Aspire Create
Centre for Rehabilitation Engineering and Assistive Technology

Coursework Submission Coversheet

Student Number: 18149489
Module Code: SURG0093
Report Title: Classification of BCI System Using Features in Eyes Closed-Eyes Open State of Healthy Subject
Submission Deadline [DD/MM/YYYY]: 26/02/2020
Word count (include all sections except references/bibliography and appendices. Exclude footnotes, tables, figures, captions, references/bibliography and appendices): 1636

NB: I confirm that by submitting this piece of coursework, I have read and understood the UCL statements and guidelines concerning plagiarism.

See: https://www.ucl.ac.uk/students/exams-and-assessments/plagiarism

By submitting this piece of coursework, I declare that:

1. This submission is entirely my own unaided work.
2. Wherever published, unpublished, printed, electronic or other information sources have been used as a contribution or component of this work, these are explicitly, clearly and individually acknowledged by appropriate use of quotation marks, citations, references and statements in the text.

Classification of BCI System Using Features in Eyes Closed-Eyes Open State of Healthy Subject

Abstract

A brain-computer interface (BCI) system was created using brain signals derived when the subject closed and opened their eyes. The largest difference in amplitude between eyes closed and eyes open occurred in the frequency range of 7-11 Hz, hence it was used as the feature. Three classifiers were used: linear discriminant analysis (LDA), quadratic discriminant analysis (QDA) and Naïve Bayes. The performance of each classifier was measured using misclassification rate and area under the curve of receiver operating characteristics (AUC). The results concluded that LDA had the best performance in both training data and test data, compared to the other classifiers.

I. Introduction

For people with severe cervical spinal cord injury, there are few available assistive technologies that could help them to interact with the environment. Devices such as sip-and-puff, eye tracking and voice recognition require muscle activation to some degree. The idea of using brain signal as a control modality has become an interesting topic for two decades, giving rise to brain-computer interface (BCI) [1].

There are several types of brain activity that can be used as an input of BCI, namely P300 event-related potential (ERP), steady state visually evoked potentials (SSVEP) and motor imagery. The first two patterns are generated in response to external stimulus, so they need additional setup to initiate the stimulus. In contrast, motor imagery relies on the user to start generating patterns. It utilises thoughts of doing motor functions without making the actions. Another useful spontaneous activity is closing-opening the eyes. A difference between eyes closed and eyes open was found at frequencies of 8 to 12 Hz [2]. This range of frequency is identified as alpha band. Although relatively simple, there was a limited number of study using this feature [3] [4]. However, these studies have a specific aim to identify certain body conditions. The classification algorithms, artificial neural network and Spearman correlation analysis are not particularly useful outside of the applications in the studies. Therefore, this study addresses the application of eyes closed and eyes open technique as a BCI system.

Brain signal of one subject was acquired using electroencephalography (EEG) due to its invasive nature and real-time data collection. The model was applied to three different kinds of classifier: linear discriminant analysis (LDA), quadratic discriminant analysis (QDA) and Naïve Bayes. The objective is to evaluate the performance of each classifier in a BCI system derived from eyes closed and eyes open activities.

II. Methods

The EEG device comprised an EEG cap, driver box and amplifier. The cap consisted of 16 electrodes on the scalp, ground electrode on the front of Fz and reference electrode on the ear. The electrodes were labelled as Fz, FC3, FC1, FCz, FC2, FC4, C3, C1, Cz, C2, C4, CP3, CP1, CPz, CP2, CP4 (Figure 1). These electrodes were 20% apart from each other. The driver box connected the electrodes with the amplifier. The amplifier included analog-to-digital converter (ADC).

EEG cap was placed on the subject sitting on a chair. To adjust the cap, Cz was used as the midpoint of the head. The length from nasion to Cz was equal to the length between Cz and inion. Similarly, the distance between left ear and Cz was equal to the distance between Cz and right ear. All electrodes were filled with conductive gel. The experiment involved the subject closing their eyes for 30 s followed by opening their eyes for 30 s alternately for a given time period. Data were recorded at 512 Hz sampling rate.

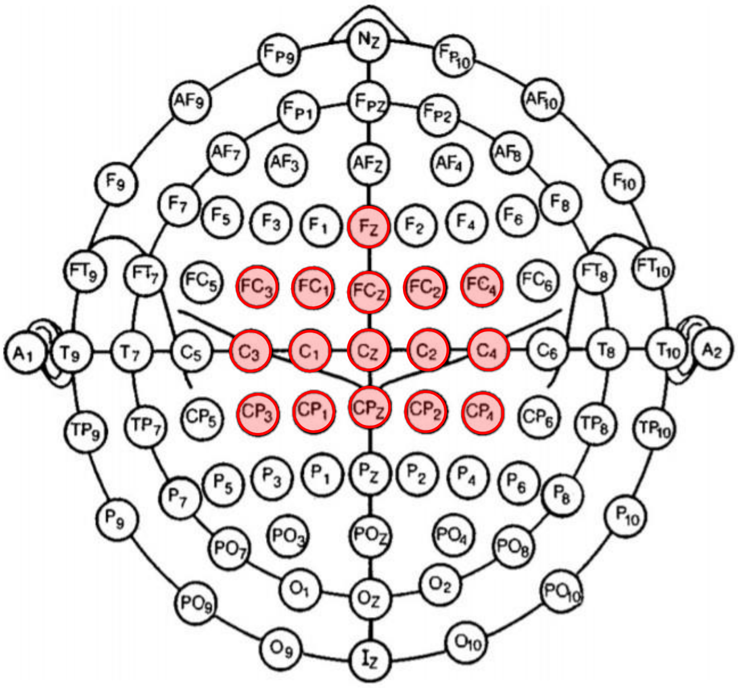


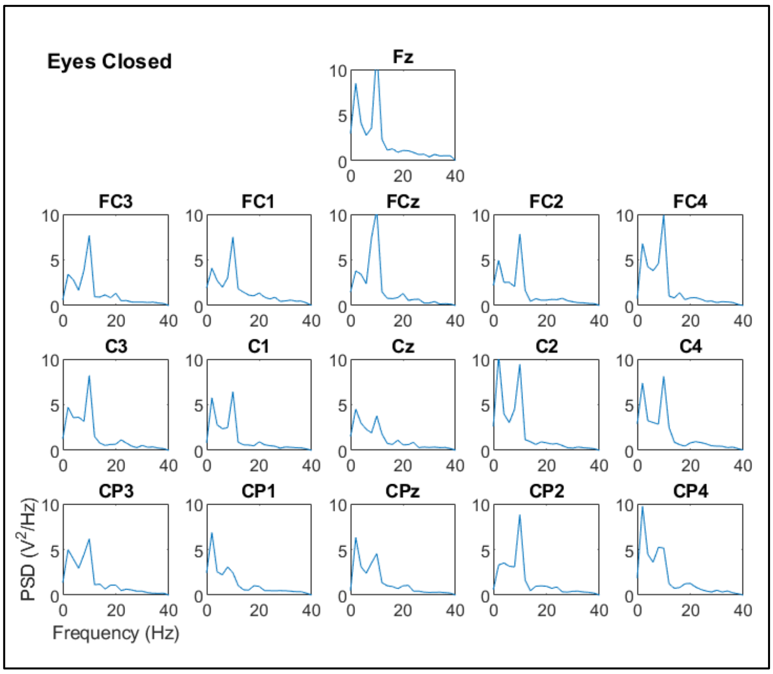
Figure 1. International 10/20 system, a standardised method that determines the placement of scalp electrodes in EEG measurement.

The processing of the signals was done in MATLAB. Epochs of 24 s were extracted from the raw signals produced by the EEG device. The signals, which was in time domain, were converted to frequency domain using fast Fourier transform (FFT) algorithm before band-pass filtered with cut-off frequencies of 2 and 40 Hz using Finite Impulse Response (FIR) filter. The potential was transformed into power spectral density (PSD).

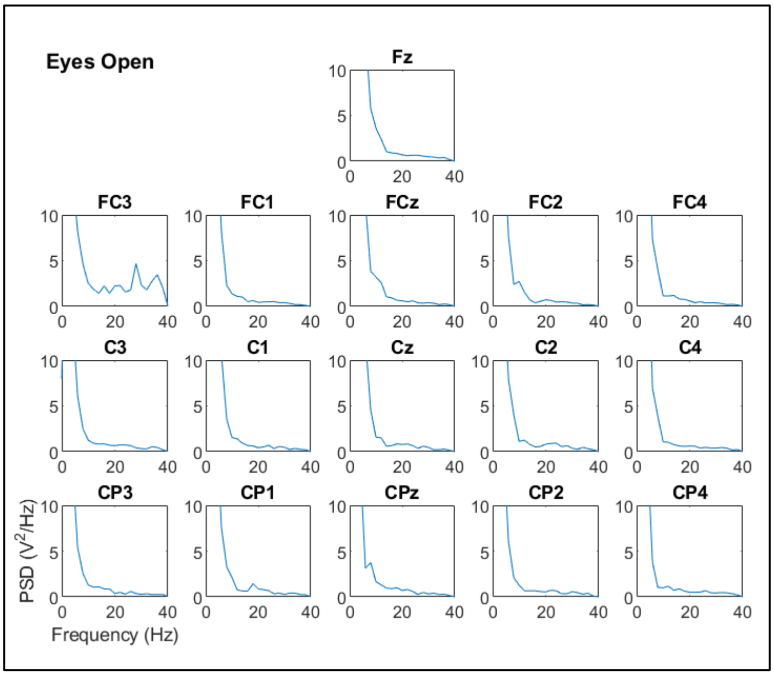
Features were extracted from a particular electrode based on the highest discriminant power between eyes closed and eyes open. LDA, QDA and Naïve Bayes were applied to the training data set. Two metrics were computed to evaluate the performance of the resulting model: misclassification rate and area under the curve (AUC) of receiver operating characteristics (ROC).

III. Results

The signals of both eyes closed and eyes open conditions from all electrodes are presented in frequency domain (Figure 2). In eyes closed condition, two peaks at low frequencies of 1-3 Hz and 8-11 Hz were found. In eyes open, a common feature was high power at lower frequency spectrum. Features were extracted from electrode CPz based on the highest discriminant power between eyes closed and eyes open in the frequency band of 7-11 Hz (Figure 3). Classification using LDA resulted in 9 misclassified data, represented as x in Figure 4. This corresponded to a resubstitution error of 25.00% (Table 1). QDA resulted in 11 misclassified data, which was 30.56%. Naïve Bayes produced 13 misclassified data, which was 36.11%.



(a)



(b)

Figure 2. EEG signals for eyes (a) closed and (b) open in frequency domain from 16 different electrodes.

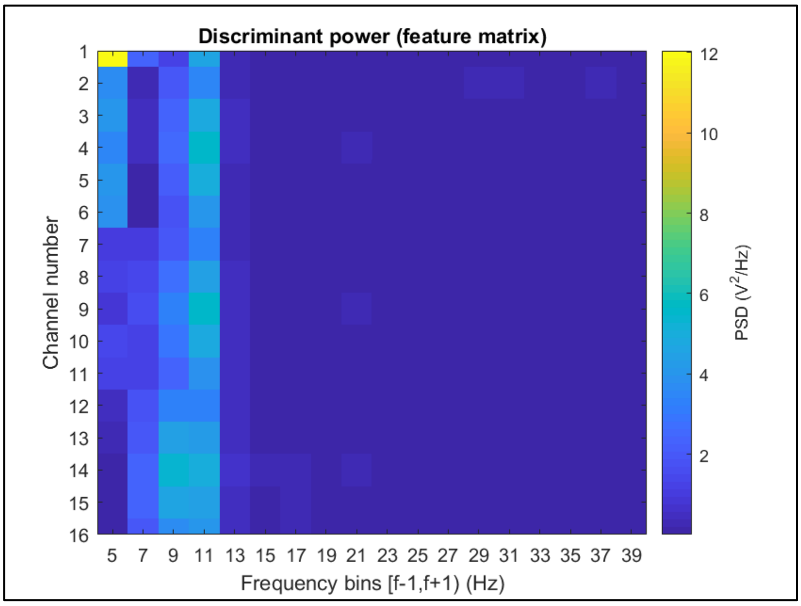
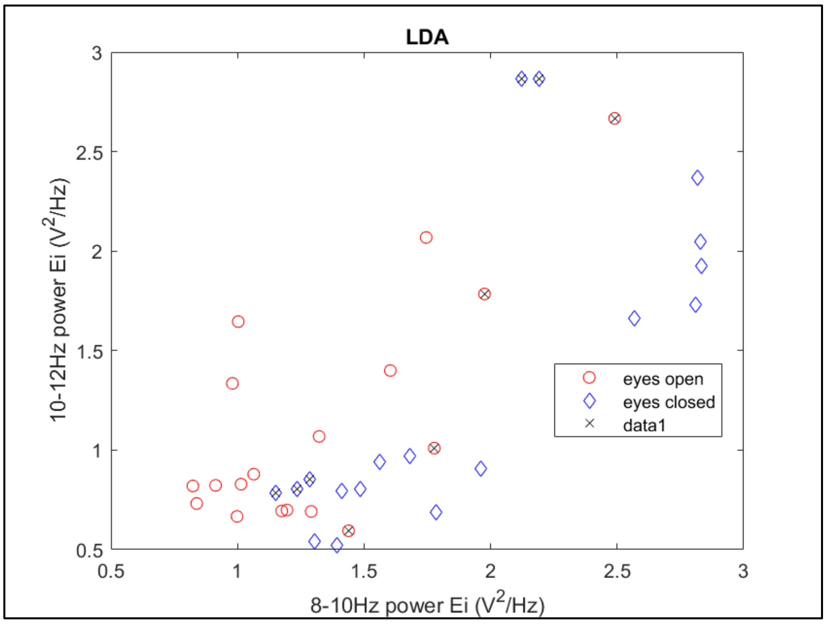
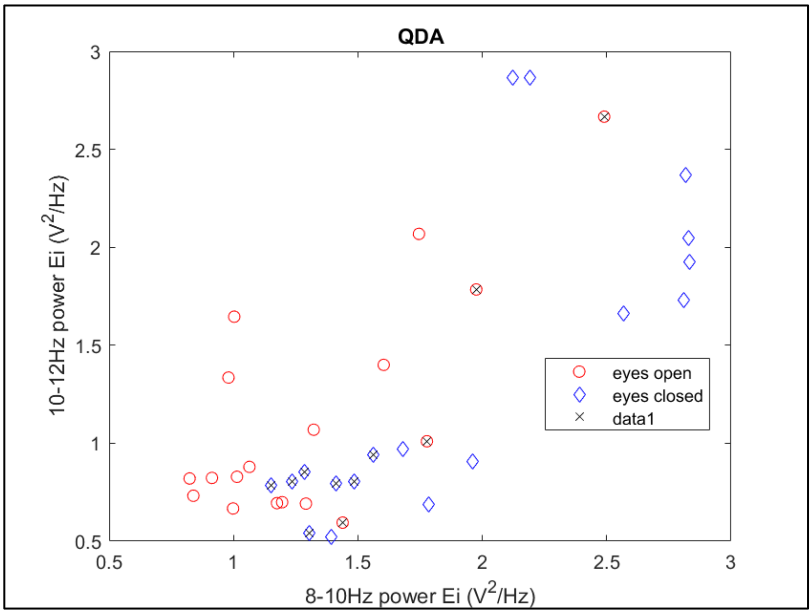


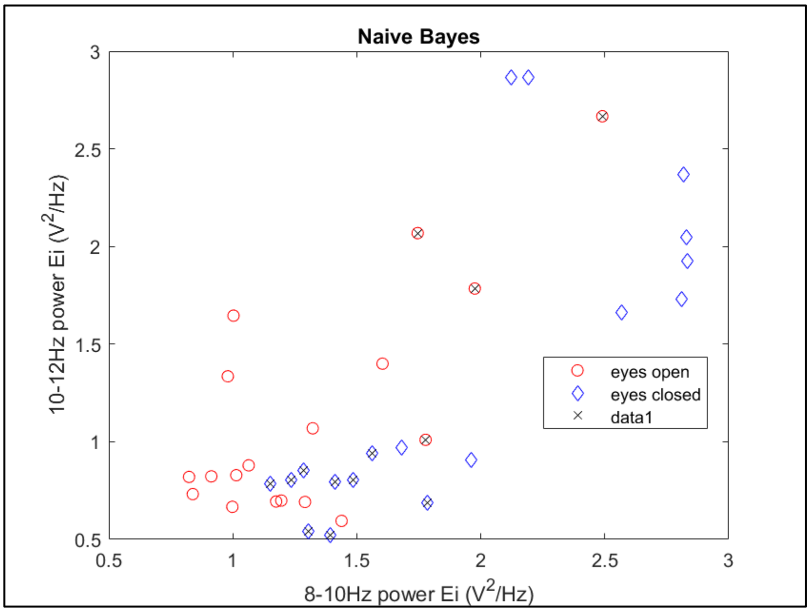
Figure 3. Discriminant power matrix illustrated the difference of amplitude between eyes closed and eyes open across 16 electrodes.


LDA
10-12Hz power Ei (V²/Hz)
8-10Hz power Ei (V²/Hz)
eyes open
eyes closed
data1
0.5
1
1.5
2
2.5
3

(a)


QDA
10-12Hz power Ei (V²/Hz)
8-10Hz power Ei (V²/Hz)
eyes open
eyes closed
data1
0.5
1
1.5
2
2.5
3

(b)



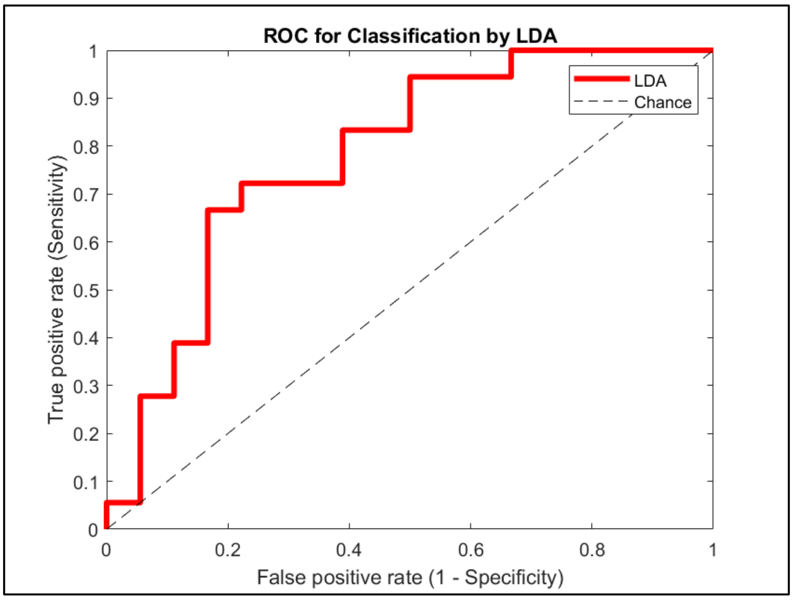
(c)

Figure 4. The application of (a) LDA, (b) QDA and (c) Naïve Bayes classifier to training data.

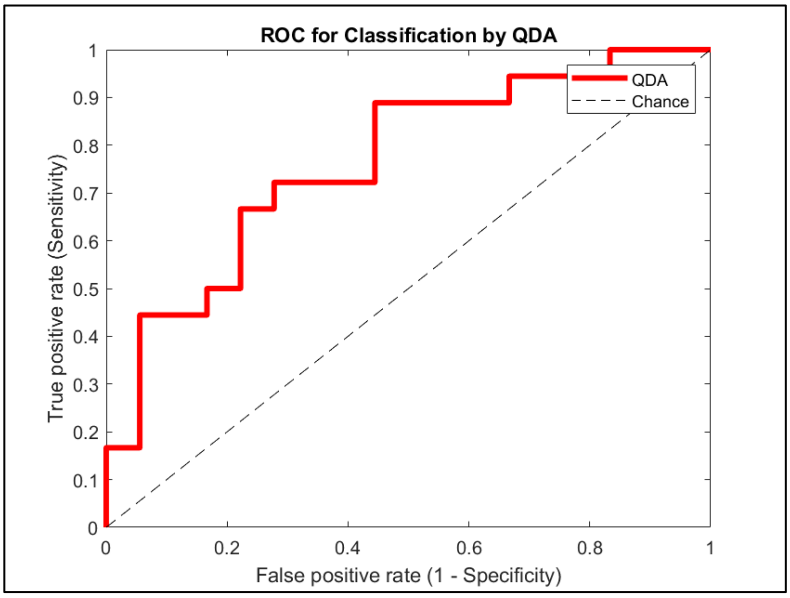
Table 1. Performance measurement of the model using different classifier.

Parameter	LDA	QDA	Naïve Bayes
Resubstitution error (%)	25.00	30.56	36.11
Cross-validation error (%)	30.56	36.11	44.44
AUC (%)	78.40	76.54	58.80

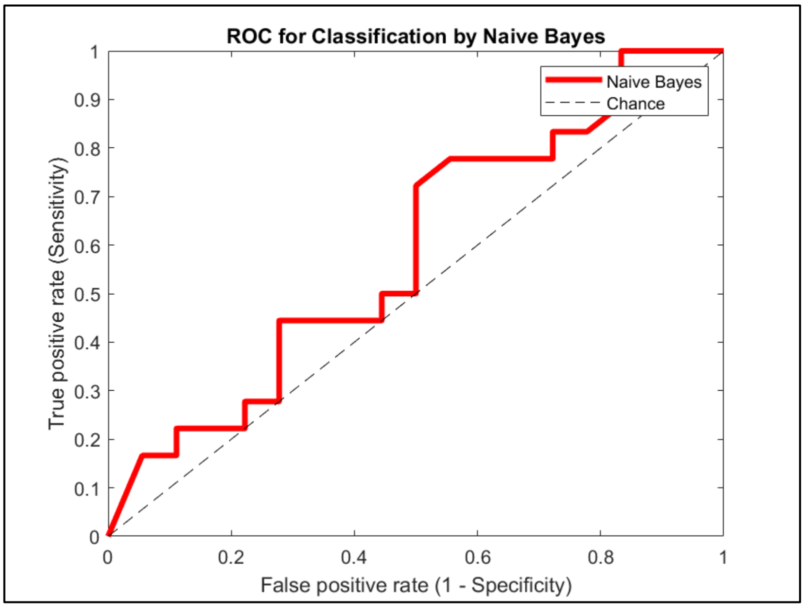
The performance of each classifier was measured with test data, generating cross-validation error. The cross-validation error of all classifiers were higher than the resubstitution error. Figure 5 showed ROC curve, which depicted the probability of classification at various settings. AUC defined the accuracy of the classifier, or the degree of separation. The graph inferred that the chance of LDA, QDA and Naïve Bayer classifiers to correctly predict the user input was 78.40%, 76.54% and 58.80% respectively (Table 1).


ROC for Classification by LDA
LDA
Chance
True positive rate (Sensitivity)
False positive rate (1 - Specificity)
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
0
0.2
0.4
0.6
0.8
1

(a)


ROC for Classification by QDA
QDA
Chance
True positive rate (Sensitivity)
False positive rate (1 - Specificity)
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
0
0.2
0.4
0.6
0.8
1

(b)



(c)

Figure 5. ROC curve for (a) LDA, (b) QDA and (c) Naïve Bayes classifier was created by plotting true positive rate against false positive rate at different scenarios.

IV. Discussion

During the experiment using manual trigger, movement artefacts might happen due to the subject blinking. Therefore, 2 seconds of data before and after each trigger were discarded. Band pass filters were used to attenuate noises and powerline interferences. When calculating PSD for every epoch and every electrode, Welch method was used instead of squared value of FFT. Welch method divides the squared values into several sections and averages them over time, therefore reduces the noise [5].

Electrode number 14 was chosen because it showed the highest distinction between eyes closed and eyes open in the alpha band frequencies, 7-11 Hz. This corresponded to studies arguing that the highest alpha waves originated from occipital region of the cerebral cortex [6]. The power amplitude in the frequency of 7-9 Hz and 9-11 Hz was set as two features. These features were used to construct training data, which was then applied to a classifier. A classifier assigns data into classes of eyes closed and eyes open. LDA generates linear axes that maximise the degree of separation between classes. In contrast, QDA generates quadratic boundaries. Naïve Bayes assumes that all training data values are independent [7]. In order to linearise nonlinear problems, the data should be projected into a new feature space [8]. Hence, new feature space consisted of 8-10 Hz and 10-12 Hz were constructed.

Training data applied to a classifier results in a model. The accuracy of the model was evaluated with resubstitution error, which is the probability of misclassification of the model [9]. Misclassification occurred if the data did not fit within the boundaries of 8-10 Hz and 10-12 Hz. Table 1 illustrated the resubstitution error of LDA was the lowest and Naïve Bayes the highest. This indicated that LDA performed better than the other two in the training data.

When an input of EEG signal (new data) is applied on the model, it recognises the feature vector and translates it to an output of commands. In BCI systems, both user and device need to adapt, because brain features change over time and it would repeatedly produce a new model that provides a prediction of user intentions (labels), forming a loop of machine learning (Figure 6) [8].

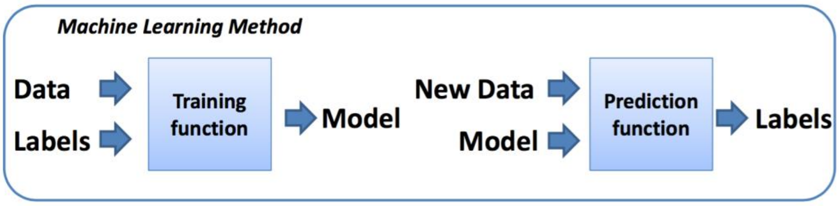


Figure 6. Machine learning loop.

As depicted in Figure 6, BCI system receives new data, therefore new metrics should be used to describe the performance of the model on future (test) data. Since there was no new data in the experiment, the test data used the same data as training data. This can be done by splitting the data into k sets of equal size. Each set is used for testing the model and the remainder is used for training. The method is called k-fold cross validation [10]. The results were averaged across the k sets, yielding a new error, cross-validation error. All results agreed that the performance decreased when the model was generalised in test data.

A better way to illustrate the performance of the model is by depicting ROC curve. The degree of separation was measured by the area under the curve. Bigger AUC means there was bigger chance that the device can separate between eyes closed and eyes open. The degree of separation was influenced by the number of samples. Doubling the samples increase the AUC by 3.88%. It also reduced the resubstitution error by 10%. Besides taking more data during the experiment, the result could also be manipulated by reducing epoch length.

The difference between resubstitution error and cross-validation error could represent the generalisation of the model within subject. The results indicated that LDA showed the most stable performance over time. For determining the generalisation of the model between subject, the study requires new data from other subjects.

V. Conclusion

Classification of a BCI system which used amplitude in alpha wave frequencies as the feature provided higher accuracy if used with LDA than QDA and Naïve Bayes. The misclassification error found in using LDA can be reduced by increasing the training data set. The limitation of this study is the use of same data for both training and test data. Using data from several subjects over a significant time period could be beneficial to create a more reliable system. By classifying the two states of eyes, the paradigm can be used to generate two commands associated with any device of the user's preference.

References

[1] T. K. C. Z. S. H. F. L. F. C. F. A. A. R. C. H. D. M. A. K. Sonja C. Kleih, "Out of the frying pan into the fire—the P300-basedBCI faces real-world challenges," *Progress in Brain Research,* vol. 194, pp. 27-46, 2011.

[2] N. M. G. Bernard J. Baars, "The brain is conscious," in *Fundamentals of Cognitive Neuroscience*, Academic Press, 2013, pp. 211-252.

[3] A. B. V. B. A. L. R. T. G. C. F. S. A. F. R. N. F. G. L. B. M. T. R. C. V. G. A. S. P. A. S. S. F. B. G. S. G. L. G. Triggiani, "Classification of Healthy Subjects and Alzheimer's Disease Patients with Dementia from Cortical Sources of Resting State EEG Rhythms: A Study Using Artificial Neural Networks," *Frontiers in Neuroscience,* vol. 604, no. 10, 2017.

[4] J.-A. R. R. M. C. A. L. C. L. a. J. V.-D. Micoulaud-Franchi, "Parieto-temporal Alpha EEG Band Power at Baseline as a Predictor of Antidepressant Treatment Response with Repetitive Transcranial Magnetic Stimulation: A Preliminary Study," *Journal of Affective Disorders,* vol. 137, pp. 156-160, 2012.

[5] J. O. Smith, Spectral Audio Signal Processing, W3K Publishing, 2011.

[6] S. Ghoreishizadeh, "Biopotential signals," UCL, London, 2020.

[7] M. C. A. L. F. L. B. A. F Lotte, "A review of classification algorithms for for EEG-based brain-computer interfaces," *Journal of Neural Engineering,* vol. 4, p. 24, 2007.

[8] E. W. W. Jonathan R. Wolpaw, Brain-Computer Interfaces. Principles and Practice., Oxford University Press, 2012.

[9] R. Rakotomalala, "Resampling methods for error rate estimation in supervised learning," July 2009. [Online]. Available: http://eric.univ-lyon2.fr/.

[10] A. Nordman, "Data Mining," 2011. [Online]. Available: http://staffwww.itn.liu.se/~aidvi/courses/06/dm/.