Classification of Error Related Potential (ErrP) in P300-Speller

Anwar Isayed

July 2015

P300-Speller is one of the most popular paradigm for constructing Brain Computer Interface (BCI) system that allows subjects to type letters by focusing on a specific target on a computer screen. When BCI system recognises a different command than the subject's intentions, an Error-Related Potentials (ErrP) occurs from the subject's brain as a response. Researchers aim to build classifiers for detecting ErrPs for single subjects. The aim of this work is to build a transfer learning classifier for detecting ErrPs across multiple subjects. We propose two different ensemble approaches for ErrP detection; Random Forest (RF) and ensemble linear Support Vector Machines (SVMs). The effect of different parameters and methods in a pre-processing stage are studied in order to find the best combination for increasing the detection sensitivity and specificity among different subjects. We obtain 68% Area Under Curve (AUC) at F3 electrode across multiple subject by using the ensemble linear SVM. We show that the F3 and C2 are the best electrodes for detecting ErrP. We also show that it is possible to extract the most useful features from centro-frontal electrodes by using 30 PCA components. We obtain 78% (AUC) by using RF with 32 features. To support our work, we compare our results with the Linear SVM classifier where our results were superior. We concluded that both RF and ensemble linear SVM can cope with the heterogeneity among different subjects.

Classification of Error Related Potential (ErrP) in P300-Spelles

Anwar Isayed

July 2015

يعتبر مهجئ ٣٠٠ ـ الايجابي واحد من أشهر النماذج المستخدمة في بناء انظمة واجهة الدماغ الحاسوبية والذي يسمح بدوره للأشخاص بكتابة الاحرف بالتركيز على هدف معين على شاشة الحاسوب. عندما يتعرف نظام واجهة الدماغ الحاسوبية بطريقة مختلفة عن نوايا المستخدم، يحدث التيار ذات العلاقة بالخطأ من دماغ الفرد كاستجابة. يهدف الباحثين لبناء مصنف لتمييز هذه الاشارات لكل شخص على حدى. يهدف هذا العمل الى بناء مصنف يمتلك القدرة على الانتقال بين الاشخاص لتمييز التيارات ذات العلاقة بالخطأ. في هذا العمل، اقترحنا استخدام مصنفين يمتلكان القدرة على بناء مجموعة من المصنفات الداخلية لتمييز التيارات ذات العلاقة بالخطأ خلال مجموعة من الأفراد، تشمل هذا المصنفات الغابة العشوائية ، ومجموعة الشعاع الدعمي الالي الخطي. تأثيرات المتغيرات المختلفة والادوات في مرحلة ما قبل المعالجة تم دراستها من أجل ايجاد أفضل حساسية وخصوصيةخلال الأفراد. حصلنا على نتيجة ٦٨ بالمئة كمساحة تحت المنحنى في القطب ف ٣ خلال الأفراد باستخدام مجموعة الشعاع الدعمي الالي الخطي. اثبتنا ايضا ان استخدم ف ٣ وسي ٢ هم أفضل الاقطاب لتمييز التيارات ذات العلاقة بالخطأ. اثبتنا ايضا احتمالية استخراج ٣٠ مميز بواسطة تحليل العنصر الرئيسي من المميزات في منطقة الأقطاب المتوسطةوالأمامية. حصلنا على نتيجة ٧٨ بالمئة كمساحة تحت المنحنى باستخدام الغابة العشوائية مع ٣٢ مميز لدعم عملنا، تم مقارنة نتائجنا مع نتائج الشعاع الدعمي الالي الخطي وكانت نتائجنا الأفضل. استخدام الغابة العشوائية ومجموعة الشعاع الدعمي الخطي التعامل مع عدم التجانس بين الأفراد.

Palestine Polytechnic University
Deanship of Graduate Studies and Scientific Research
Master of informatics

Classification of Error Related Potential (ErrP) in P300-Speller

Submitted by:

Anwar Isayed

Thesis submitted in partial fulfilment of requirements of the degree Master of Science in Informatics
July, 2015

The undersigned hereby certify that they have read, examined and recommended to the Deanship of Graduate Studies and Scientific Research at Palestine Polytechnic University the approval of a thesis entitled: **Classification of error-related potentials(ErrP) in P300-Speller**, submitted by**Anwar Isayed** in partial fulfilment of the requirements for the degree of Master in Informatics.

Graduate Advisory Committee:

Dr. Hashem Tamimi (Supervisor),Palestine Polytechnic University.

Signature:_____ Date:_____

Dr. Sami AbuSneineh (Internal committee member), Palestine Polytechnic University.

Signature:_____ Date:_____

Dr. Abualseoud Hanani (External committee member), Birzeit university .

Signature:_____ Date:_____

Thesis Approved

Dr. Murad Abusubaih Dean of Graduate Studies and Scientific Research Palestine Polytechnic University

Signature:_____ Date:_____

DECLARATION

I declare that the Master Thesis entitled **"Classification of error-related potentials(ErrP) in P300-Speller"** is my original work, and herby certify that unless stated, all work contained within this thesis is my own independent research and has not been submitted for the award of any other degree at any institution, except where due acknowledgement is made in the text.

Anwar Isayed

Signature:_____ Date:_____

STATEMENT OF

PERMISSION TO USE

In presenting this thesis in partial fulfillment of the requirements for the master degree in Informatics at Palestine Polytechnic University, I agree that the library shall make it available to borrowers under rules of the library.

Brief quotations from this thesis are allowable without special permission, provided that accurate acknowledgement of the source is made.

Permission for extensive quotation from, reproduction, or publication of this thesis may be granted by my main supervisor, or in his absence, by the Dean of Graduate Studies and Scientific Research when, in the opinion of either, the proposed use of the material is for scholarly purposes.

Any coping or use of the material in this thesis for financial gain shall not be allowed without my written permission.

Anwar Isayed

Signature:_____ Date:_____

DEDICATION

I dedicate my thesis to my god and lord. Also, I dedicate this thesis to my beloved mother, father, brothers and sisters. I would like to thank Dr. Hashem Tamimi who has supported me all the way since the beginning of my studies. I also dedicate this thesis to my friends, especially: Hamza Al-takroury, Mohammad Juba, Samer Isayed, Ibrahem Iqdemat and Ruba Jarrar. Thank you for your moral support, everlasting guidance and endless inspiration. Also, I dedicate this thesis and give special thanks to my best doctor Dana Al-Hamwi. I will always appreciate all you have done. Finally, this thesis is dedicated to all those who believe in the richness of philosophy and learning and to the memory of my grandparents and Tawfiq Tawil.

ACKNOWLEDGEMENT

I give thanks to god for his protection and giving me the ability to do the work. I also thank professor Jrmie Mattout for providing me the data-set I needed.

الملخص

يعتبر مهجئ ٣٠٠ ـ الايجابي واحد من أشهر النماذج المستخدمة في بناء انظمة واجهة الدماغ الحاسوبية والذي يسمح بدوره للأشخاص بكتابة الاحرف بالتركيز على هدف معين على شاشة الحاسوب. عندما يتعرف نظام واجهة الدماغ الحاسوبية بطريقة مختلفة عن نوايا المستخدم، يحدث التيار ذات العلاقة بالخطأ من دماغ الفرد كاستجابة. يهدف الباحثين لبناء مصنف لتمييز هذه الاشارات لكل شخص على حدى. يهدف هذا العمل الى بناء مصنف يمتلك القدرة على الانتقال بين الاشخاص لتمييز التيارات ذات العلاقة بالخطأ. في هذا العمل، اقترحنا استخدام مصنفين يمتلكان القدرة على بناء مجموعة من المصنفات الداخلية لتمييز التيارات ذات العلاقة بالخطأ خلال مجموعة من الأفراد، تشمل هذا المصنفات الغابة العشوائية ، ومجموعة الشعاع الدعمي الالي الخطي. تأثيرات المتغيرات المختلفة والادوات في مرحلة ما قبل المعالجة تم دراستها من أجل ايجاد أفضل حساسية وخصوصيةخلال الأفراد. حصلنا على نتيجة ٦٨ بالمئة كمساحة تحت المنحنى في القطب ف ٣ خلال الأفراد باستخدام مجموعة الشعاع الدعمي الالي الخطي. اثبتنا ايضا ان استخدم ف ٣ وسي ٢ هم أفضل الاقطاب لتمييز التيارات ذات العلاقة بالخطأ. اثبتنا ايضا احتمالية استخراج ٣٠ مميز بواسطة تحليل العنصر الرئيسي من المميزات في منطقة الأقطاب المتوسطةوالأمامية. حصلنا على نتيجة ٧٨ بالمئة كمساحة تحت المنحنى باستخدام الغابة العشوائية مع ٣٢ مميز لدعم عملنا، تم مقارنة نتائجنا مع نتائج الشعاع الدعمي الالي الخطي وكانت نتائجنا الأفضل. استخدام الغابة العشوائية ومجموعة الشعاع الدعمي الخطي التعامل مع عدم التجانس بين الأفراد.

Abstract

P300-Speller is one of the most popular paradigm for constructing Brain Computer Interface (BCI) system that allows subjects to type letters by focusing on a specific target on a computer screen. When BCI system recognises a different command than the subject's intentions, an Error-Related Potentials (ErrP) occurs from the subject's brain as a response. Researchers aim to build classifiers for detecting ErrPs for single subjects. The aim of this work is to build a transfer learning classifier for detecting ErrPs across multiple subjects. We propose two different ensemble approaches for ErrP detection; Random Forest (RF) and ensemble linear Support Vector Machines (SVMs). The effect of different parameters and methods in a pre-processing stage are studied in order to find the best combination for increasing the detection sensitivity and specificity among different subjects. We obtain 68% Area Under Curve (AUC) at F3 electrode across multiple subject by using the ensemble linear SVM. We show that the F3 and C2 are the best electrodes for detecting ErrP. We also show that it is possible to extract the most useful features from centro-frontal electrodes by using 30 PCA components. We obtain 78% (AUC) by using RF with 32 features. To support our work, we compare our results with the Linear SVM classifier where our results were superior. We concluded that both RF and ensemble linear SVM can cope with the heterogeneity among different subjects.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

ACC	Anterior Cingulated Cortex
AUC	Area Under Curve
BCI	Brain Computer Interface
CRP	Corneal-Retinal Potential
BSS	Blind Source Separation
EEG	Electroencephalogram
EMG	Electromyography
ERP	Event-related potential
ErrP	Error-related potential
ICA	Independent Component Analysis
IIR	Infinite Impulse Response
LDA	Linear Discriminant Analysis
PCA	Principle Component Analysis
RF	Random forest
ROC	Receiver Operating Characteristic
SCP	Slow Cortical Potentials
SMR	Sensori Motor Rhythms
SNR	Signal-to-Noise Ratio
SSVP	Steady-State Visually Evoked Potentials
SVM	Support Vector Machine

Chapter 1

Introduction

Brain Computer Interface (BCI) is a communication system that can be used to detect characteristics and extract the patterns of the brain signals from a technique to determine the subject's intent [72, 58]. There are a variety of techniques which are used to read the activity of a brain such as Electroencephalography(EEG), Electromyography (EMG), etc. EEG is one of the most popular techniques available to record signals from the nerves in the scalp[68]. One type of BCI that uses natural responses of the subject's brain to external visual stimuli is called a P300-Speller [57]. This type allows subjects to type letters by focusing on a specific target on a computer screen.

Unfortunately, the nature of the brain signals in non-invasive techniques, especially EEG, has highly complex and low signal-to-noise ratio (SNR) in addition to misinterpret the signals from the computer sometimes [25, 46]. Thus, it is difficult to use BCI outside the research laboratory. To overcome these problems, several approaches have been explored. One approach that has been proposed for detecting errors in EEG-based BCI is an Error related-potential (ErrP) by means of cerebral orders that generates a response after

the subject receives feedback.

In general, BCI systems require extensive subject training [39]. The main aim of this work is to build a transfer learning classifier for detecting the ErrP across multiple subjects (no need for previous training among subjects). This classifier is used for detecting ErrP within the Context of P300-speller. It also aims to improve the final decision of P300 classifier to achieve a reliable BCI system. This may allow disabled individuals to interact with their environment in real life and a robust BCI system with high accuracy which might be achieved.

In this work, we apply Random Forest (RF) as one of the state of the art techniques that are used in classification to discriminate the ErrP across multiple human subjects. RF is an ensemble machine learning technique provides embedded classification and feature selection approach. Since RF can capture interaction and complex relationships in the data [22], it can cope with the heterogeneity among different subjects.

In addition to the RF, we propose a new ensemble machine learning technique that can effectively detect the ErrP by combining multiple linear Support Vector Machines (SVMs). We refer to it as ensemble linear SVM. In addition we use Principal Component Analysis (PCA) to extract the most useful features from Anterior Cingulated Cortex region.

1.1 Contributions

This thesis contributes to the field of BCI by detecting ErrP signals. We achieved the following contributions:

1. Unlike the existing work which focus on single subjects when detecting ErrP, we apply comprehensive methodology for building a transfer learning classifier to detect ErrP among a wide array of subjects. To deal with the possible heterogeneity of signals across multiple human subjects, we study two ensemble approaches RF and ensemble linear SVM.

2. We apply state of the art techniques that are used in pre-processing steps exhibit different performance to figure out the most useful parameters across multiple human subjects.

3. We compare our results with a linear SVM technique, which is commonly used in a BCI system to support this study. Unlike many existing studies which rely on accuracy as a performance measure, we use the area under ROC curves since this measure is more comprehensive than accuracy.

1.2 Thesis Outline

This research is organised as the following: we provide an introduction and our contributions in chapter 1. The rest of this thesis is organised as what follows: In chapter 2, we provide a background related to this work and a summary of some previous works related to our work. In chapter 3, we describe a benchmark and our methodology of this thesis. In chapter 4, we

provide our results and discussion. Finally, we provide conclusion, two block diagrams of how a real-time detector works and future work in chapter 5.

Chapter 2

Background

In this chapter, the theoretical background needed to understand the remaining parts of the thesis is presented. This chapter is organised as follows. In the first section, the concepts of BCI are explained. In the second section, we describe Electroencephalogram (EEG) then we describe the P300-Speller, including the concepts of ErrP in the third section. The signal pre-processing methods (BSS and Buterworthfilter) will be describe in the section four. The fifth section describes features extraction that contains the concepts of Principle Component Analysis (PCA) as a feature extractions tool. The sixth section describes Linear Support Vector Machine (SVM) and Random Forest(RF). The seventh section describes a Receiver Operating Characteristics (ROC) that will be used in this work to evaluate our results. The final section covers the literature review related to this work.

2.1 Brain Computer Interface(BCI)

Brain Computer Interface (BCI), also referred to as a brain machine interface (BMI), Human Computer Interaction (HCI) or sometimes called a Mind-Machine Interface (MMI), is mainly used as one of the most promis-

ing systems, which provides a direct communication link between the subject's brain to a computer for reflecting their intentions to the external world [26, 49, 2].

There are three main types of BCI: invasive BCI, partially invasive BCI and non-invasive BCI[30, 34]. An invasive BCI is any BCI technology which involves implanting a foreign device into the subject's brain, while noninvasive BCI is recorded from the scalp directly such as EEG. Partially invasive BCI are implanted inside the skull of the subject.

There are certain set of phases that must be considered when designing a BCI system: data acquisition, pre-processing, features extraction and classification[55]. Figure 2.1 illustrates the block diagram of the BCI. The first step represents an EEG technique in signal acquisition stage, which are used to capture the brain activity. In the pre-processing stage, filters are used to increase the signal-to-noise ratio and to avoid artifact in a suitable form for feature extraction. The feature vectors are often of high dimensionality[50, 14]. To reduce the dimensionality of the features, feature selection can be performed. In this step, the most useful features are chosen, while other are omitted. Finally, the classification stage classifies the features from feature extraction stage in order to decipher the subject's intentions.

2.2 Electroencephalogram (EEG)

An electroencephalogram (EEG) is a non-invasive technique that reads electrical activity generated by the brain from the scalp surface after being

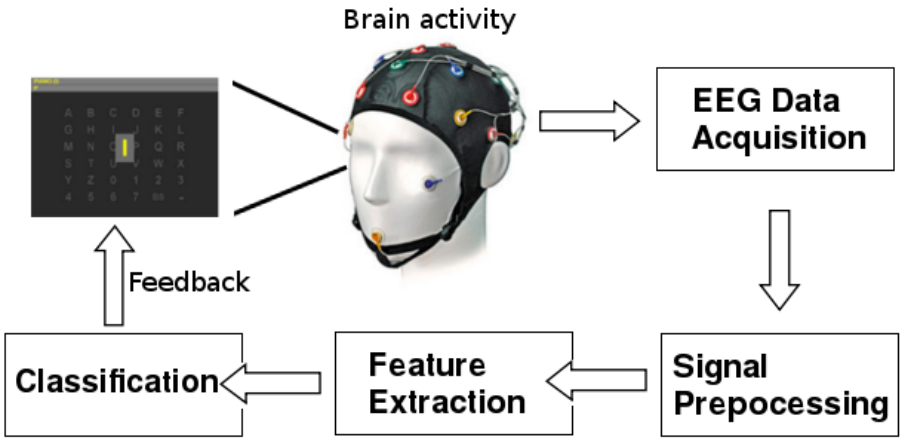


Figure 2.1: Block diagram of the BCI work process

attached to electrodes on the subject's scalp [72]. In Figure 2.2, the non-invasive EEG electrodes on the subject's scalp is shown. It measures the brain's electrical activity to be viewed as brain waves and provide them to a BCI system. In the field of neuroscience, EEG reflects the firing of neurons in the subject's brain[70]. It is characterised by easy of use, high temporal resolution and has low setup cost[3]. It is the most common type of non-invasive BCI[30, 76].

Figure 2.2: EEG electrodes [42]

In general, BCI has four different types of EEG activity patterns, which

are used as the BCI system inputs: The P300 event-related potentials (ERP), sensorimotor rhythms (SMR), slow cortical potentials (SCP) and steady-state visually evoked potentials (SSVEP) [24, 43]. The ERP is one of the most reliable and popular activity pattern in the EEG recording used for constructing BCI systems, but it is not easy to be detected [72]. It occurs as changes in EEG in the response of the stimulus such as visual stimulus. It has different components as shown in Figure 2.3.

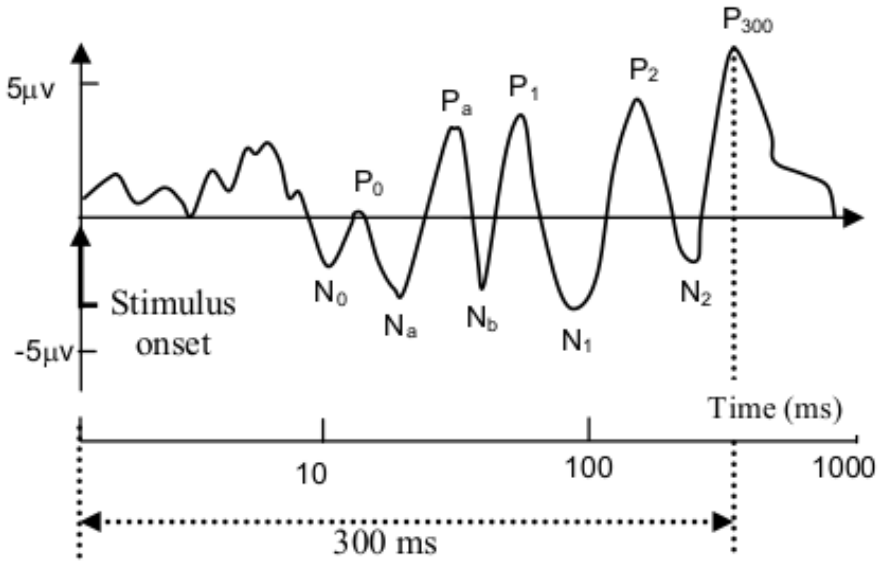


Figure 2.3: Components of Event Related Potentials (ERP) [20]

2.3 BCI P300-Speller

P300 is one type of ERP components typically that evoke a positive peak, visible in EEG at about 300 ms after the stimulus onset[62, 51, 52, 11]. One of the greatest advantages of the P300 is that it does not require intensive subject training [20]. It appears as a positive peak about 300 ms after the stimulus onset.

P300-Speller paradigm presents one of the most widely BCI system [20]. The first one was designed by Farwell and Donchin [18]. It exploits the P300 component of ERP to enable a subject to type letters by focusing on a specific target on a computer screen. It is very effective in detecting the characters[51].

In this paradigm, there is a 6×6 matrix of symbols. It has 26 letters of alphabet and 10 digits (See figure 2.4). Each row and column of the matrix is flashed in random order[59]. By analysing the P300 using a machine learning classifier, it is possible to identify the target letter by identifying the row and column. The selected row and column have one common letter.

P300-Speller paradigm has been modified in many ways such as modify the way of random flashing. [73] proposed a random flashing way to improve the spelling accuracy by constructing six pairs of groups of six letters instead of the traditional way of the random flashing. After the analyses of the P300, the target letter is shown on the centre of the screen to get feedback.

There are several barriers to be handled in P300: First, detecting the P300 in a single trial is very difficult, and the number of repetitions is predetermined for each user to get the best trade-off between speed and accuracy. Increasing the number of trials leads us to increase the time needed to obtain stable and reliable (ERP)[54]. The second problem is the nature of the brain signals, especially non-invasive techniques (e.g. EEG) have a high complexity and low signal-to-noise ratio of the ERP. Noisy EEG data can be challenging due to the low signal-to-noise ratio[45]. The third problem is related to the system, sometimes BCI misinterprets the signal related to

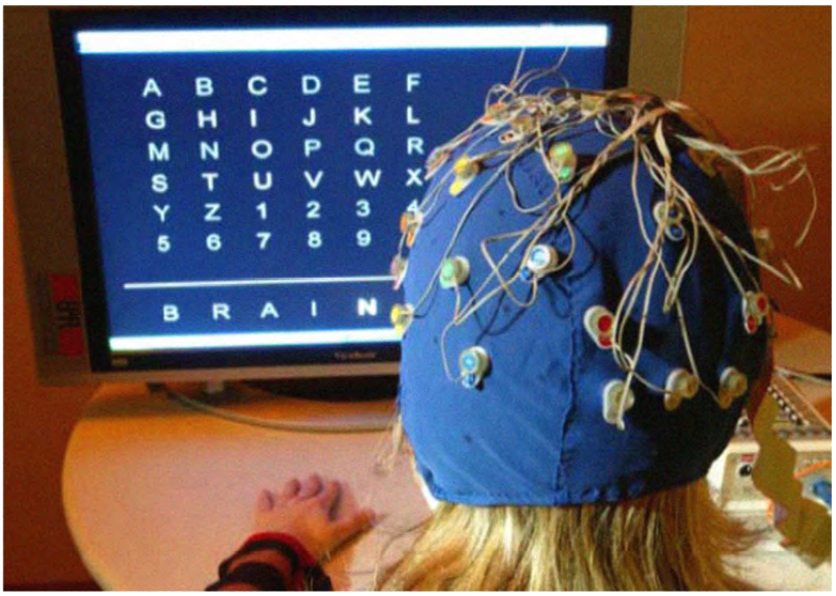


Figure 2.4: P300-Speller[63]

P300 classification[46]. All these factors increased the need to look for a way to detect errors by a technique that gets the P300 accurately and fast.

2.3.1 Error Related Potential(ErrP)

The error-related potential (ErrP) is an ERP generated from the subject's brain as a response when a BCI recognises a different command than the subject's intentions. It was first proposed by Falkenstein in 1991[16]. A different command is generated when a subject makes a mistake or the BCI misinterprets the subject's intent[61]. It is generated at the end of each trial after the feedback onset[60, 46]. Recently, ErrPs have been used as feedback instructions for some devices to correct the errors[23]. It has been proposed for detecting errors in BCI system to reduce the number of errors and trials[11]. The shape and latency times of ErrP depending on the kind of paradigm[9].

There are different types of ErrPs characteristics that were discovered: response ErrP is the most known one . It occurs when subjects perform an

incorrect response action to a stimulus. Feedback ErrP occurs when subject has to be made aware of the mistake through an external feedback. Interaction ErrP (ErrPi) occurs when the system behaves differently from the subject intent, or the subject generates an error by himself (eg. changes his mind)[74, 71]. Some authors make no distinction between them [60]. In this thesis, we will make no distinction between them, either.

The vast majority of the authors reported that ErrP is centralised from a few electrodes: [Cz], [Cz and FCz] and fronto-central areas[9]. See figure 2.5 for the location of these electrodes. Each electrode has a letter and a number. The letter identifies the lobe (F: Frontal, T: Temporal, C: Central, P: Parietal and O:Occipital) and the number identifies the hemisphere location.

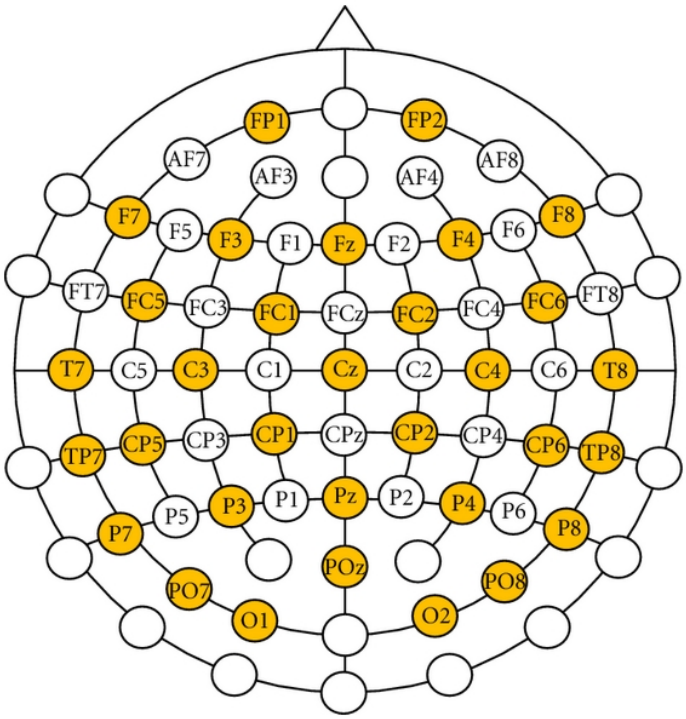


Figure 2.5: EEG channel montage[46]

To analyse ErrP, we follow the same steps which are shown in Figure 2.1 to find out the ErrP characteristics by extracting signals. The main aim of the classifier is to make a decision about the ErrP signals in two classes, one of them is for correct trial, when the system behaves the same as intent of the subject, and other is for error trial. If there is an error, other options will be taken into account to correct it like using language models[1]. It is used to correct the target letter. In our research, we will focus on detecting ErrP signals without focusing on error correction.

The length of signals depend on an experiment, but typically is 600ms[8]. Recently, Spuler [69] shows that a window of 200 ms to 900 ms for classification ErrPs gave the best results in motor-imagery. The motor-imagery is a mental process in which a subject imagines a movement [47].

2.4 Pre-processing of ErrP data

The classification of ErrP in EEG data is difficult due to the low signal-to-noise ratio (SNR) and artifacts noise often produced by the eye movements or blink of the subject during EEG recordings [32]. To overcome these problems, there are methods to remove noise without removing the useful information of EEG signals, such as Blind Source Separation (BSS) and butterworth filter types, which are described in the next sections.

2.4.1 Blind Source Separation (BSS)

BSS is a common way used to recover the source signals from mixtures of unknown signals without prior information about the sources[35, 12]. The basic block diagram of the BSS process is shown in Figure 2.6. In this figure,

X is a mixture of unknown signals that has the mixed source vector $S_{(n)}$, noise and interference signals to produce $X_{(n)}$.

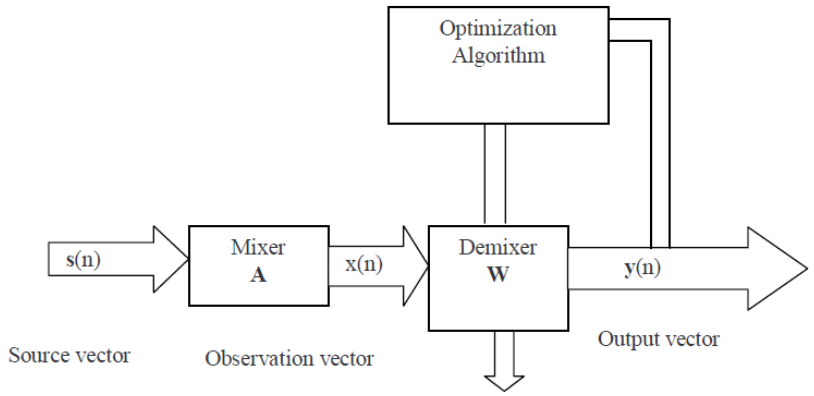


Figure 2.6: The basic block diagram of the BSS[12]

An optimisation algorithm like Independent Component Analysis (ICA)[13], which is now a standard to find a matrix W(the inverse of A) of BSS. It has the capability to separate and extract the source signals $Y_(n)$[12, 64]. W is used to restore the source signals including noise and artifacts. The basic idea of ICA is shown in figure 2.7. BSS can be expressed as[48]: $A_{m\times m}.S_{m\times n}=X_{m\times n}$ Where A represents an mixing matrix $m \times m$, and $Y = WX \cong S$.

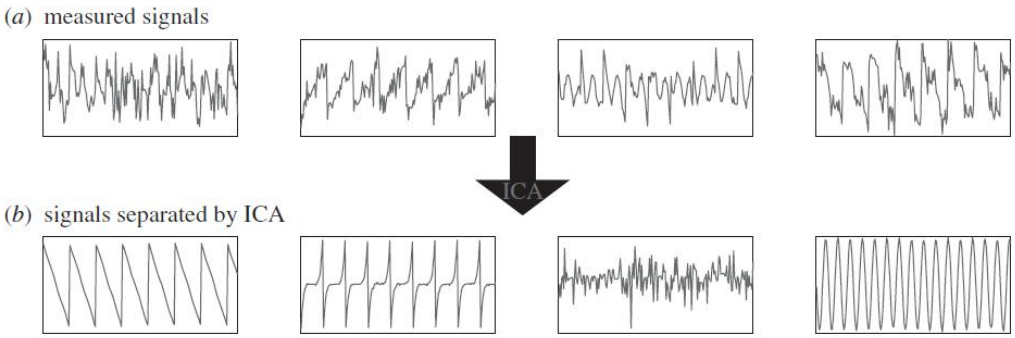


Figure 2.7: The basic idea of ICA[27]

ICA has many algorithms. One of them is called fastICA[28], which is used to estimate independent components for solving the BSS problem. It supports good accuracy and fast speed and it can be easily considered as a scalar shifted version of an ICA algorithm [64]. The algorithm of fast ICA is[12, 4]:

Algorithm 1 The algorithm of fast ICA

1. Initialise randomly vector $w(0)$ and normalise it to unity. Let $k = 1$.
2. Update $w(k) = \hat{E}\{x(w(k-1)^T x)^3\} - 3w(k-1)$ by using Newton phase. The expectation operator can be estimated using a large number of samples.
3. Normalise w(k) to unity length. $w(k) = \frac{w(k)}{||w(k)||_2}$.
4. If $|w^T(k)w(k-1)|$ is not close to 1, then set $k = k+1$ and repeat step 2, Otherwise, output vector $w(k)$.
5. If there are n independent components to be estimated, then the above algorithm run for n times.

The BSS algorithm is used to overcome the eye noise (eye movements and blink) in ERP during multi-channel EEG recording[77, 33]. Joyce[31] described a completely automated method for eliminating eye noise using BSS algorithm and Electrooculography(EOG) channels. EOG is a technique used for measuring the standing corneal-retinal potential (CRP) [66]. Flag those components that invert.

This algorithm demonstrates how to use the BSS algorithm for cleaning the eye noise from EEG data. It consists of five steps:

1. Decompose the EEG data onto a set of components using a BSS algo-

rithm.

2. Reverse the sign on all lower and horizontal EOG channels (i.e., multiply signals by -1) and again decompose data onto components using a BSS algorithm.

3. identify the components that correlate above a certain level (correlation threshold) with the pre-processed lower and horizontal EOG channel data.

4. Flag BSS components with high power in the low frequency band.

5. Remove from the data those components identified in Step 2, and those that were identified in both Steps 3 and 4.

2.4.2 Filtering

Filtering is a pre-processing step. It is used to improve the signal-to-noise ratio[75]. Infinite Impulse Response (IIR) filters are widely used in many applications, such as EEG signal processing[75]. The general equation of an N^{th} order IIR filter can be expressed as follows:

$$y(n) = \sum_{i=0}^{M} b_i x(n-i) - \sum_{i=0}^{N} a_i y(n-j) \tag{2.1}$$

Where b_i and a_i are the coefficients that define the filter, N is the order of filter, n: the point that is to be calculated, x: input signal. A filter can be characterised by its frequency response : low pass, high pass, band stop and band pass.

The transfer function of IIR filters can be expressed as follows:

$$H(z) = \frac{\sum_{i=0}^{M} b_i z^{-i}}{1 + \sum_{j=0}^{N} a_j z^{-j}} \tag{2.2}$$

The Butterworth filter is a type of the IIR filters. This type is an analogue filter that was described by Stephen Butterworth in 1930[7]. It has been originally used to allow a certain frequency to pass to its output and reject all other frequencies. The transfer function for butterworth filter is[67]:

$$|H(\Omega)| = \frac{1}{\sqrt{1 + (\frac{\Omega}{\Omega_p})^{2N}}} \tag{2.3}$$

Where Ω_p is the cutoff frequency and Ω is the analog frequency of the first specifications.

2.5 Features extraction

In EEG signals, the features can be applied to a classifier directly from the raw data, but in this method there is no clear patterns [38, 41]. It is important to understand what features are used. A great variety of features extraction methods are used in the ERP data. The most known of these methods that are used to enhance the classification accuracy is[42]: ICA used as a features extraction tool and Principal Component Analysis (PCA).

2.5.1 Principle Component Analysis (PCA)

PCA is a statistical algorithm that is used to extract the important features, reveal features that are sometimes hidden from confusing data-sets and reduce a complex data to a lower dimension as a set of principal components [65]. Figure 2.8 illustrates the basic idea of PCA.

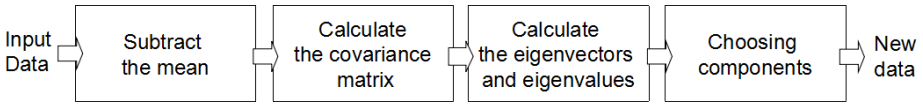


Figure 2.8: The basic idea of PCA.

As a first step, the mean of all the data points in each of the data dimensions is computed and subtracted for each dimension defined as:

$$\overline{X} = \frac{\sum_{i=0}^{N} x_i}{N} \tag{2.4}$$

Where x: data point, $\overline{X}$ is the mean and N is the number of data points.

Secondly, calculating a covariance matrix of the result data from the first step. The covariance between X and Y is:

$$Cov(X,Y) = \frac{\sum_{i=0}^{N}(X_i - \overline{X})(Y_i - \overline{Y})}{N} \tag{2.5}$$

Since the data is m dimensional, the covariance matrix will be m x m between dimensions. The covariance matrix for an imaginary square dimensional $(m \times m)$ will be:

$$C = \begin{pmatrix} Cov(X,X) & Cov(X,Y) & \cdots & Cov(X,m) \\ Cov(Y,X) & Cov(Y,Y) & \cdots & Cov(Y,m) \\ \vdots & \vdots & \ddots & \vdots \\ Cov(m,X) & Cov(m,Y) & \cdots & Cov(m,m) \end{pmatrix} \tag{2.6}$$

Where Cov(X,X) is also the variance.

In the third step, we calculate the eigenvectors and eigenvalues from a covariance matrix. The eigenvalues of C are defined as:

$$|C - \lambda I| = 0 \tag{2.7}$$

where I is the m x m identity matrix. Let λ is the eigenvalue. The eigen-

vectors of C are defined as:

$$\lambda_i u_i = C u_i \tag{2.8}$$

where u_i is the eigenvector. Each of the eigenvector is closely related to an eigenvalue. The eigenvectors are sorted in descending order by the eigenvalue, and are called principal components. The highest eigenvector values are the components of the greater significance and the lower ones can be ignored. This is called the dimensional reduction [44]. Finally, we choose k components to form a projection matrix W (every column represents an eignen vector). We use this matrix to transfer the original data-set onto the new subspace can be defined as:

$$y = W^T \times x \tag{2.9}$$

2.6 Classification of ErrP

This section gives a brief overview of linear Support Vector Machine (SVM) and Random Forest (RF) in order to better understand the classifiers that are used in this thesis.

2.6.1 linear Support Vector Machine(SVM)

In this section, we briefly describe the formulation of linear SVM. SVM is a machine learning technique that was initiated in [10] for binary classification. Let $(x1, y1), (x2, y2), ..., (x_N, y_N)$ is the training dataset , where $x_i \in \mathbf{R}^n$ and $y_i \in \{-1, +1\}$.

Linear SVM cycles a hyperplane across the training dataset to design an optimal hyperplane that separates classes(See figure 2.9). This hyperplane

creates a margin between the classes.

The optimal hyperplane is defined by the equation of:

$$F(x) = \vec{w}.x - b = 0 \tag{2.10}$$

where $\vec{w}$ is a vector of weights and b is the bias that would classify the training data correctly. The margin m is defined by:

$$m = \frac{1}{||w||_2} \tag{2.11}$$

The total margin is:

$$m = \frac{2}{||w||_2} \tag{2.12}$$

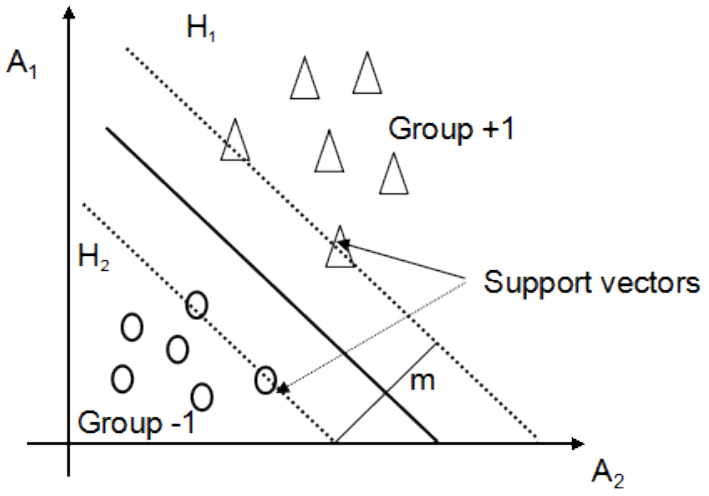


Figure 2.9: Linearly Separable SVM[5]

Minimising the $\vec{w}$ will maximise the separability. The maximising margin can be expressed as a *Primal* optimisation problem:

$$\min_{w,b} \frac{1}{2}||\vec{w}||^2 \quad subject \quad to \quad y_i(\vec{w}.\vec{x}_i - b) \geq 1, \forall_i \tag{2.13}$$

In the boundary of the margin, there are supporting hyperplanes parallel

and equidistant vectors called support vectors (α). Any data on or above the boundary is class +1, and any data on or bellow the boundary is class -1. They are defined by the following equations of:

$$F(x) = \vec{w}.x - b = +1 \tag{2.14}$$

$$F(x) = \vec{w}.x - b = -1 \tag{2.15}$$

A new data x_i can be classified by evaluating the sign of $F(x)$ with:

$$\begin{aligned} F(x) &= sgn(\vec{w}^T x + b) \\ &= sgn(\sum_{i=0}^{N} \alpha_i y_i (x_i.x) - b) \end{aligned} \tag{2.16}$$

2.6.2 Random Forest (RF)

Random Forest (RF) is an ensemble learning method developed by [6] in 2001 for classification and regression. It runs efficiently on large data, handles thousands of input variables and gives estimates of important variables [36]. It generates a different subset of variables from the training data (See Figure 2.10). It aims to train many decision trees and let them vote for the popular class [17]. A randomly selected subset of variables is used to split each node.

In RF, the part of training samples is taken out to estimate error and variable importance. The class assignment is made by the number of votes from all of the trees. In the regression the average of the results is used.

The random forests algorithm is as follows[40]:

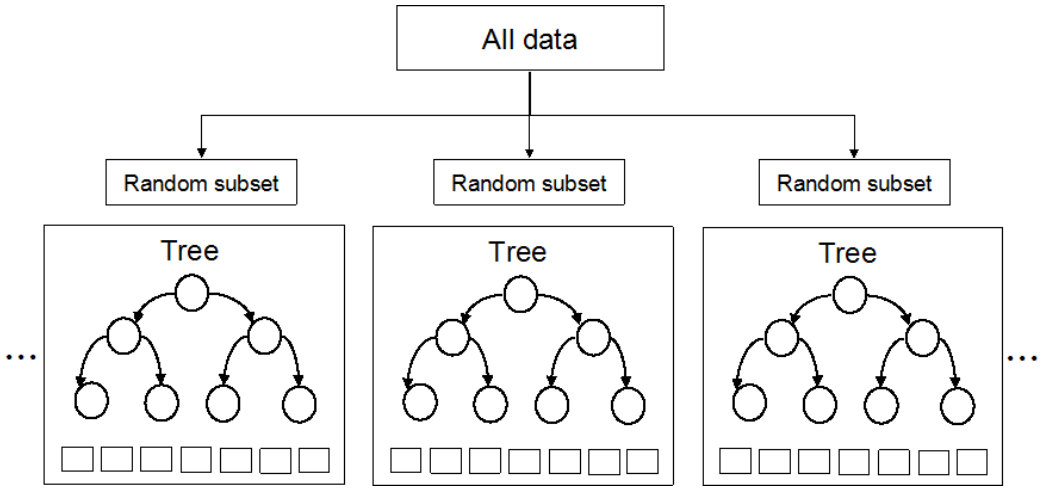


Figure 2.10: Random Forest

1. Draw *ntree* bootstrap samples from the original data.

2. For each of the bootstrap samples, grow a tree with the following: at each node, rather than choosing the best split among all predictors, randomly sample m_{try} of the predictors and choose the best split from among those variables.

3. Predict new data by aggregating the predictions of the ntree trees (i.e., majority votes for classification, average for regression).

An estimate of the error rate can be obtained, depending on the training data, by the following[40]:

1. At each bootstrap iteration, predict the data which is not in the bootstrap sample (what Breiman calls "out-of-bag", or OOB, data) using the tree grown with the bootstrap sample.

2. Aggregate the OOB predictions as an on the average. Each data point would be out-of-bag around 36% of the times, so aggregate these predictions.

3. Calculate the error rate, and call it the OOB estimate of error rate.

2.7 A receiver operating characteristics (ROC)

A Receiver Operating Characteristics (ROC) is a graphical tool that used for evaluating, organising and selecting a binary classifier based on their performance [19]. There are two class labels in the binary classifier (positive and negative) and they are not totally separated, so each classifier has four possible outcomes (True Positive, false negative, true negative, false positive). These possible outcomes are shown as a matrix called common confusion matrix. See table 2.1.

Table 2.1: A confusion matrix

	Positive	Negative
Yes	True Positives	False Positives
No	False Negatives	True Negatives

Based on the common confusion matrix, we can calculate sensitivity and specificity as shown in equations 2.1 and 2.2:

$$Sensitivity = \frac{True \quad Positive}{True \quad Positive + False \quad negative} \tag{2.17}$$

$$Specificity = \frac{True \quad Negative}{True \quad Negative + False \quad Positive} \tag{2.18}$$

ROC plots sensitivity on y-axis and (1 specificity) on x-axis for varying cut-off points of test values [37]. In ROC analysis, distributions of test results for positive and negative defined by threshold and shown by its sensitivity and specificity. Each threshold ranging from 0 to 1. To find optimal threshold point from ROC curve, we can use methods to give equal weight to sensitivity and specificity.

We use the area under ROC in this study because ROC curve are more comprehensive than other measures, such as accuracy. ROC lets use observe the trade-off between sensitivity and specificity for all possible thresholds in the study, rather than one value. So looking at the ROC curve is a way to assess the model independent of the choice of a threshold.To obtain the optimal cut-off point, we calculate the distance for each observed cut-off point, and locate the point where the distance is minimum. The value of Area Under Curve (AUC) is between 0 and 1 (or 0 and 100%). This value can be used to evaluate the classifier. The larger this value is the better classifier is. For a random classifier (random guess) the area under the ROC curve equals to 50%.

2.8 Literature review

Some studies have already presented ErrP detection and a few studies have been conducted so far on the online ErrP detection. Since most of the data for this thesis comes from Margaux et al.[46], it is important to start with their achievements. They presented the first online P300-Speller that can detect and correct ErrP. They used the 5DAWN algorithm[56] that provides orthogonal linear spatial filter. There were five sessions and used the first four spelling sessions for detecting ErrP, the last session is used ErrP to correct the trial. The length of the window in their work ranges from 200ms to 600ms. They also used a mixture of multi-dimensional Gaussian model as a classifier. To evaluate the ErrP classification, the common confusion matrix was used, and three quantitative measures were also used to evaluate

the error correction. They used the second best guess of a trial to correct a target letter.

Margaux et al.[46] also used bit rate imperfect measure of BCI performance to compare the spelling accuracy with and without online correction. It gives the opportunity to spell the letter again. They also computed the difference between responses to correct and incorrect feedback at Cz electrode. They found a negative ErrP between 250 and 500 ms followed by a positive Errp between 350 and 550 ms in the feedback signal. For the spelling of the letter, the performance was 64% for fast mode and 80% $\pm$18 in slow mode, while the sensitivity, specificity and accuracy in error detection was obtained 63%, 88% and 78%, respectively.

In other study by Margues et al.[53], the data by 32 electrodes was recorded, using 10-10 system. They used a Bandpass filter between 0.1 and 20Hz and down sampling data at 100Hz. They also corrected the errors of the eye movements using ICA algorithm. The length of the window was 1-600ms. They used a xDWAN as a spatial filter. They used a two-class naive Bayes classifier. Their result was that specificity was above 90%, while sensitivity depended on the size of the training set.

Combaz et al.[9] had two groups of subjects. Each group was used in an experiment. The first group consisted of six subjects, whose data was observed by the researchers to elicit ErrP and compare their results with their previous study. It aimed to study the possibility of detecting ErrP using several classifiers and measuring their accuracy. In their two experiments, they used ultra low-power 8-channel wireless placed over the frontal, central and

parietal areas of the brain. The data was sampled-down to 80 data points. In the detection of P300 ERP, the signal was filtered between 0.3 and 15 Hz using Butterworth filter with zero-phase 3rd order, and then they cut the feedback into 800 ms from the stimulus onset. Finally, they built a linear support vector classifier with a 10 cross-validation to detect ERP.

The second group had two new subjects, they used the classifier from the first experiment online to detect the P300-Speller. In the detecting ErrP, the length of the feedback was 2 seconds and the signal was filtered by zero-phase 3rd Butterworth filter between 0.5 and 15 Hz. They explored that the ErrP was 320 ms negative peak followed by 450 positive peak in Fz and FCz. They analysed the difference between the responses to the erroneous feedback. They corrected the feedback by the sampled-down the signals from 1000 Hz to 100 Hz. They studied the influence of the amount of the training data of the classification of ErrP. To test the classifying of the feedback responses, they used a Fisher Linear Discriminant Analysis (FLDA) and a linear SVM. Before that, the signals were filtered between 0.5 and 30 Hz using zero-phase 4th order Butterworth filter. They finally confirmed the SVM outperformed the FLDA for all the subjects except one subject.

Mainsah et al.[45] compared the improvement of various corrective algorithms accuracy, and they compared the performance of spelling correction with ErrP and non-ErrP offline. In P300 classification, they used data from the first session to train a Linear Discriminant Analysis(LDA) classifier. In the second session, They used a LDA classifier to train ErrP classification.

Seno et al.[11] worked on the first online ErrP detection method. They

applied a dual classifier processing module to handle both P300 Speller and ErrP classification. In P300, a genetic algorithm was used to extract features. For training the features, they used a logistic classifier. During the training, the speller was used in a copy mode only. In the ErrP detection, they first used data from Fz to Cz electrodes in the 10-20 system. Then, they filtered the EEG signals by 1-10Hz bandpass filter. After that, they developed an determined significant intervals automatically by using the t-test, and Density-based spatial clustering of applications with noise (DBSCAN) algorithm[15] to fill holes and small intervals from t-test for classification into a logistic classifier. Finally, they reached the accuracy of about 60%.

The summary of literature review is described in table 2.2.

Table 2.2: Summary of literature review: papers with quantitative results are mentioned in the table.

Cite	Pre-processing	Window	Classifier	Results
[46]	- Downsampling into 100Hz - 1-20Hz Bandpass filter - 5DWAN spatial filter	200-600 ms	Multidimensional Gaussian model	Accuracy: 78%
[53]	- Downsampling into 100Hz - 0.1-20Hz Bandpass filter - xDWAN spatial filter	1-600ms	Two-class naive Bayes classifier	- Specificity is above 90% - Sensitivity depending on the size of training set.
[9]	- Downsampling into 100Hz Two experiments: - 0.5-15Hz 3rd Butterworth - 0.5-30Hz 3rd Butterworth	2 sec	Two experiments: FLDA Linear SVM	linear SVM outperforms FLDA
[45]	- Downsampling to 80 points (800 ms to 80 points) - 0.3-15Hz 3rd Butterworth	1-800ms	Linear SVM	Not available
[11]	- 1-20Hz Bandpass filter	-	Logistic classifier	Accuracy: 60%

Chapter 3

Data and Methods

This chapter introduces the data and our methodology that are used in this thesis. In the first section of this chapter, we describe the data-sets. In the second section, we describe our methodology to detect ErrP.

3.1 Data-set

In this section, we introduce the data sets that we use to conduct our experiments. These data was taken from Perrin et al. [46] who collected EEG data from twenty six healthy subjects. The data sample rate was 600 Hz and down-sampled at 200 Hz. It was recorded at 56 electrodes placed at standard positions of the 10-20 system. EOG also was recorded to track the eye movements that cause eye noise. The subjects were asked not to blink their eyes during feedback presentation. The size of data equals 13.8 GB.

For each subject, the EEG data is collected under five sessions. The description of the data is shown in table 3.1. The details of the sessions of each subject are shown in appendix 1.

After the last flash of each trial during each session, feedback are presented about 1.3 seconds (260 data points) at the centre of a visual field. This period is used to detect the ErrP. The feedback is displayed in the middle of the screen in a large font. The subject is instructed to keep looking at the screen and wait for the feedback. For each subject, the sessions from one to four consist of 240 samples, while the fifth session consists of 100 samples. Each subject has 340 samples. The subjects participate in the same sessions but they did not show the same amount of errors. A standard matrix 6×6 is applied in P300-Speller.

There are two types of spelling copy conditions in our benchmark, corresponding to short and long trials:

1. A fast mode (each item was flashed 4 times).

2. A slow mode (each item was flashed 8 times).

Table 3.1: Description of the dataset

Subject Number	**Total corrects**	**Total errors**	**Subject Number**	**Total corrects**	**Total errors**
1	273	67	**14**	215	125
2	220	120	**15**	320	20
3	141	199	**16**	211	129
4	273	67	**17**	226	114
5	142	198	**18**	261	79
6	316	24	**19**	224	116
7	307	33	**20**	236	104
8	246	94	**21**	312	28
9	286	54	**22**	313	27
10	309	31	**23**	221	119
11	225	115	**24**	239	101
12	189	151	**25**	197	143
13	178	162	**26**	181	159

3.2 Methodology

To detect the ErrP across multiple subjects, it is important to design an appropriate methodology which enable us to identify the error or correct the feedback across multiple human subjects. This section introduces to the reader the methodology, which is used in this thesis. Figure 3.1 shows our entire methodology details and the steps that are performed in this thesis. Our methodology is based on two main classifiers: RF and a proposed new ensemble linear SVM. We employ RF to figure out the best parameters of the pre-processing steps in addition to detect the ErrP. We also apply our ensemble linear SVM to detect the ErrP as an alternative. In addition, we compare these two classifiers with the state of the art method for Errp detection, which is based on a single linear SVM.

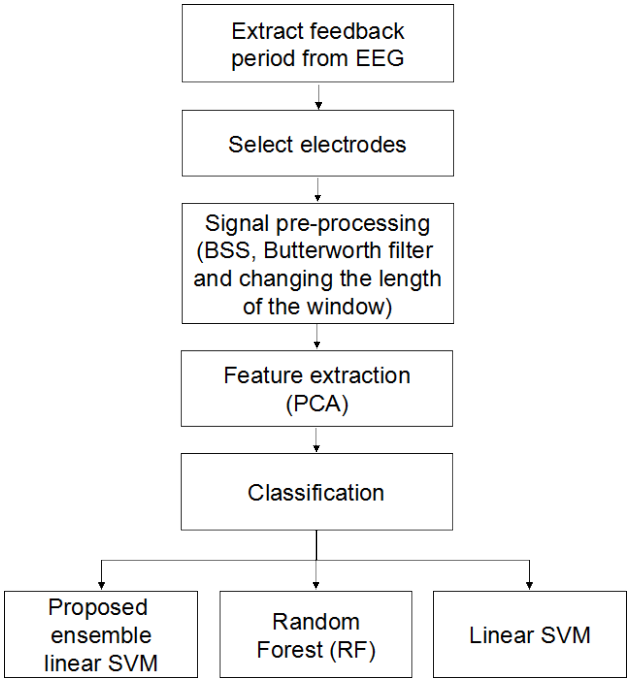


Figure 3.1: The working process of this thesis.

The first part of the research focuses on extracting the feedback period

that contains ErrP. We select the electrodes over frontal-central regions, which generate ErrP[21] and we ignore other data. The frontal-central electrodes namely are in the positions in figure 2.5: (F7, F5, F3, F1, Fz, F2, F4, F6, F8, FC5, FC3, FC1, FCz, FC2, FC4, FC6, C5, C3, C1, Cz, C2, C4 and C6). We use this data in our experiments.

Depending on the outcome of the first portion, we apply the pre-processing steps. The following are the key methodologies utilised in the process of the pre-processing steps (see figure 3.2).

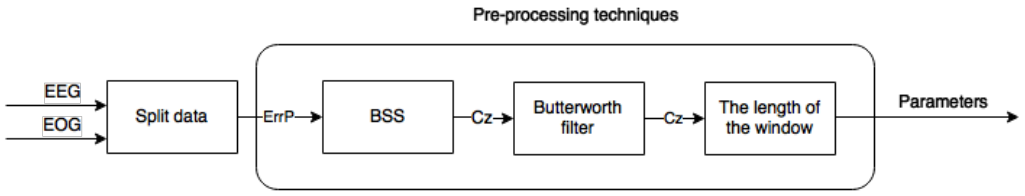


Figure 3.2: The block diagram of the pre-processing steps.

The BSS algorithm, which is explained in chapter 2, is used to eliminate the eye noise from EEG electrodes. We study the impact of changing the different correlation thresholds on the ability of BSS algorithm to correct data across multiple human subjects. To study the correlation threshold impact, we use data from Cz electrode, which is usually selected to detect ErrP[29]. We evaluate the impact by using 10-fold cross validation with RF as follows:

1. Apply RF to the raw data.
2. Apply BSS algorithm with different correlation thresholds with RF.
3. Compare their results and identify the impact of BSS algorithm to correct the data across multiple human subjects.

After applying the BSS algorithm, we apply Butterworth filter in order to improve the signal-to-noise ratio. In this phase, we apply a butterworth filter with different parameter (order and cutoff frequencies) that we have identified by research to figure out the most useful parameters across multiple human subjects. Most of the previous studies use 10Hz or 20Hz lowpass butterworth filter with 5th order[8]. The (0.5Hz - 30Hz) bandpass butterworth filter and (0.3Hz - 15 Hz) bandpass butterworth filter with 3rd order also is applied in [9, 45]. To figure out the best butterworth filter, we use data from previous experiment that is used for correcting data. We used the same steps in the previous evaluation process (Cz electrode, 10-fold cross validation with RF).

After we use BSS and Butterworth filter to enhance signals, we study the effects of changing the length of the window in ErrP across multiple human subjects. According to the literature review, the researchers use 600 ms as the length of the window after the feedback onset, and sometimes get rid of the first 200 ms. They suggest that it is enough to detect ErrP signals in P300-Speller. Recently, [69] shows that a window of 200 ms to 900 ms for classification ErrPs can give the best results in the motor imagery, but not in P300-Speller. In this phase, we study how changing the length of the window impacts the ability of RF classifier across multiple human subjects, by using the same steps of the previous evaluation process. In this phase, we apply three steps:

1. Study how getting rid of the first 200 ms impacts the classification.

2. After that, we study adding the length of the window till it reaches 900ms and 1300ms on the classification.

3. Compare their results to identify the best length of the window of ErrP in P300-Speller across multiple human subjects.

After the pre-processing parameters are obtained, we apply them on other electrodes to evaluate the importance of each electrode under frontal-central regions by using RF and a new ensemble linear SVM that we propos. Here, we aim to identify the most useful electrode across multiple human subjects. This gives the researchers useful information to build a P300-Speller classifier with a few number of electrodes.

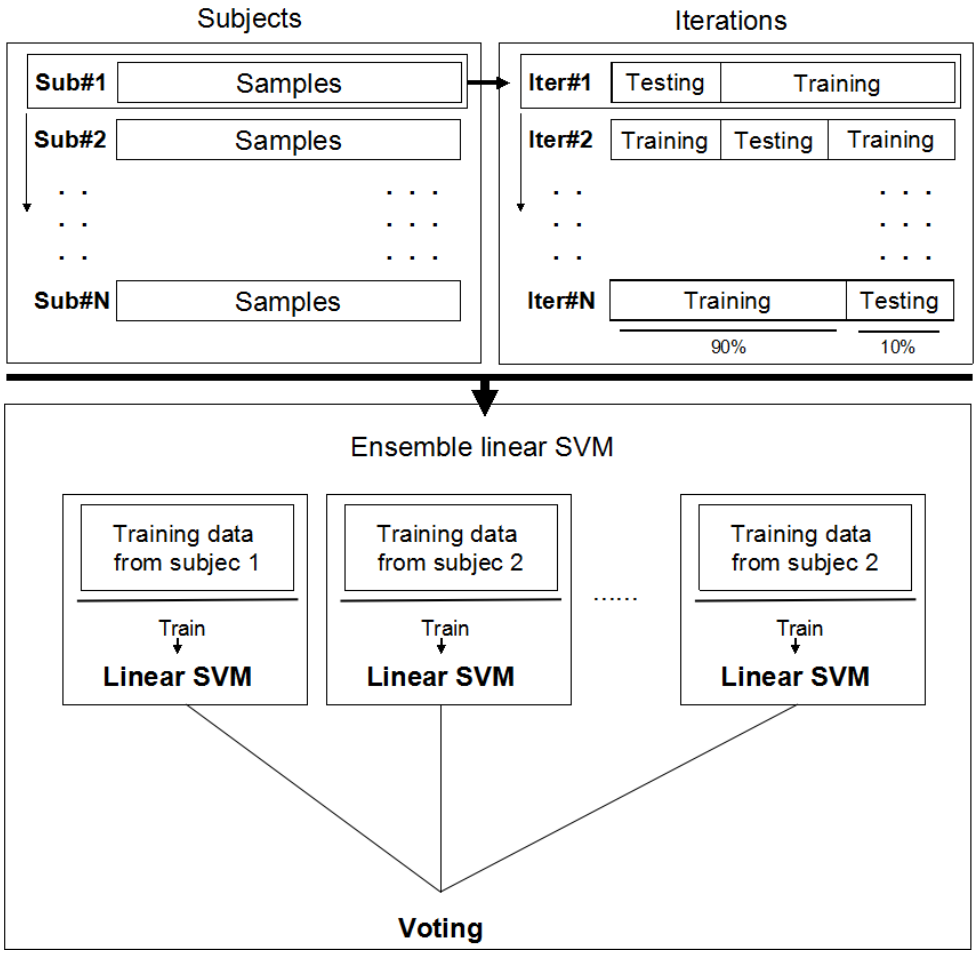


Figure 3.3: The block diagram of our new ensemble linear SVM.

The main idea of our new ensemble linear SVM is applying many linear

SVM classifiers. The basic block diagram of the ensemble linear SVM is shown in figure 3.3. We divide the collected data on each subject into training and testing data by applying 10-fold cross validation. The training data from each subject is used to build a linear SVM classifier. We use them to build the ensemble linear SVM. We choose the best score from mutli linear SVMs by voting. We use the testing data from the other subjects each time to test our ensemble linear SVM. We exclude the test data from the subject who was involved in the training as figure 3.4. We repeat the cross validation in our ensemble linear SVM 10 times and all previous steps for fronto-central electrodes. The number of the linear SVMs equals the number of the subjects. This ensemble classifier is used as a transfer learning classifier to detect ErrP among a wide array of subjects.

Like the existing work, which focuses on single subjects when detecting ErrP, we divide data collected on each subject into training and testing data by applying 10-fold cross validation, then we apply the linear SVM for each subject independently to evaluate the importance of each electrode under frontal-central regions.

In our benchmark, we have 23 electrodes that cover frontal-central regions. The number of features over these regions is high. To reduce and extract the most useful features, we use the PCA algorithm, which is described in chapter 2. We study different numbers of the components and the number of RF trees. Once the best features are obtained, we apply the linear SVM across multiple subjects to compare its result with our results to support this work. Finally, we try to add meta data to features.

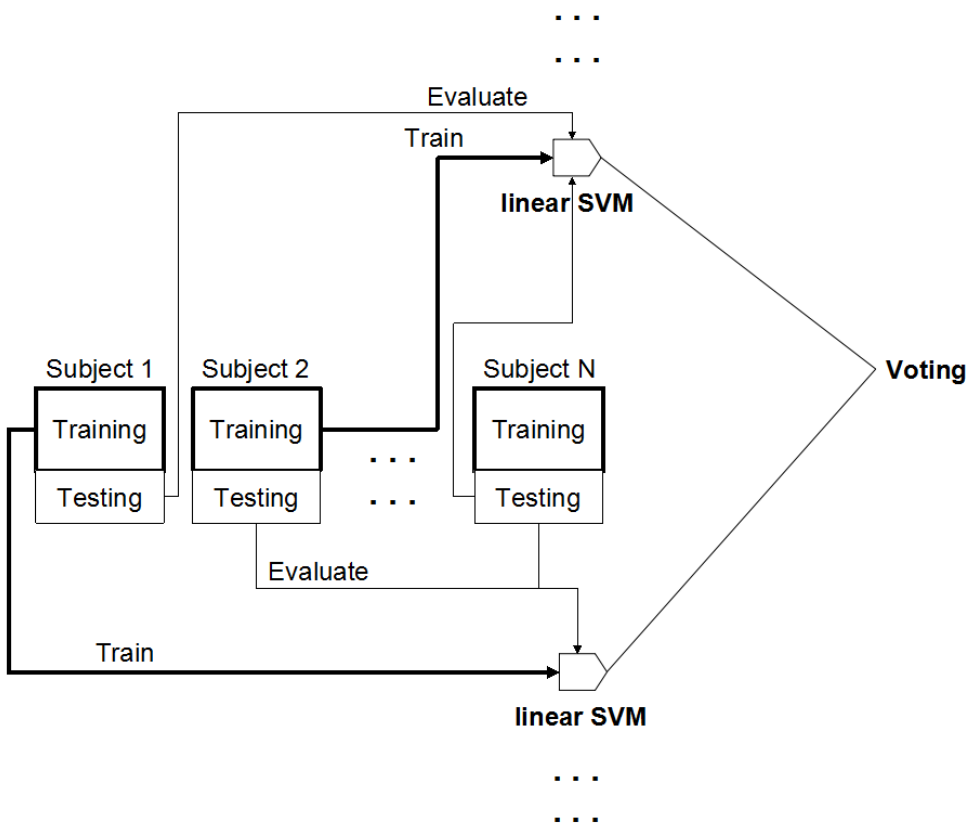


Figure 3.4: Building an ensemble linear SVM classifier and choosing the best value by voting.

Chapter 4

Experiments and Results

This chapter presents the results of the experiments that are conducted on the benchmark. We begin by presenting the results of splitting data. In the second section, we view the effects of correcting data. In the third section, the results of using butterworth filter are shown. The fourth section presents the results of changing the length of the window. The fifth section gives the results of using RF over central-frontal electrodes. In section six, we shows the results of applying the feature extractions. Finally, we compare our results of our ensemble linear SVM and RF with the results of linear SVM. All the experiments are performed by using the freely available R packages: Signal, randomForest, pROC, caret, LinearSVM and AUC. We use the same steps in the previous evaluation process to evaluate PCA.

4.1 Splitting data

In our methodology, we intentionally splited the data to get feedback period that contains the ErrP. For each subject, we extracted 340 samples. Each sample has 6240 features collected from 23 EEG electrodes (260x23=5980 data points), and EOG electrode (260 data points). The number of samples

is 340x26=8840 for all the subjects. They equal the number of the labels in our benchmark. We described our data as two dimensional matrix that contains 8840 samples x 6240 features. Each 260 features reflect the data of an electrode. We applied RF for an experiment on the raw data from Cz electrode. Figure 4.1 shows the ROC of this experiment. The AUC of the features from raw Cz data equals 61% when m_{try}=65.

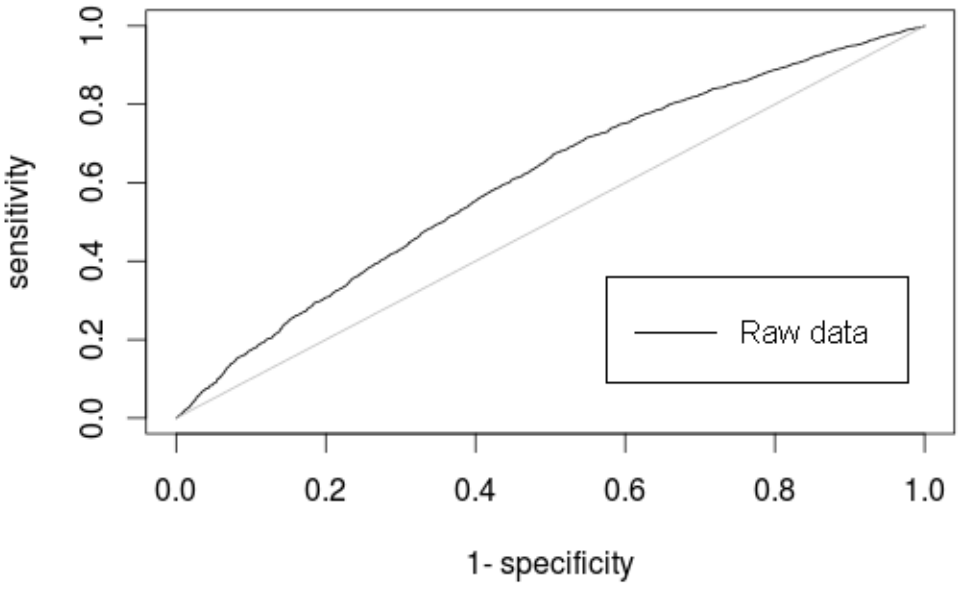
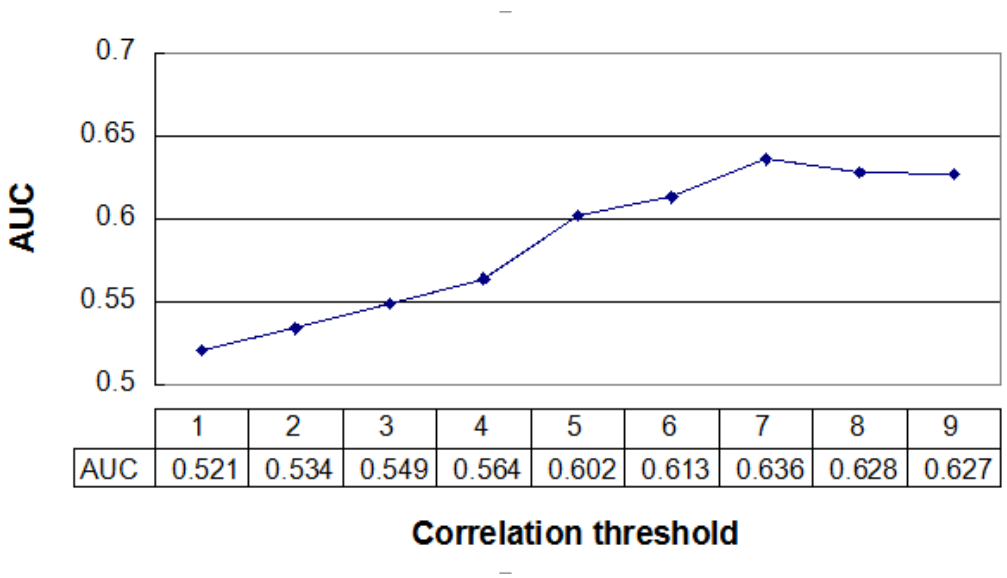


Figure 4.1: The ROC of the raw data at Cz using RF with AUC=61%.

4.2 BSS algorithm

In this section, after we got ErrP, nine experiments are performed. We tested different values of correlation thresholds to study the ability of BSS algorithm to correct the data from eye noise across multiple subjects. The values are range from 0.1 to 0.9. Figure 4.2 shows the results of these experiments. We

noticed that using BSS algorithm that is based on EOG can correct EEG data when the threshold equals 0.7 across multiple human subjects. The high value of the threshold confirms that the subjects did not blink during the feedback presentation.



	1	2	3	4	5	6	7	8	9
AUC	0.521	0.534	0.549	0.564	0.602	0.613	0.636	0.628	0.627

Figure 4.2: The AUC of BSS thresholds over Cz using RF.

We improved the result using BSS algorithm. Figure 4.3 shows the AUC for the raw data and the corrected data using RF. The AUC of the corrected data equals 63.6% when m_{try} is 130.

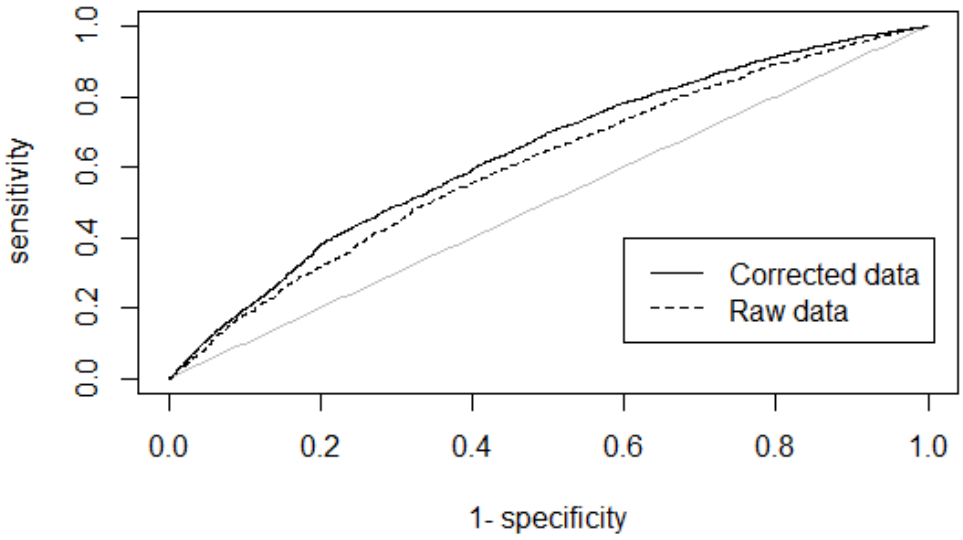


Figure 4.3: The ROC of BSS over Cz using RF with AUC=63.6%.

4.3 Filtering

After the corrected data (data without eye noise) was performed, we applied different parameters of Butterworth filter in four experiments using:

1. 5th order 10 Hz Lowpass Butterworth filter.
2. 5th order 20 Hz Lowpass Butterworth filter.
3. 3th (0.5Hz - 30 Hz) Bandpass Butterworth filter.
4. 3th (0.3Hz - 15 Hz) Bandpass Butterworth filter.

The AUC of these experiments are: 63.8% (m_{try}=65), 64.2% (m_{try} =130), 58.8%(m_{try}=130) and 62.4% (m_{try}=130) respectively. Figure 4.4 shows the ROC of these experiments. It can be clearly noticed that the best parameters

of butterworth filter are 5th order 20 Hz low Butterworth filter on all the subject. Figure 4.5 shows the AUC difference between the corrected data and the data after applying the butterworth filter.

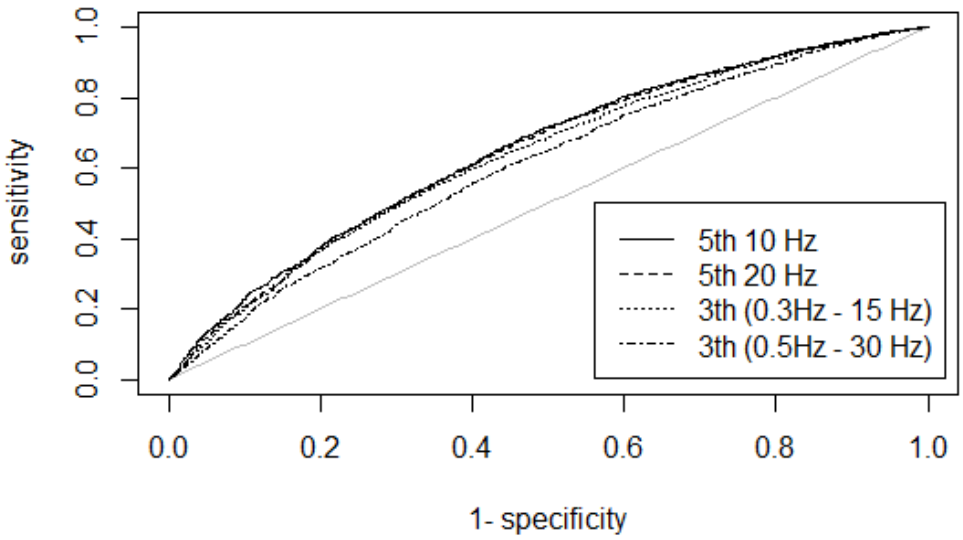


Figure 4.4: The ROC of Butterworth filters using RF with different orders and frequencies.

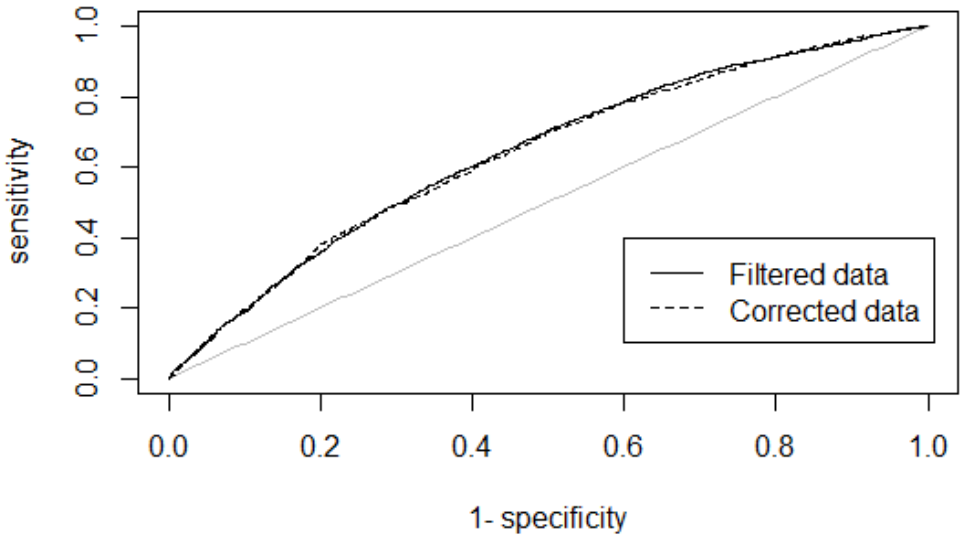


Figure 4.5: The ROC difference between the filtered and corrected data using RF with AUC=63.6% and 64.2%.

4.4 The length of the window

After we applied the butterworth filter, we studied the impact of the change of the length of the window. We intentionally applied two experiments for the length that the researchers prefer (1ms-600ms and 200ms-600ms). In these experiments, we also aimed to study the impact of the first 200 ms. The result of the experiments are shown in Figure 4.6. The AUCs are 63.5% (m_{try}=60) and 62.8% (m_{try}=81) respectively. We noticed that the first 200 ms have important features, and we can not avoid this part of the window.

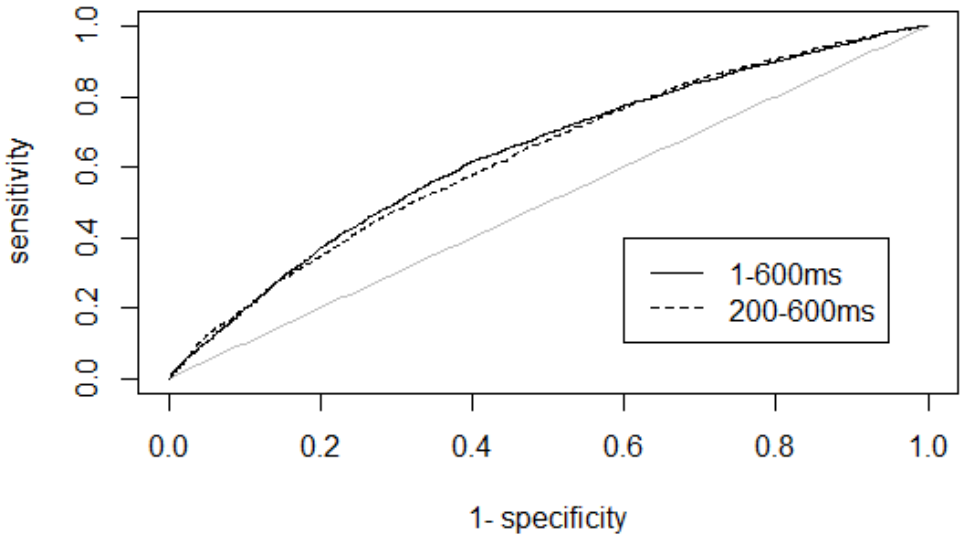


Figure 4.6: The ROC curve over the first 200 ms using RF with AUC=63.5% and 62.8%.

Recently, [69] shows that the window of 200 ms to 900 ms for classification ErrPs gives the best results in the motor-imagery. Depending on the previous experiments, we added the first 200ms to this length and applied it with the previous parameters. Figure 4.7 shows the difference between the 1ms-900ms and the feedback period. The AUC of these experiments are 64.3% (m_{try}=90) and 64.3% (m_{try}=130) respectively. We noticed that the optimal length of the window on all the subjects is from 1ms to 900ms and we reduced the number of the features from 260(1300 ms) to 180(900ms) on each electrode.

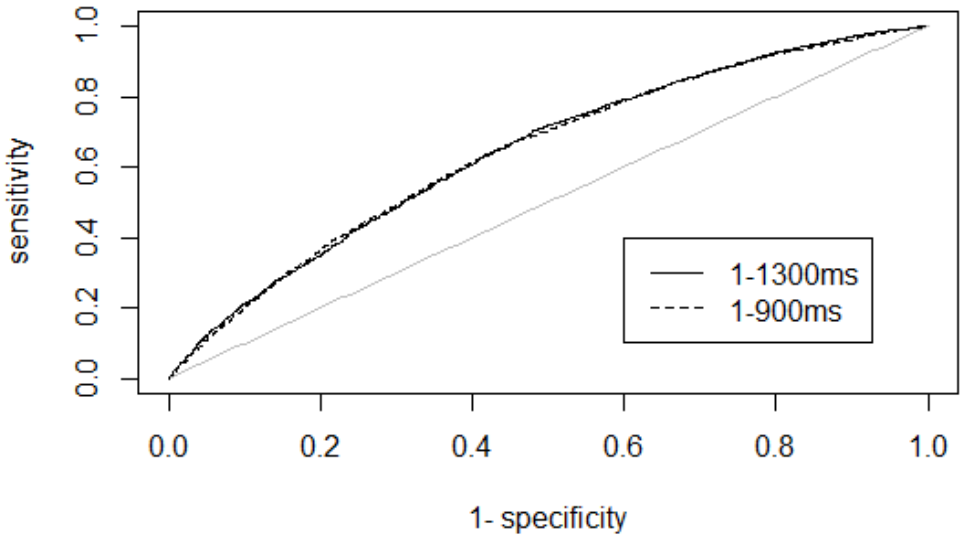


Figure 4.7: The ROC curve over the length of the windows 1-1300ms and 1-900ms using RF.

Depending on the previous pre-processing steps, figure 4.8 shows the AUC difference between the raw data and the data after prepossessing. We noticed the importance of the pre-processing steps.

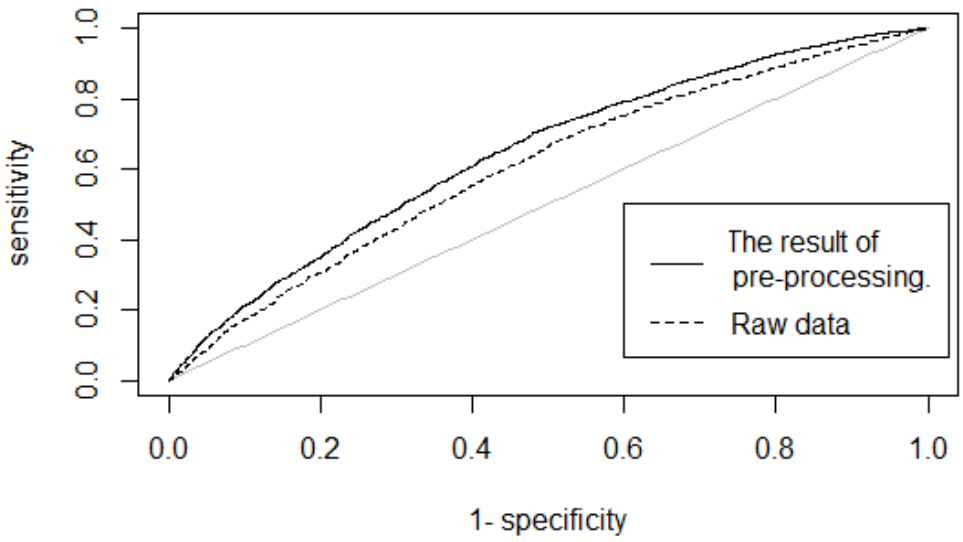


Figure 4.8: The ROC of data before and after pre-processing steps using RF with AUC=61.6% and 64.3%.

4.5 Selecting electrodes

According to the parameters that we found in the previous experiments, we applied RF and our ensemble linear SVM for each electrode over the frontal-central regions to evaluate the importance of each one across multiple human subjects. Besides, our ensemble linear SVM shows the result of applying the testing data from the other subject on each electrode. The result of our linear SVM are shown in tables 4.1, 4.2 and 4.3.

The results of 23 experiments using RF on each electrode across multiple human subjects are described in table 4.4.

Table 4.1: The AUC of our ensemble linear SVM for different electrodes under multiple human subjects.

Iter#	**F7**	**F5**	**F3**	**F1**	**Fz**	**F2**	**F4**	**F6**
1	0.661	0.671	**0.68**	0.659	0.663	0.655	0.636	0.662
2	0.651	0.668	0.663	0.651	0.647	0.649	0.643	0.646
3	0.641	0.652	0.653	0.648	0.642	0.641	0.629	0.631
4	0.653	0.657	0.657	0.654	0.650	0.649	0.634	0.643
5	0.653	0.655	0.656	0.650	0.646	0.648	0.634	0.641
6	0.659	0.663	0.664	0.658	0.654	0.656	0.640	0.648
7	0.661	0.666	0.666	0.659	0.654	0.657	0.639	0.649
8	0.658	0.665	0.665	0.658	0.649	0.653	0.635	0.646
9	0.657	0.664	0.663	0.655	0.645	0.648	0.632	0.644
10	0.658	0.665	0.664	0.656	0.646	0.649	0.632	0.646

Table 4.2: The AUC of our ensemble linear SVM for different electrodes under multiple human subjects.

Iter#	**F8**	**FC5**	**FC3**	**FC1**	**FCz**	**FC2**	**FC4**	**FC6**
1	0.674	0.671	0.661	0.642	0.645	0.613	0.625	0.657
2	0.661	0.667	0.644	0.642	0.638	0.624	0.621	0.642
3	0.651	0.653	0.639	0.627	0.627	0.612	0.617	0.623
4	0.660	0.658	0.645	0.637	0.633	0.617	0.625	0.635
5	0.659	0.652	0.641	0.637	0.628	0.617	0.620	0.633
6	0.666	0.661	0.649	0.646	0.639	0.625	0.624	0.640
7	0.665	0.662	0.649	0.646	0.639	0.625	0.619	0.637
8	0.661	0.660	0.647	0.646	0.636	0.624	0.619	0.634
9	0.657	0.656	0.642	0.642	0.632	0.620	0.618	0.632
10	0.657	0.657	0.641	0.642	0.632	0.622	0.615	0.632

Table 4.3: The AUC of our ensemble linear SVM for different electrodes under multiple human subjects.

Iter#	**C5**	**C3**	**C1**	**Cz**	**C2**	**C4**	**C6**
1	0.658	0.648	0.651	0.613	0.622	0.616	0.598
2	0.643	0.647	0.647	0.634	0.628	0.626	0.606
3	0.633	0.635	0.629	0.62	0.611	0.610	0.595
4	0.646	0.642	0.637	0.627	0.621	0.615	0.603
5	0.645	0.640	0.634	0.624	0.618	0.607	0.594
6	0.653	0.648	0.642	0.631	0.624	0.614	0.598
7	0.654	0.649	0.640	0.632	0.624	0.613	0.594
8	0.652	0.649	0.639	0.634	0.623	0.610	0.591
9	0.648	0.645	0.635	0.629	0.618	0.607	0.589
10	0.649	0.646	0.636	0.630	0.619	0.609	0.591

Table 4.4: The AUC for RF across multiple human subjects on each electrode

Electrode	**AUC**	**Electrode**	**AUC**
C1	0.64	F6	0.62
C2	0.65	F7	0.57
C3	0.62	F8	0.61
C4	0.65	FC1	0.63
C5	0.61	FC2	0.65
C6	0.64	FC3	0.63
Cz	0.64	FC4	0.64
F1	0.64	FC5	0.62
F2	0.64	FC6	0.65
F3	0.62	FCz	0.63
F4	0.63	Fz	0.64
F5	0.58		

The result of our ensemble linear SVM is 68% and the result of RF equals 65% across multiple human subjects. We notice that the result of our ensemble linear SVM outperforms the RF when we use one electrode across multiple human subjects.

Like the existing work [46], which focus on single subjects when detecting ErrP, we divide data collected on each subject into training and testing data by applying 10-fold cross validation, then we apply a linear SVM classifier for each subject independently to evaluate the importance of each electrode under frontal-central regions. The result of these experiments are shown in appendix 2.

The average of AUCs of the linear SVM is 87% at C2 electrode using the testing data from the same subject independently. We noticed that the C2 or F3 electrode are the best electrodes for detecting ErrP across one subject. Based on the results, we suggest using them instead of the Cz electrode.

4.6 Feature extractions

After the pre-processing steps, we can perform features extraction to improve our results. It is an important step to extract the most useful features. We used features over frontal-central regions. In this phase, we applied a variety of features extraction methods like Autoregressive (AR), ICA ... and we found that PCA is the best one. After that, we test a different number of components to figure out the best parameter of PCA. Figure 4.9 shows the AUCs of RF with running different number of components. We noticed that the results are improved when the number of the components equals 30 and the AUC of RF is 71% when m_{try}=30.

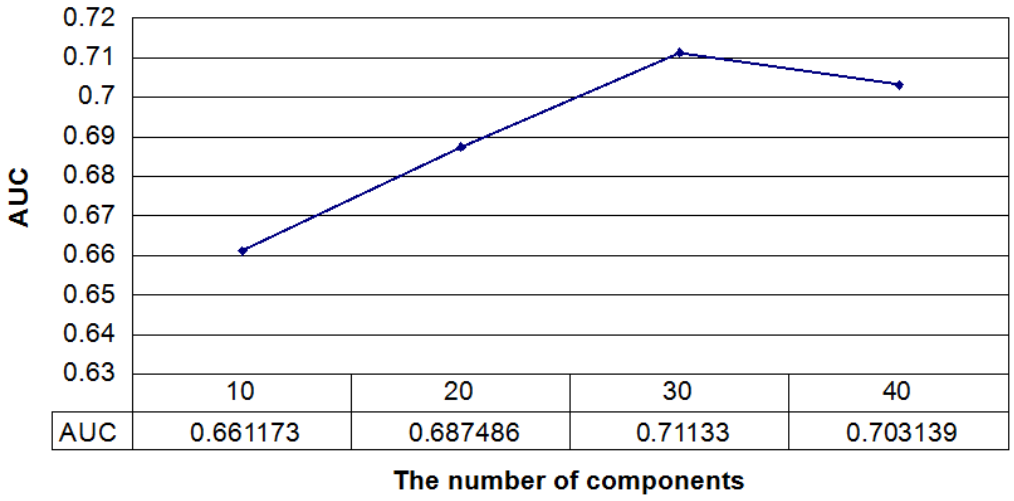


Figure 4.9: The AUC of PCA components over frontal-central regions.

4.7 Comparison of our results with Linear SVM

We used the parameters that we have found in the pre-processing and feature extraction steps to compare our results of our ensemble linear SVM and RF

with the results of linear SVM, which is commonly used in ErrP to support this study. The AUC of our ensemble linear SVM is 68% and the AUC of RF equals 71% when m_{try}=30 with extracted features using PCA across multiple human subjects. Thus, we apply linear SVM on the data of all the subjects. The AUC of the linear SVM equals 60%. Our result of RF outperforms the result of the linear SVM and our ensemble linear SVM. Figure 4.10 shows the AUC difference between RF and linear SVM. This result supports our methodology of work.

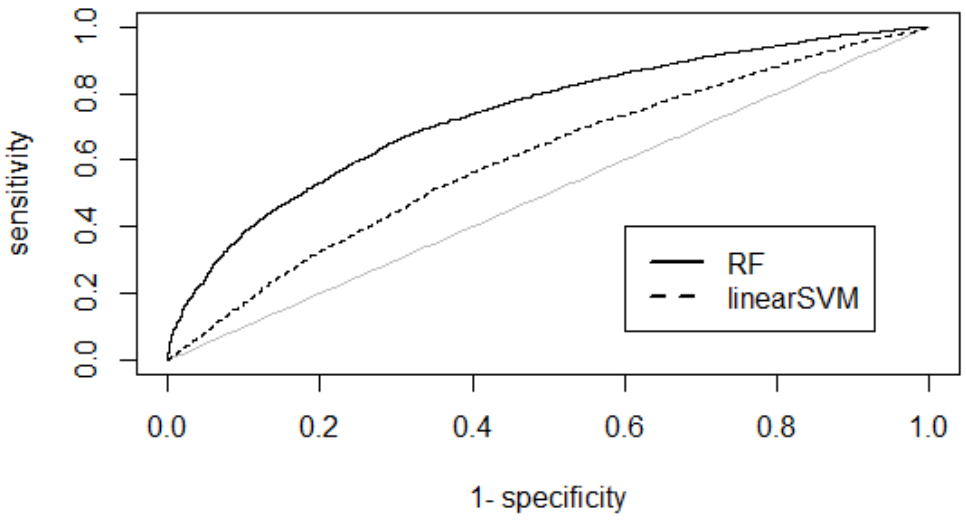
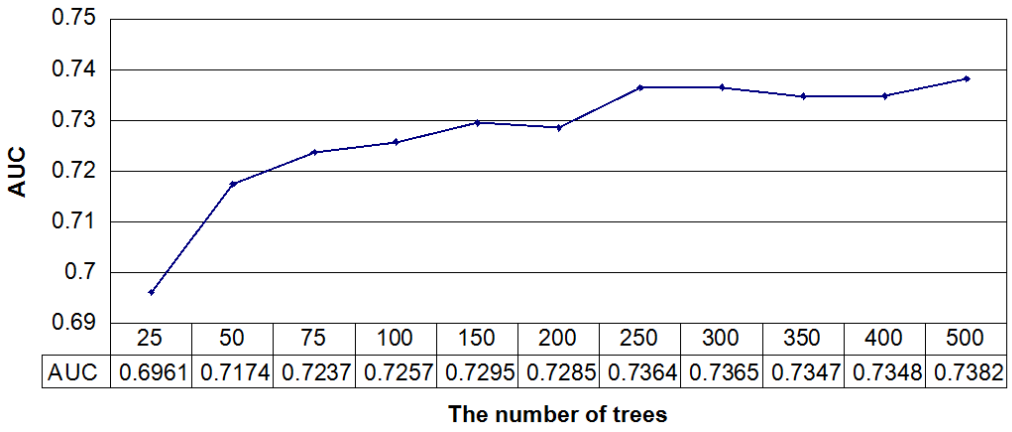


Figure 4.10: The comparison between linear SVM and RF without any meta data with the AUC=71% and 60%.

In a different way, we tried to change the number of trees in RF. Figure 4.11 shows the AUCs of the number of different trees. We notice that when we increase the number of trees, the AUC of RF is increased. The AUC stays after 300 and the result almost remain the same with AUC equals 74% when

m_{try}=30.



	25	50	75	100	150	200	250	300	350	400	500
AUC	0.6961	0.7174	0.7237	0.7257	0.7295	0.7285	0.7364	0.7365	0.7347	0.7348	0.7382

Figure 4.11: The AUC over difference RF trees.

Our results confirms that the RF can cope with the heterogeneity among different subjects and our ensemble linear SVM can figure out the best electrode.

In a different way, we tried to add some meta data to the PCA components. We added the type of the spelling mode as a meta feature. Figure 4.12 shows the AUC difference of the RF before and after adding the meta data to the PCA components. The AUC of RF after adding meta data equals 78% when m_{try}=32. Again we used the linear SVM to compare its result with our result of RF. Figure 4.13 shows the AUC difference between the linear SVM and the RF. RF outperforms the linear SVM and this supports our methodology. Finally, we detected the ErrP with 78% AUC using 32 features.

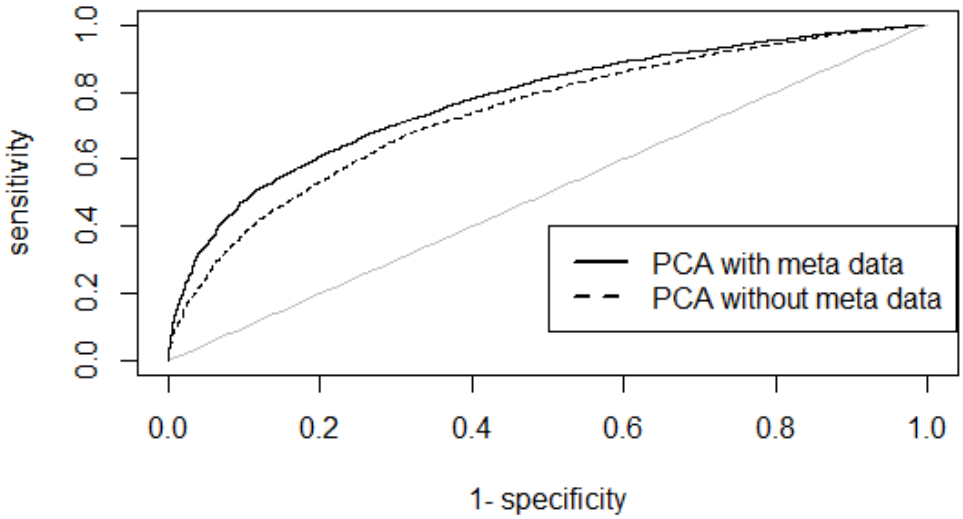


Figure 4.12: The comparison between PCA components with and without meta data with AUC=74% and 78%.

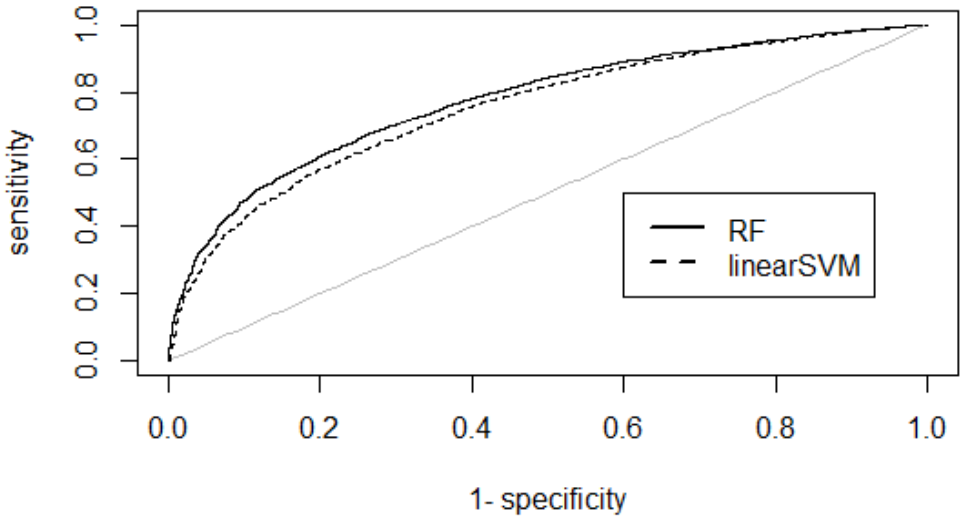


Figure 4.13: The comparison between RF and linear SVM with meta data. The AUC=76% and 78%.

Chapter 5

Conclusion and Future Work

In this chapter, we conclude the work performed in this thesis by summarising our contributions. In second section, we introduce two block diagrams of how a real-time detector works. In the third section, we will also suggest some future research directions that can provide the others with new steps.

5.1 Conclusion

In this thesis, we proposed new ensemble methods to detect the ErrP using two main classifiers: RF and ensemble linear SVM. We applied the RF machine learning technique to figure out of the best of the parameters of the pre-processing steps that can work across multiple subjects. We corrected the data by using BBS algorithm to avoid the noise often produced by the eye movements or the eye blinks. We found that the best value of the correlation threshold equals 0.7 across multiple subjects. After that, we noticed that 20Hz lowpass butterworth filter with 5th order can improve the signal-to-noise ratio without removing the useful information of ErrP signals. In the next step, we tried to find the best length of the window of ErrP. We found that the length of the window is between 1ms to 900ms after the feedback

onset is the optimal one. We evaluated the importance of each electrode under frontal-central regions. Our ensemble linear SVM outperforms the RF for each electrode in the frontal-central regions. We suggested using F3 as a main electrode instead of the Cz electrode. We obtained 68% AUC by using our ensemble linear SVM with the data at F3. We showed that it is possible to extract the most useful features from centro-frontal electrodes by using PCA with 30 components. To support our work, we compared our results with Linear SVM. We noticed that RF can detect the ErrP signals, because it can cope with the heterogeneity among different subjects. We added the type of spelling as meta data. Finally, we obtained 78% Area Under Curve (AUC) by using RF with features that is extracted from the centro-frontal electrodes across multiple subjects.

5.2 A real-time detector

In this section, we introduce two block diagrams of how a real-time detector works. Based on our methodology, there are two main block diagrams of how a real-time detector works. First, using our ensemble linear SVM as a main classifier with the data from the F3 electrode, and using RF for the electrodes over frontal-central regions, which generate ErrP[21].

Figure 5.1 shows the block diagram of how a real-time detector works when we use RF as main classifier. The first step focuses on extracting the feedback period that contains ErrP from EEG. In this step, we select the electrodes over frontal-central regions, which generate ErrP[21] and we ignore other data. We select the window is between 1ms to 900ms after the feedback onset.

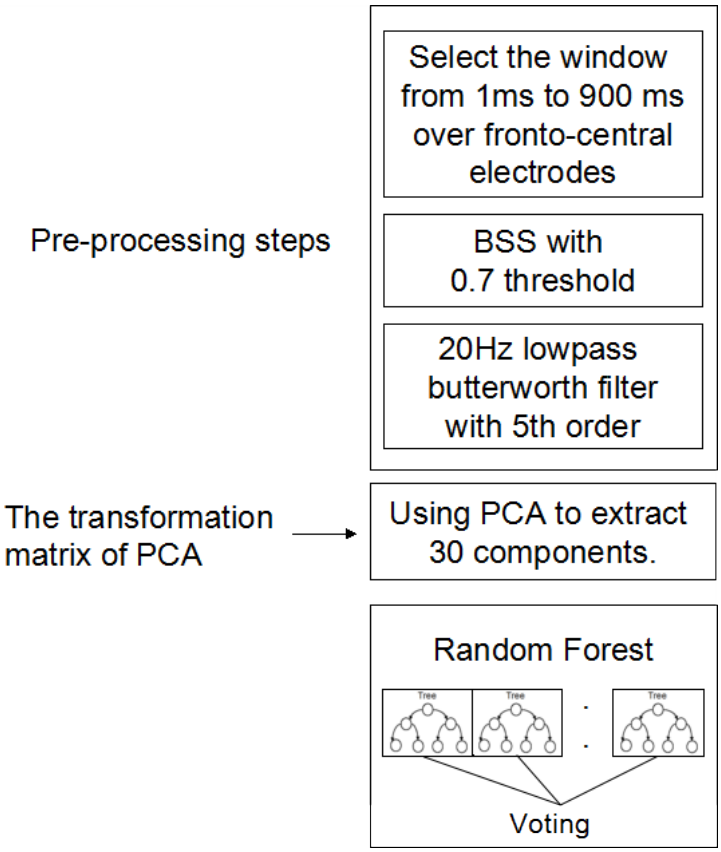


Figure 5.1: Block diagram of how a real-time detector works when using RF

Depending on the outcome of the first portion, we apply the pre-processing methods over these electrodes. We use BSS algorithm to correct data with the correlation threshold value equals 0.7. After applying the BSS algorithm, we apply 20Hz lowpass butterworth filter with 5th order in order to improve the signal-to-noise ratio. After we use BSS and Butterworth filter to enhance signals, we use the transformation matrix of PCA components to transfer the original data-set onto 30 components, then we apply the trained RF classifier to classify the components.

Figure 5.2 shows another block diagram of how a real-time detector when we use our ensemble linear SVM as main classifier. THe first step we extract

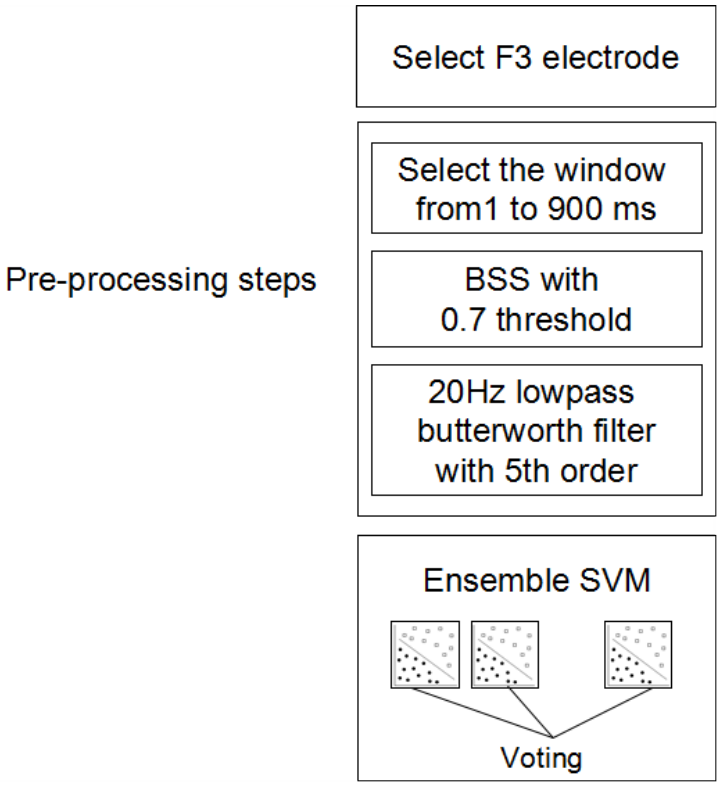


Figure 5.2: Block diagram of how a real-time detector works when using our ensemble linear SVM

the window is between 1ms to 900ms after the feedback period that contains ErrP from F3 electrode. Depending on the outcome of the first portion, we apply the same methods which are shown in Figure 5.1: BSS algorithm and lowpass butterworth filter. Finally, we apply the trained our ensemble linear SVM to classify the data after applying pre-processing methods.

5.3 Future Work

In this thesis, we have succeeded in implementing a classifier that can be applied across multiple subjects without any need for prior training sessions. This classifier is to be used in the P300-speller. This classifier could be connected and tested on a real mechanical device in the future. Also, rather than using RF and our ensemble linear SVM classifiers to detect the ErrP,

it may be a good idea to use these classifiers to correct the target letter.

Appendices

Appendix A

Appendix

A.1 The description of the subjects' sessions

Table A.1: The description of the subjects' sessions

Subject name	Session #	Target	%	Non target	%
S01	1	50	0.83	10	0.17
S01	2	49	0.82	11	0.18
S01	3	45	0.75	15	0.25
S01	4	49	0.82	11	0.18
S01	5	80	0.80	20	0.20
S02	1	50	0.83	10	0.17
S02	2	43	0.72	17	0.28
S02	3	36	0.60	24	0.40
S02	4	35	0.58	25	0.42
S02	5	56	0.56	44	0.44
S03	1	31	0.52	29	0.48
S03	2	30	0.50	30	0.50
S03	3	20	0.33	40	0.67
S03	4	24	0.40	36	0.60
S03	5	36	0.36	64	0.64
S04	1	52	0.87	8	0.13
S04	2	53	0.88	7	0.12
S04	3	46	0.77	14	0.23
S04	4	50	0.83	10	0.17
S04	5	72	0.72	28	0.28
S05	1	25	0.42	35	0.58
S05	2	34	0.57	26	0.43
S05	3	28	0.47	32	0.53
S05	4	24	0.40	36	0.60
S05	5	31	0.31	69	0.69
S06	1	59	0.98	1	0.02
S06	2	57	0.95	3	0.05
S06	3	55	0.92	5	0.08
S06	4	56	0.93	4	0.07
S06	5	89	0.89	11	0.11
S07	1	56	0.93	4	0.07
S07	2	53	0.88	7	0.12
S07	3	57	0.95	3	0.05
S07	4	53	0.88	7	0.12
S07	5	88	0.88	12	0.12
S08	1	46	0.77	14	0.23
S08	2	48	0.80	12	0.20
S08	3	45	0.75	15	0.25
S08	4	42	0.70	18	0.30
S08	5	65	0.65	35	0.35
S09	1	53	0.88	7	0.12
S09	2	48	0.80	12	0.20
S09	3	49	0.82	11	0.18
S09	4	57	0.95	3	0.05
S09	5	79	0.79	21	0.21

Table A.2: The description of the subjects' sessions

Subject name	Session #	Target	%	Non target	%
S10	1	53	0.88	7	0.12
S10	2	55	0.92	5	0.08
S10	3	55	0.92	5	0.08
S10	4	58	0.97	2	0.03
S10	5	88	0.88	12	0.12
S11	1	46	0.77	14	0.23
S11	2	42	0.70	18	0.30
S11	3	42	0.70	18	0.30
S11	4	39	0.65	21	0.35
S11	5	56	0.56	44	0.44
S12	1	41	0.68	19	0.32
S12	2	41	0.68	19	0.32
S12	3	33	0.55	27	0.45
S12	4	33	0.55	27	0.45
S12	5	41	0.41	59	0.59
S13	1	38	0.63	22	0.37
S13	2	30	0.50	30	0.50
S13	3	25	0.42	35	0.58
S13	4	35	0.58	25	0.42
S13	5	50	0.50	50	0.50
S14	1	37	0.62	23	0.38
S14	2	44	0.73	16	0.27
S14	3	42	0.70	18	0.30
S14	4	42	0.70	18	0.30
S14	5	50	0.50	50	0.50
S15	1	58	0.97	2	0.03
S15	2	58	0.97	2	0.03
S15	3	57	0.95	3	0.05
S15	4	56	0.93	4	0.07
S15	5	91	0.91	9	0.09
S16	1	44	0.73	16	0.27
S16	2	41	0.68	19	0.32
S16	3	38	0.63	22	0.37
S16	4	35	0.58	25	0.42
S16	5	53	0.53	47	0.47
S17	1	54	0.90	6	0.10
S17	2	43	0.72	17	0.28
S17	3	36	0.60	24	0.40
S17	4	34	0.57	26	0.43
S17	5	59	0.59	41	0.41
S18	1	51	0.85	9	0.15
S18	2	46	0.77	14	0.23
S18	3	46	0.77	14	0.23
S18	4	45	0.75	15	0.25
S18	5	73	0.73	27	0.27

Table A.3: The description of the subjects' sessions

Subject name	**Session #**	**Target**	**%**	**Non target**	**%**
S19	1	44	0.73	16	0.27
S19	2	45	0.75	15	0.25
S19	3	43	0.72	17	0.28
S19	4	38	0.63	22	0.37
S19	5	54	0.54	46	0.46
S20	1	43	0.72	17	0.28
S20	2	46	0.77	14	0.23
S20	3	42	0.70	18	0.30
S20	4	41	0.68	19	0.32
S20	5	64	0.64	36	0.36
S21	1	58	0.97	2	0.03
S21	2	58	0.97	2	0.03
S21	3	56	0.93	4	0.07
S21	4	55	0.92	5	0.08
S21	5	85	0.85	15	0.15
S22	1	57	0.95	3	0.05
S22	2	56	0.93	4	0.07
S22	3	54	0.90	6	0.10
S22	4	55	0.92	5	0.08
S22	5	91	0.91	9	0.09
S23	1	42	0.70	18	0.30
S23	2	36	0.6	24	0.4
S23	3	39	0.65	21	0.35
S23	4	43	0.72	17	0.28
S23	5	61	0.61	39	0.39
S24	1	54	0.90	6	0.10
S24	2	51	0.85	9	0.15
S24	3	46	0.77	14	0.23
S24	4	37	0.62	23	0.38
S24	5	51	0.51	49	0.49
S25	1	42	0.70	18	0.30
S25	2	43	0.72	17	0.28
S25	3	26	0.43	34	0.57
S25	4	32	0.53	28	0.47
S25	5	54	0.54	46	0.46
S26	1	35	0.58	25	0.42
S26	2	37	0.62	23	0.38
S26	3	30	0.50	30	0.50
S26	4	39	0.65	21	0.35
S26	5	40	0.40	60	0.60
		6261	**0.70**	**2579**	**0.30**

A.2 The results for applying linear SVMs on each subject independently

Table A.4: The AUC of different electrodes using a linear SVM on each subject

Sub #	**F7**	**F5**	**F3**	**F1**	**Fz**	**F2**	**F4**	**F6**
1	0.74	0.77	0.74	0.81	0.89	0.88	0.82	0.81
2	0.83	0.91	0.86	0.85	0.84	0.88	0.88	0.84
3	0.84	0.88	0.94	0.95	0.94	0.92	0.93	0.98
4	0.79	0.74	0.67	0.66	0.69	0.67	0.75	0.83
5	0.71	0.75	0.84	0.86	0.82	0.88	0.83	0.78
6	0.84	0.88	0.84	0.83	0.87	0.88	0.92	0.88
7	0.83	0.83	0.94	0.98	0.99	0.98	0.93	0.76
8	0.88	0.81	0.85	0.86	0.9	0.96	0.97	0.95
9	0.91	0.98	0.95	0.99	1	0.99	0.98	0.99
10	1	1	1	1	0.99	0.97	0.95	0.95
11	0.67	0.74	0.75	0.78	0.8	0.85	0.87	0.85
12	0.89	0.86	0.76	0.74	0.72	0.7	0.74	0.74
13	0.82	0.88	0.93	0.91	0.92	0.85	0.85	0.84
14	0.82	0.84	0.81	0.91	0.93	0.91	0.89	0.75
15	1	1	1	1	1	1	0.98	0.97
16	0.67	0.74	0.75	0.74	0.75	0.76	0.9	0.75
17	0.77	0.83	0.81	0.85	0.85	0.84	0.85	0.83
18	0.79	0.78	0.75	0.72	0.77	0.74	0.75	0.73
19	0.78	0.93	0.87	0.83	0.82	0.81	0.82	0.78
20	0.64	0.67	0.65	0.76	0.76	0.78	0.79	0.67
21	0.87	0.96	0.89	0.96	0.88	0.88	0.97	1
22	0.67	0.77	0.76	0.78	0.86	0.88	0.78	0.82
23	0.7	0.74	0.75	0.86	0.85	0.87	0.81	0.8
24	0.69	0.84	0.95	1	0.94	0.89	0.82	0.77
25	0.61	0.6	0.64	0.6	0.6	0.62	0.67	0.57
26	0.69	0.64	0.64	0.58	0.69	0.65	0.64	0.64
Avg.	**0.79**	**0.82**	**0.82**	**0.84**	**0.85**	**0.85**	**0.85**	**0.82**

Table A.5: The AUC of different electrodes using a linear SVM on each subject

Sub #	**F8**	**FC5**	**FC3**	**FC1**	**FCz**	**FC2**	**FC4**	**FC6**
1	0.57	0.77	0.75	0.86	0.85	0.88	0.9	0.89
2	0.9	0.88	0.86	0.83	0.89	0.86	0.86	0.83
3	0.85	0.9	0.92	0.9	0.92	0.93	0.93	0.91
4	0.82	0.74	0.71	0.66	0.7	0.72	0.78	0.77
5	0.82	0.8	0.83	0.88	0.85	0.87	0.85	0.79
6	0.89	0.87	0.82	0.79	0.82	0.88	0.92	0.93
7	0.91	0.84	0.96	0.98	0.98	0.98	1	0.97
8	0.94	0.82	0.88	0.88	0.94	0.95	0.95	0.96
9	0.97	0.97	0.99	0.98	1	1	0.99	0.98
10	1	0.97	1	1	1	0.98	1	1
11	0.73	0.76	0.78	0.79	0.85	0.89	0.88	0.82
12	0.78	0.76	0.77	0.78	0.76	0.77	0.75	0.78
13	0.84	0.9	0.9	0.87	0.85	0.81	0.83	0.88
14	0.82	0.78	0.84	0.88	0.88	0.89	0.87	0.83
15	0.97	1	0.98	1	1	0.98	0.98	0.97
16	0.71	0.76	0.74	0.76	0.77	0.84	0.86	0.84
17	0.85	0.89	0.9	0.93	0.91	0.87	0.85	0.91
18	0.77	0.77	0.73	0.75	0.78	0.76	0.76	0.76
19	0.57	0.81	0.88	0.85	0.8	0.82	0.86	0.8
20	0.7	0.65	0.75	0.78	0.72	0.79	0.74	0.72
21	0.89	0.94	0.94	0.99	0.98	0.94	1	0.98
22	0.77	0.82	0.76	0.79	0.85	0.92	0.87	0.85
23	0.77	0.75	0.78	0.78	0.88	0.78	0.79	0.76
24	0.81	0.84	0.94	0.98	1	1	0.79	0.84
25	0.6	0.7	0.69	0.67	0.61	0.71	0.74	0.67
26	0.74	0.58	0.64	0.53	0.71	0.65	0.72	0.68
Average	**0.81**	**0.82**	**0.84**	**0.84**	**0.86**	**0.86**	**0.86**	**0.85**

Table A.6: The AUC of different electrodes using a linear SVM on each subject

Sub #	**C5**	**C3**	**C1**	**Cz**	**C2**	**C4**	**C6**
1	0.77	0.82	0.82	0.79	0.86	0.81	0.83
2	0.86	0.86	0.84	0.83	0.9	0.83	0.82
3	0.88	0.9	0.91	0.89	0.91	0.93	0.93
4	0.76	0.73	0.68	0.73	0.77	0.76	0.87
5	0.84	0.85	0.84	0.86	0.87	0.85	0.88
6	0.88	0.88	0.78	0.83	0.86	0.94	0.93
7	0.91	0.94	0.96	0.98	0.98	0.98	0.98
8	0.95	0.89	0.94	0.93	0.97	1	0.97
9	1	0.98	0.99	1	0.99	0.98	0.95
10	1	0.97	0.97	1	1	1	1
11	0.8	0.81	0.81	0.81	0.81	0.85	0.81
12	0.82	0.75	0.79	0.75	0.77	0.82	0.74
13	0.87	0.85	0.88	0.88	0.91	0.9	0.88
14	0.8	0.84	0.91	0.84	0.87	0.85	0.9
15	1	1	1	0.98	0.97	0.97	0.94
16	0.76	0.81	0.83	0.85	0.77	0.86	0.73
17	0.82	0.94	0.91	0.93	0.9	0.88	0.9
18	0.79	0.97	0.78	0.76	0.8	0.74	0.88
19	0.79	0.86	0.85	0.85	0.89	0.86	0.75
20	0.72	0.79	0.88	0.73	0.82	0.8	0.72
21	1	0.96	0.96	0.97	1	1	0.96
22	0.7	0.76	0.82	0.93	0.94	0.94	0.91
23	0.69	0.78	0.86	0.9	0.84	0.8	0.77
24	0.81	0.78	0.81	0.84	0.94	0.73	0.7
25	0.64	0.72	0.77	0.71	0.74	0.73	0.72
26	0.61	0.56	0.65	0.68	0.73	0.63	0.64
Average	**0.83**	**0.85**	**0.86**	**0.86**	**0.88**	**0.86**	**0.85**

Bibliography

[1] Moving away from error-related potentials to achieve spelling correction in p300 spellers. *Neural Systems and Rehabilitation Engineering, IEEE Transactions on*, PP(99):1–1, 2014.

[2] Setare Amiri, Reza Fazel-Rezai, and Vahid Asadpour. Brain computer interface and its type - a study. *A Review of Hybrid Brain-Computer Interface Systems*, 2013.

[3] HS Anupama, NK Cauvery, and GM Lingaraju. Real-time eeg based object recognition system using brain computer interface. In *Contemporary Computing and Informatics (IC3I), 2014 International Conference on*, pages 1046–1051. IEEE, 2014.

[4] Sergio Bermejo. Finite sample effects of the fast ica algorithm. *Neurocomputing*, 71(1):392–399, 2007.

[5] P Bhuvaneswari and J Satheesh Kumar. Support vector machine technique for eeg signals. *Group*, 1:H1, 2013.

[6] Leo Breiman. Random forests. *Machine learning*, 45(1):5–32, 2001.

[7] Stephen Butterworth. On the theory of filter amplifiers. *Wireless Engineer*, 7(6):536–541, 1930.

[8] Ricardo Chavarriaga, Aleksander Sobolewski, and José del R Millán. Errare machinale est: the use of error-related potentials in brain-machine interfaces. *Frontiers in neuroscience*, 8, 2014.

[9] Adrien Combaz, Nikolay Chumerin, Nikolay V Manyakov, Arne Robben, Johan AK Suykens, and Marc M Van Hulle. Towards the detection of error-related potentials and its integration in the context of a p300 speller brain–computer interface. *Neurocomputing*, 80:73–82, 2012.

[10] Corinna Cortes and Vladimir Vapnik. Support-vector networks. *Machine learning*, 20(3):273–297, 1995.

[11] Bernardo Dal Seno, Matteo Matteucci, and Luca Mainardi. Online detection of p300 and error potentials in a bci speller. *Computational intelligence and neuroscience*, 2010:11, 2010.

[12] Niva Das, Aurobinda Routray, Pradipta Kishore Dash, and D India. Ica methods for blind source separation of instantaneous mixtures: A case study. *Neural Information process. Letters and Reviews*, 11(11):225–46, 2007.

[13] Lieven De Lathauwer, Bart De Moor, and Joos Vandewalle. An introduction to independent component analysis. *Journal of chemometrics*, 14(3):123–149, 2000.

[14] Antonio Dourado, Bruno Direito, César Teixeira, João Duarte, Bjoern Schelter, Michel Quyen, Andreas Bonhage, and Francisco Sales. Feature selection in high dimensional eeg feature spaces for epileptic seizure prediction. In *World Congress*, volume 18, pages 6206–6211, 2011.

[15] Martin Ester, Hans-Peter Kriegel, Jörg Sander, and Xiaowei Xu. A density-based algorithm for discovering clusters in large spatial databases with noise. In *Kdd*, volume 96, pages 226–231, 1996.

[16] Michael Falkenstein, Joachim Hohnsbein, Jörg Hoormann, and Ludger Blanke. Effects of crossmodal divided attention on late erp components. ii. error processing in choice reaction tasks. *Electroencephalography and clinical neurophysiology*, 78(6):447–455, 1991.

[17] Hee-Sok Han Faraz Akram and Tae-Seong Kim. A p300-based word typing brain computer interface system using a smart dictionary and random forest classifier.

[18] Lawrence Ashley Farwell and Emanuel Donchin. Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials. *Electroencephalography and clinical Neurophysiology*, 70(6):510–523, 1988.

[19] Tom Fawcett. An introduction to roc analysis. *Pattern recognition letters*, 27(8):861–874, 2006.

[20] Reza Fazel-Rezai. Human error in p300 speller paradigm for brain-computer interface. In *Engineering in Medicine and Biology Society, 2007. EMBS 2007. 29th Annual International Conference of the IEEE*, pages 2516–2519. IEEE, 2007.

[21] Pierre W Ferrez et al. Error-related eeg potentials generated during simulated brain–computer interaction. *Biomedical Engineering, IEEE Transactions on*, 55(3):923–929, 2008.

[22] Benjamin A Goldstein, Eric C Polley, and Farren Briggs. Random forests for genetic association studies. *Statistical applications in genetics and molecular biology*, 10(1):1–34, 2011.

[23] Jonathan Grizou, Inaki Iturrate, Luis Montesano, Pierre-Yves Oudeyer, and Manuel Lopes. Calibration-free bci based control. In *Twenty-Eighth AAAI Conference on Artificial Intelligence*, pages 1–8, 2014.

[24] C Guger, G Krausz, and G Edlinger. *Brain-computer interface control with dry EEG electrodes*. Proceedings of the 5th Int. BrainComputer Interface Conference, Verlag der Technischen Universitat Graz, Graz, Austria, 2011.

[25] Bin He, Lin Yang, Christopher Wilke, and Han Yuan. Electrophysiological imaging of brain activity and connectivitychallenges and opportunities. *Biomedical Engineering, IEEE Transactions on*, 58(7):1918–1931, 2011.

[26] Anupama H.S, N.K Cauvery, and Lingaraju G.M. Brain computer interface and its type - a study. *international journal advanced engineering technology (IJAET)*, 3(2):739–745, 2012.

[27] Aapo Hyvärinen. Independent component analysis: recent advances. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 371(1984):20110534, 2013.

[28] Aapo Hyvärinen and Erkki Oja. Independent component analysis: algorithms and applications. *Neural networks*, 13(4):411–430, 2000.

[29] Inaki Iturrate, Luis Montesano, and Javier Minguez. Single trial recognition of error-related potentials during observation of robot operation. In *Engineering in Medicine and Biology Society (EMBC), 2010 Annual International Conference of the IEEE*, pages 4181–4184. IEEE, 2010.

[30] Iosr Journals, Subha Austalekshmi.T.V, and Vijaya Sarathi.S. A novel brain computer interface on its types on a future look. 08 2014.

[31] Carrie A Joyce, Irina F Gorodnitsky, and Marta Kutas. Automatic removal of eye movement and blink artifacts from eeg data using blind component separation. *Psychophysiology*, 41(2):313–325, 2004.

[32] Tzyy-Ping Jung, Colin Humphries, Te-Won Lee, Scott Makeig, Martin J McKeown, Vicente Iragui, Terrence J Sejnowski, et al. Extended ica removes artifacts from electroencephalographic recordings. *Advances in neural information processing systems*, pages 894–900, 1998.

[33] Tzyy-Ping Jung, Scott Makeig, Colin Humphries, Te-Won Lee, Martin J Mckeown, Vicente Iragui, and Terrence J Sejnowski. Removing electroencephalographic artifacts by blind source separation. *Psychophysiology*, 37(02):163–178, 2000.

[34] Mandeep Kaur, P Ahmed, and M Qasim Rafiq. Analyzing eeg based neurological phenomenon in bci systems. *International Journal of Computer Applications*, 57(17):40–49, 2012.

[35] Jaya Kulchandani and Kruti J. Dangarwala. Blind source separation using independent component analysis: Algorithms and applications. *International of Computer Science and Information Technologies (IJCSIT)*, 5:6739–3743, 2014.

[36] Vrushali Y Kulkarni and Pradeep K Sinha. Effective learning and classification using random forest algorithm. *International Journal of Engineering and Innovative Technology (IJEIT)*, 3:267–273, 2014.

[37] Rajeev Kumar and Abhaya Indrayan. Receiver operating characteristic (roc) curve for medical researchers. *Indian pediatrics*, 48(4):277–287, 2011.

[38] Ludmila I Kuncheva and Juan J Rodríguez. Interval feature extraction for classification of event-related potentials (erp) in eeg data analysis. *Progress in Artificial Intelligence*, 2(1):65–72, 2013.

[39] Eric C Leuthardt, Gerwin Schalk, Jonathan R Wolpaw, Jeffrey G Ojemann, and Daniel W Moran. A brain–computer interface using electrocorticographic signals in humans. *Journal of neural engineering*, 1(2):63, 2004.

[40] Andy Liaw and Matthew Wiener. Classification and regression by randomforest. *R news*, 2(3):18–22, 2002.

[41] Fabien Lotte, Marco Congedo, Anatole Lécuyer, and Fabrice Lamarche. A review of classification algorithms for eeg-based brain–computer interfaces. *Journal of neural engineering*, 4, 2007.

[42] Dr. T. V. Prasad M. Rajya Lakshm and Dr. V. Chandra Prakash. Survey on eeg signal processing methods. *International Journal of Advanced Research in Computer Science and Software Engineering(IJARCSSE)*, 4:84–91, January 2014.

[43] Sergio Machado, Leonardo Ferreira Almada, and Ramesh Naidu Annavarapu. Progress and prospects in eeg-based brain-computer interface: Clinical applications in neurorehabilitation. *Journal of Rehabilitation Robotics*, 1(1):28–41, 2013.

[44] Kavita Mahajan, M. R. Vargantwar, and Sangita M. Rajput. Classification of eeg using pca, ica and neural network. *IJCA Proceedings on International Conference in Computational Intelligence (ICCIA 2012)*, ICCIA(6):1–4, March 2012.

[45] B Mainsah, K Morton, L Collins, E Sellers, and C Throckmorton. Moving away from error-related potentials to achieve spelling correction in p300 spellers. *IEEE transactions on neural systems and rehabilitation engineering: a publication of the IEEE Engineering in Medicine and Biology Society*, 2014.

[46] Perrin Margaux, Maby Emmanuel, Daligault Sébastien, Bertrand Olivier, and Mattout Jérémie. Objective and subjective evaluation of online error correction during p300-based spelling. *Advances in Human-Computer Interaction*, 2012:4, 2012.

[47] Th Mulder. Motor imagery and action observation: cognitive tools for rehabilitation. *Journal of neural transmission*, 114(10):1265–1278, 2007.

[48] Ganesh R Naik and Wenwu Wang. *Blind Source Separation: Advances in Theory, Algorithms and Applications*. Springer, 2014.

[49] Luis Fernando Nicolas-Alonso and Jaime Gomez-Gil. Brain computer interfaces, a review. *Sensors*, 12(2):1211–1279, 2012.

[50] Eunho Noh and Virginia R de Sa. Discriminative dimensionality reduction for analyzing eeg data. *the annual meeting of the cognitive science society*, 2014.

[51] Jiahui Pan, Yuanqing Li, Zhenghui Gu, and Zhuliang Yu. A comparison study of two p300 speller paradigms for brain–computer interface. *Cognitive neurodynamics*, 7(6):523–529, 2013.

[52] Jaeyoung Park, Kee-Eung Kim, and Yoon-Kyu Song. A pomdp-based optimal control of p300-based brain-computer interfaces. In *AAAI*, 2011.

[53] M Perrin, E Maby, R Bouet, O Bertrand, and J Mattout. *Detecting and interpreting responses to feedback in BCI*. 5th International BCI Conference, 2011.

[54] Matthew B Pontifex, Mark R Scudder, Michael L Brown, Kevin C O'Leary, Chien-Ting Wu, Jason R Themanson, and Charles H Hillman. On the number of trials necessary for stabilization of error-related brain activity across the life span. *Psychophysiology*, 47(4):767–773, 2010.

[55] Chunchu Rambabu and B Rama Murthy. Eeg signal with feature extraction using svm and ica classifiers. *Int J Comput Appl*, 85(3):1–7, 2014.

[56] Bertrand Rivet, Antoine Souloumiac, Virginie Attina, and Guillaume Gibert. xdawn algorithm to enhance evoked potentials: application to brain–computer interface. *Biomedical Engineering, IEEE Transactions on*, 56(8):2035–2043, 2009.

[57] Bertrand Rivet, Antoine Souloumiac, Guillaume Gibert, and Virginie Attina. p300 speller brain-computer interface: Enhancement of p300 evoked potential by spatial filters. In *Signal Processing Conference, 2008 16th European*, pages 1–5. IEEE, 2008.

[58] Avid Roman-Gonzalez. Eeg signal processing for bci applications. In *Human–Computer Systems Interaction: Backgrounds and Applications 2*, pages 571–591. Springer, 2012.

[59] Ricardo Ron-Angevin, Sergio Varona-Moya, Leandro da Silva-Sauer, and Trinidad Carrión-Robles. A brain–computer interface speller with a reduced matrix: A case study in a patient with amyotrophic lateral sclerosis. 2014.

[60] Sandra Rousseau, Christian Jutten, and Marco Congedo. The error-related potential and bcis. In *20th European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning (ESANN 2012)*.

[61] Sandra Rousseau, Christian Jutten, and Marco Congedo. Time window selection for improving error-related potential detection. In *5th International Conference on Neural Computation Theory and Applications (NCTA 2012)*, page 1, 2012.

[62] Nico M Schmidt, Benjamin Blankertz, and Matthias S Treder. Online detection of error-related potentials boosts the performance of mental typewriters. *BMC neuroscience*, 13(1):19, 2012.

[63] Abeer E Selim, Manal Abdel Wahed, and Yasser M Kadah. Machine learning methodologies in p300 speller brain-computer interface systems. In *Radio Science Conference, 2009. NRSC 2009. National*, pages 1–9. IEEE, 2009.

[64] Hao Shen, Knut Hüper, and Martin Kleinsteuber. On fastica algorithms and some generalisations. In *Numerical Linear Algebra in Signals, Systems and Control*, pages 403–432. Springer, 2011.

[65] Jonathon Shlens. A tutorial on principal component analysis. *arXiv preprint arXiv:1404.1100*, 2014.

[66] Uzma Siddiqui and AN Shaikh. An overview of electrooculography. *International Journal of Advanced Research in Computer and Communication Engineering*, 2(11):4328–4330, November 2013.

[67] Yaduvir Singh, Sweta Tripathi, and Manoj Pandey. Analysis of digital iir filter with labview. *Analysis*, 10(6), 2010.

[68] Le Song and Julien Epps. Classifying eeg for brain-computer interface: Learning optimal filters for dynamical system features. *Computational intelligence and neuroscience*, 2007, 2008.

[69] Martin Spüler and Christian Niethammer. Error-related potentials during continuous feedback: using eeg to detect errors of different type and severity. *Frontiers in human neuroscience*, 9, 2015.

[70] Dan Szafir and Robert Signorile. An exploration of the utilization of electroencephalography and neural nets to control robots. In *Human-Computer Interaction–INTERACT 2011*, pages 186–194. Springer, 2011.

[71] J Teeuw. Comparison of error-related eeg potentials. In *13th Twente Student Conference on IT*, 2010.

[72] Michal Teplan. Fundamentals of eeg measurement. *Measurement science review*, 2(2):1–11, 2002.

[73] G Townsend, BK LaPallo, CB Boulay, DJ Krusienski, GE Frye, CKea Hauser, NE Schwartz, TM Vaughan, JR Wolpaw, and EW Sellers. A novel p300-based brain–computer interface stimulus presentation paradigm: moving beyond rows and columns. *Clinical Neurophysiology*, 121(7):1109–1120, 2010.

[74] Tsvetomira Tsoneva, Jordi Bieger, and Gary Garcia-Molina. Towards error-free interaction. In *Engineering in Medicine and Biology Society (EMBC), 2010 Annual International Conference of the IEEE*, pages 5799–5802. IEEE, 2010.

[75] Andreas Widmann, Erich Schröger, and Burkhard Maess. Digital filter design for electrophysiological data–a practical approach. *Journal of neuroscience methods*, 2014.

[76] Beste F Yuksel, Michael Donnerer, James Tompkin, and Anthony Steed. A novel brain-computer interface using a multi-touch surface. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 855–858. ACM, 2010.

[77] Weidong Zhou and Jean Gotman. Automatic removal of eye movement artifacts from the eeg using ica and the dipole model. *Progress in Natural Science*, 19(9):1165–1170, 2009.