.281	11.251
-5.185	-9.381	-13.577
-5.453	-9.866	-14.279
1.864	-2.164	-6.191
6	**2**	**1**

FIGURE 52: Probability - Yellow V Colours - Frontal Lobes (at 95% Confidence)

Likelihood > 75.00	Likelihood > 80.00	Likelihood > 85.00
-11.989	-15.829	-19.670
-1.318	-4.642	-7.966
-0.009	-4.207	-8.406
-11.847	-16.101	-20.356
0.062	-3.757	-7.575
5.207	0.816	-3.575
-6.086	-10.444	-14.802
1.223	-2.767	-6.757
9.170	4.450	-0.270
-11.989	-15.829	-19.670
-11.989	-15.829	-19.670
-2.212	-5.471	-8.731
-11.989	-15.829	-19.670
-2.489	-5.849	-9.209
0.809	-3.762	-8.334
-11.989	-15.829	-19.670
3.892	1.387	-1.119
19.753	14.983	10.213
-11.989	-15.829	-19.670
-11.984	-15.824	-19.664
3.431	-0.035	-3.501
10	**5**	**3**

FIGURE 53: Probability - Yellow V Visual Distractions - Frontal Lobes (at 95% Confidence)

Likelihood > 75.00	Likelihood > 80.00	Likelihood > 85.00
-2.893	-7.966	-13.038
-2.883	-7.938	-12.993
-2.879	-7.927	-12.974
-2.893	-7.966	-13.038
-2.847	-7.844	-12.841
2.592	-5.403	-13.398
-2.872	-7.908	-12.945
-1.229	-7.315	-13.400
5.484	-2.190	-9.863
-2.893	-7.966	-13.038
-2.893	-7.966	-13.038
-2.929	-8.064	-13.199
-2.893	-7.966	-13.038
-2.930	-8.066	-13.202
-2.827	-7.813	-12.800
-2.893	-7.966	-13.038
-2.730	-7.932	-13.133
15.831	10.725	5.620
-2.893	-7.966	-13.038
-2.893	-7.966	-13.038
1.508	-4.847	-11.202
5	**1**	**1**

FIGURE 54: Probability - Blue v Red - Frontal Lobes (at 95% Confidence)

Likelihood > 75.00	Likelihood > 80.00	Likelihood > 85.00
-3.042	-6.567	-10.091
-2.448	-7.113	-11.778
-1.330	-6.967	-12.603
-3.155	-6.817	-10.479
-1.337	-7.423	-13.509
6.477	0.065	-6.347
-2.722	-7.078	-11.435
2.778	-4.218	-11.214
7.670	2.593	-2.484
-3.042	-6.567	-10.091
-3.042	-6.567	-10.091
-3.002	-6.971	-10.940
-3.042	-6.567	-10.091
-3.117	-6.953	-10.789
0.429	-7.445	-15.318
-3.042	-6.567	-10.091
1.176	-7.999	-17.174
14.270	10.252	6.233
-3.042	-6.567	-10.091
-3.042	-6.567	-10.091
2.242	-2.745	-7.732
9	**3**	**1**

FIGURE 55: Probability - Yellow V Red w Audio - Frontal Lobes (at 95% Confidence)

.5.3 Temporal Lobes

Likelihood > 75.00	Likelihood > 80.00	Likelihood > 85.00
4.490	0.863	-2.764
5.123	1.010	-3.104
5.871	1.191	-3.488
4.490	0.863	-2.764
5.724	1.166	-3.391
8.720	2.271	-4.178
4.532	0.875	-2.782
6.762	1.674	-3.414
10.216	4.152	-1.913
4.490	0.863	-2.764
4.490	0.863	-2.764
4.490	0.863	-2.764
4.490	0.863	-2.764
4.490	0.863	-2.764
6.630	1.379	-3.871
4.490	0.863	-2.764
8.148	1.801	-4.546
13.799	9.597	5.395
4.490	0.863	-2.764
4.508	0.866	-2.775
6.577	2.083	-2.411
21	**21**	**1**

FIGURE 56: Probability - Yellow V Red - Temporal Lobes (at 95% Confidence)

Likelihood > 75.00	Likelihood > 80.00	Likelihood > 85.00
-5.125	-9.272	-13.419
-5.824	-11.307	-16.789
-5.746	-11.624	-17.502
-5.367	-9.709	-14.052
-6.217	-12.400	-18.582
4.896	-2.799	-10.494
-5.580	-10.291	-15.001
-2.394	-8.471	-14.548
12.254	4.360	-3.533
-5.125	-9.272	-13.419
-5.125	-9.272	-13.419
-5.476	-9.933	-14.389
-5.125	-9.272	-13.419
-5.495	-9.932	-14.369
-3.300	-9.603	-15.906
-5.125	-9.272	-13.419
3.333	-0.817	-4.967
23.107	17.531	11.954
-5.125	-9.272	-13.419
-5.075	-9.250	-13.424
1.460	-2.726	-6.911
5	**3**	**1**

FIGURE 57: Probability - Yellow V Colours - Temporal Lobes (at 95% Confidence)

Likelihood > 75.00	Likelihood > 80.00	Likelihood > 85.00
-12.068	-15.934	-19.800
-7.615	-12.681	-17.748
-7.467	-12.977	-18.487
-12.135	-16.030	-19.925
-8.219	-13.636	-19.052
7.006	-1.595	-10.197
-11.202	-15.951	-20.700
-5.319	-11.772	-18.226
15.360	6.702	-1.957
-12.068	-15.934	-19.800
-12.068	-15.934	-19.800
-10.982	-15.654	-20.325
-12.068	-15.934	-19.800
-10.948	-15.571	-20.193
-4.400	-11.450	-18.501
-12.068	-15.934	-19.800
4.439	1.929	-0.580
21.750	16.340	10.931
-12.068	-15.934	-19.800
-12.110	-16.004	-19.899
-1.724	-6.003	-10.282
4	**4**	**2**

FIGURE 58: Probability - Yellow V Visual Distractions - Temporal Lobes (at 95% Confidence)

Likelihood > 75.00	Likelihood > 80.00	Likelihood > 85.00
-2.892	-7.963	-13.034
-2.885	-7.943	-13.002
-2.909	-8.014	-13.118
-2.892	-7.963	-13.034
-2.798	-7.991	-13.184
-1.797	-7.332	-12.867
-2.893	-7.966	-13.038
-2.397	-7.744	-13.091
4.520	-4.431	-13.382
-2.892	-7.963	-13.034
-2.892	-7.963	-13.034
-2.927	-8.060	-13.193
-2.892	-7.963	-13.034
-2.917	-8.032	-13.146
-2.452	-7.788	-13.123
-2.892	-7.963	-13.034
-2.636	-7.776	-12.917
17.921	11.119	4.318
-2.892	-7.963	-13.034
-2.917	-8.032	-13.147
-1.569	-7.400	-13.231
3	**1**	**1**

FIGURE 59: Probability - Blue v Red - Temporal Lobes (at 95% Confidence)

Likelihood > 75.00	Likelihood > 80.00	Likelihood > 85.00
-4.086	-7.886	-11.686
-2.392	-8.605	-14.818
-1.472	-8.873	-16.273
-4.121	-7.954	-11.787
-1.353	-8.771	-16.189
5.361	-6.764	-18.889
-4.139	-8.240	-12.341
-0.985	-9.057	-17.129
9.924	-2.657	-15.238
-4.086	-7.886	-11.686
-4.086	-7.886	-11.686
-4.285	-8.270	-12.255
-4.086	-7.886	-11.686
-4.282	-8.264	-12.246
-1.278	-9.238	-17.198
-4.086	-7.886	-11.686
1.432	-8.746	-18.923
18.066	11.971	5.876
-4.086	-7.886	-11.686
-4.086	-7.886	-11.686
5.437	-3.713	-12.863
9	**1**	**1**

FIGURE 60: Probability - Yellow V Red w Audio - Temporal Lobes (at 95% Confidence)