ABSTRACT

Objectives

This study explored the architecture of a machine learning application in a Brain Computer Interface (BCI) system and experimented various classification algorithms in a P300 Speller BCI.

Methods

This study used the P300 Speller BCI datasets from the BCI Competition. The datasets were pre-processed and segmented for analysing and fitting purpose. Wyrm Python toolbox and Scikit-learn Python package were used to help with the exploration. Various classifiers were also experimented: Linear Discriminant Analysis (LDA), Gradient Boosting, Ada Boost, K-neighbours, Decision Tree, Random Forest, Quadratic Discriminant Analysis (QDA), Linear Support Vector Classifier (SVC) and SVC. The predicted characters were compared with true labels downloaded from BCI Competition test results to calculate the percentage of correctly predicted characters.

Results

Main components of a machine learning architecture in a BCI system were explored: EEG signals pre-processing: artifacts removal, features selection, features extraction and predictive model training and results analysing.

For the P300 Speller datasets explored in this study, LDA overall produced better results, with the best score at 94.5% while other classifiers were in a much lower range: <30%, except for the case of Linear SVC with 56% for subject A and 90% for subject B. All those results were observed without Principal component analysis (PCA) and used 64 channels. However, the accuracy dropped when testing with 8 channels, single channel and when testing with fewer trials.

Conclusions

Even though this study only explored machine learning architecture in Speller BCI systems, the knowledge and methodologies can be used to explore P300 BCIs in general. Some classification algorithms were tested, and among those, LDA from Wyrm Python toolbox produced the highest prediction score. However, future work is needed to improve accuracy when using fewer channels or fewer trials.