Physiological study on online education: Feedback classification using EEG signals

Adithya Ajith, Muzzamil Bashir
ISEN 630, Texas A&M, College Station, TX

Physiological feedback acquired by students while watching online educational videos was analyzed to assess the state of confusion. EEG data from a single electrode headset device was taken as input to a classifier with binary response. A comparative evaluation of different popular classifiers was reported. The chosen technique showed a remarkable improvement in cross validated classification rate from previous study on the same dataset. The classifier performed better when considering only theta and alpha bands in the EEG as predictors. The study demonstrates the effectiveness and potential of small scale EEG devices for assessing mental state during learning processes.

INTRODUCTION

Massive online open courses (MOOC) have increased access of education to people across the globe. These courses provide cost effective opportunities to enhance knowledge and develop in-demand skills. Since its advent, MOOCs have grown in popularity, as it provides alternate educational offering to both students and working professionals. These courses cover some of the emerging fields which adult population might need to cross train in. There has been however a significantly higher dropout rate associated with MOOCs in comparison to conventional education. Park et al. attributed this dropout to social factors like family support and organization structure (Park, & Choi, 2009).

Since the target audience of an online system come from diverse backgrounds, there is a need to customize these videos to reduce cognition anxiety among the participants. The lack of real time feedback in an online system doesn't give the instructor an opportunity to improvise and improve the delivered content. Yang et al. reported confusion as a potential reason for dropouts among students (Yang, Wen, Howley, Kraut, Rose, 2015). Dropout rate of 86% was reported in one of the studies conducted at a Mexican private university (Zermeno, & Garza, 2016). Better understanding of cognition loads associated with online learning systems can help in development of more responsive online platforms.

Confusion can be measured using either a subjective or a physiological feedback. Subjective approach relies on self-reporting by the participants on a scale of confusion. Zermeno reported in the study that only 28.6% of the registered participants gave feedback surveys (Zermeno, & Garza, 2016). The participants who dropout are less likely to take part in the subjective feedback survey. Physiological parameters like cardiac, brain and eye activities can also be used to assess psychological states. It has been reported that the physiological aspect of online learning has been comparatively less studied (Nikolaeva, & Roberts, 2016). This study tries to correlate the physiological response of the brain to the learning process.

Electroencephalography (EEG) is an electrophysiological technology that maps the electrical activity in the brain. EEG has proven to provide insightful information that correlates obtained signals to student engagement and anxiety in learning activities (Scotti, Cerutti, Mainardi, Villamira, 2005). Classical EEG method involves data collection under strict laboratory conditions using multi-channel electrode net (Milnik, 2009). These sensors require gel or saline solution to work and are inconvenient for the user. In the past decade, Brain-computer interface (BCI) has rapidly evolved and commercially available wearable EEG headsets are available at cheaper rates. They provide a comfortable way to measure EEG outside lab settings. These devices requires no gel and are user friendly. In past BCI use has been limited to patients for helping them interact with external environment. A growing interest of these devices for motor imagery applications has been reported (Leon, Cano, Ibarrola, 2016). These devices have a potential to be an everyday device of the future. The data recorded through these devices can help us in better understanding of cognition processes associated with attending online courses.

Poltavski (2015) investigated the usefulness of a low cost single electrode wearable wireless EEG headset by Nerosky Inc. The device "Mindset" comes with powerful noise filtering capabilities. This device has been validated and demonstrates comparable performance to Biopac system, a well known wet electrode EEG system, widely used in medical field. This device has shown reasonable comfort for the users for research and experimentation purposes (Ekandem, Davis, Alvarez, James, Gilbert, 2012).

This paper involves the use of "Mindset" generated EEG signal for interpretation of state of confusion among students watching MOOC videos. Studies have shown that EEG can be used to determine the student's state of confusion - information perception and absorption through online learning. EEG output is a combination of various frequency bands. The power output in theta and alpha frequency bands is related to cognitive and memory performance of an individual (Klimesch, 1999). However these relationships are comparatively complex and harder to quantify. Supervised machine learning has become an effective tool in correlating complex patterns within data to a response. The data is usually divided into subsets of training and test data. Data modelling is then carried out using the training data, a relation among the predictors and responses is developed as a result. The model is then validated with the test data to estimate how well the model classify new data.

This study follows an earlier attempt to classify confusion among students watching online videos. Wang et al. reported two classification accuracies which will be described in later section. Accuracies of 51% and 57% were achieved for student independent and student specific classifiers respectively (Wang, Li, Hu, Yang, Meng, Chang, 2011). The study establishes a weak relationship between state of confusion and brain wave signals. These low values of accuracy limits the potential of the BCI device for future studies.

We address the question - how to improve the accuracy of detection of confusion based on the EEG signal data. The data needs to be transformed using available transformation techniques. Secondly, we explore the various machine learning methods for this classification problem.

With an improved classification model, in future, the providers of the online courses can use EEG headsets to generate a subjective rating of course materials to provide a simulation of real world classroom responses, such as when a teacher is given feedback from an entire class. Then teachers can improve video clips based on these impressions. Also with sufficient study, we can we develop algorithms that can be added as a built in feature in the EEG headset, which will output the level of confusion of the user.

METHOD

Dataset

Due to the time constraint, instead of conducting experiments ourselves, we will be using the EEG dataset of a study conducted on a group of 10 college students with the NeuroSky MindSet device (Wang, Li, Hu, Yang, Meng, Chang, 2011). The MindSet makes use of a forehead sensor placed at Fp1 position as defined by International 10-20 system and a reference contact point A1 on the left ear (Milnik, 2009), as shown in Figure 1. NeuroSky's API was used to collect the EEG data.

Each student was shown 10 videos of 2 minutes duration each. Keeping in mind the educational background of the students, the videos were divided into two categories such that half of them would create confusion for the participants. These videos were then randomly shown to all of the students. The students were asked to relax for 30 seconds, after which they were shown the video. They were instructed to learn as much as they can from the video that was about to be shown. At the end of the video, students were asked to rate his/ her perceived confusion level on a scale of 1-7. Assessing the median among the responses, the scale was mapped into binary response, with 0 as 'no confusion' and 1 as 'confusion state'. A total of 100 possible scenarios are recorded and corresponding confusion level was marked by the respective user.

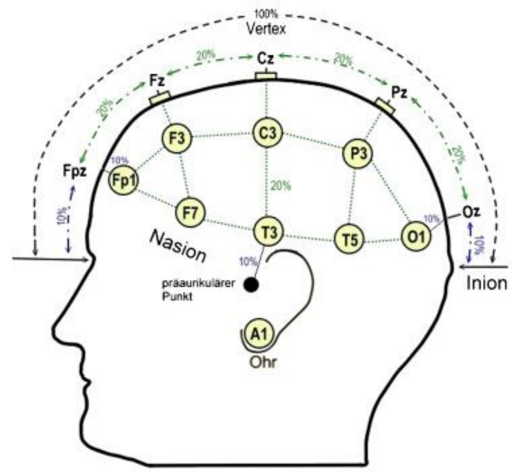

*Figure 1. E*lectrode placement *locations (International 10-20 system)*

The sampling frequency for the raw data is 512 Hz and the associated software transforms this raw data by applying Fourier transformation to split it into its constituent sine waves. Mindset applies an amplitude filter to further decompose the constituent waves into lower and higher spectrum. This transformation leads to a total of 9 features which are listed in Table 1. The software comes with an in-built noise filter algorithm. The efficacy of the device and the software is reasonably suited for experimental and research studies (Ekandem, Davis, Alvarez, James, Gilbert, 2012).

Features	Description
Raw	Original EEG signal
Delta	Delta 1-3 Hz of power spectrum
Theta	Theta 4-7 Hz of power spectrum
Alpha1	Alpha1 Lower 8-11 Hz of power spectrum
Alpha	Alpha 2 Higher 8-11 Hz of power spectrum
Beta1	Beta1 Lower 12-29 Hz of power spectrum
Beta	Beta 2 Higher 12-29 Hz of power spectrum
Gamma1	Gamma1 Lower 30-100 Hz of power spectrum
Gamma2	Gamma2 Higher 30-100 Hz of power spectrum

Table 1. Features of EEG data

Data Processing

The given data points were checked for anomalies, ranges and central tendency. It was noticed that subject 6 had much higher extreme values for raw data as shown in the Figure 2. It is suspected that the data acquisition system was not calibrated during the collection of data for this subject. We have excluded the stated subject from further analysis.

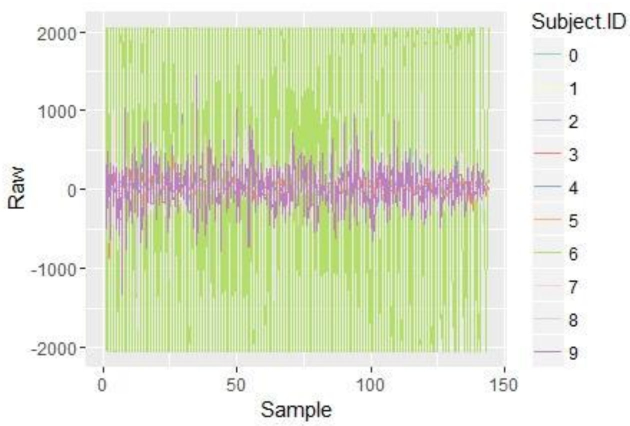

Figure 2. Raw EEG values for all subjects

In the previous study, mean values of all features for each video was considered as the predictor to the developed classification model. It is known that range of EEG power varies across different individuals. Inputting absolute value of features without normalizing inherits error into the model.

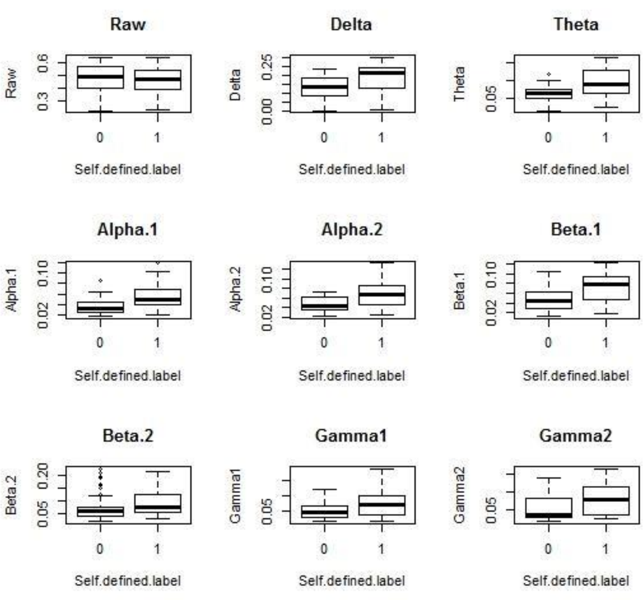

Figure 3. Boxplot of normalized values for all features

In order to avoid this, we devised a normalized scale for each of the subject under study. Minimum and maximum values of each characteristic were determined across the entire spectrum of the study. These minimum values were then subtracted from the individual values and then divided by the range of entire set to give an overall normalized dataset for all the subjects involved. Box plot of normalized values over the entire data points is shown in Figure 3. This approach improves on the previous study as it takes into account the EEG power variations among various individuals. After normalizing, we found the mean values of all the features for each of the 90 student video combinations. These mean values were used as predictors to develop the classification model. Subject-wise boxplot of individual predictors show a distinct division among the two states, as shown in Figure 4.

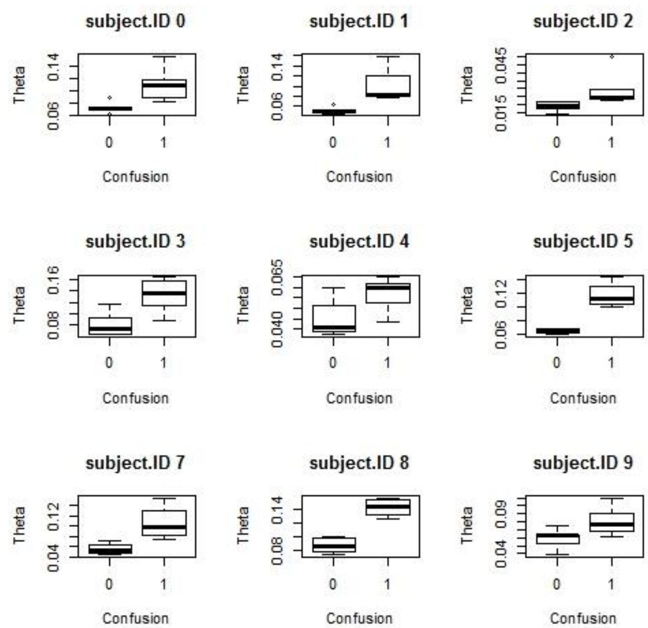

Figure 4. Boxplot of normalized theta values for all subject IDs

Data Analysis

It is attempted to make use of popular classification techniques to improve the cross validated results of the model (James, Witten, Hastie, Tibshirani, 2013). A comparative analysis of the data using different techniques was performed. The data analysis methods applied for the given study includes Logistic Regression, Linear Discriminant Analysis LDA, Quadratic Discriminant Analysis QDA, K Nearest Neighbor, Support vector machine and Naïve Bayes.

We used leave one student out CV to train the model on data of all but one subject ID. The trained model is then tested on this excluded subject, and the comparison between predicted and actual state give us the classification rate. This process is then repeated for all possible sets of training and test data, noting the classification rate of each. A mean value of the classification rate is then reported as the overall accuracy of the model.

In the previous study, two classifiers were developed namely student specific and student independent. Student specific classifier is built on a specific subject ID, which is then used to estimate the response for that individual on test conditions. Student independent classifier is more generalized and uses the entire dataset across all subject IDs for model development and cross validation. For the purpose of this study, accuracies reported are for student independent classifier. Moreover the previous study had included all features as predictor variables. Different combinations of feature can be used to improve the accuracy of the classifier. Based on the findings from literature review, we tested for each classification method with all the features as well as with just Theta, Alpha.1 and Alpha.2 as predictors.

Since some of the applied data analytics techniques assumes Gaussian distribution for the predictors. Normality of the data set was checked using Q-Q plots. The deviation from the Gaussian distribution in case of Alpha.1 predictor was corrected through log transformation as shown in Figure 5.

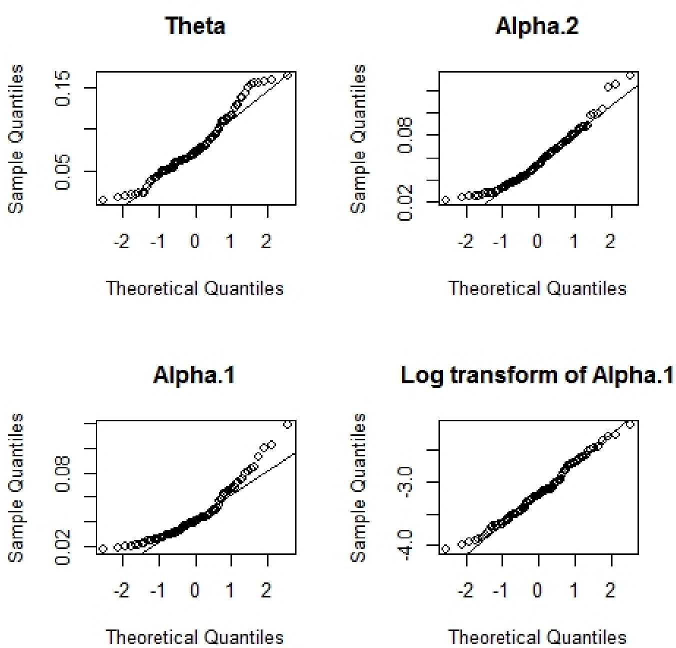

Figure 5. Q-Q plot of the normalized data

RESULTS

Except for SVM method, all classification models showed improvement in classification accuracy with three predictors (Theta, Alpha.1 and Alpha.2). A comparative evaluation of the classification accuracies using the previously stated classification models is given in Figure 6. It was noticed that Naïve Bayes has the best accuracy of 78%. This superior performance can be attributed to the relatively smaller dataset for this study. The function naïve Bayes was used from the library e1071 in R software. It uses the Bayes rule to compute a-posterior probabilities of each associated categories. It assumes a normal distribution and independence among all predictor variables.

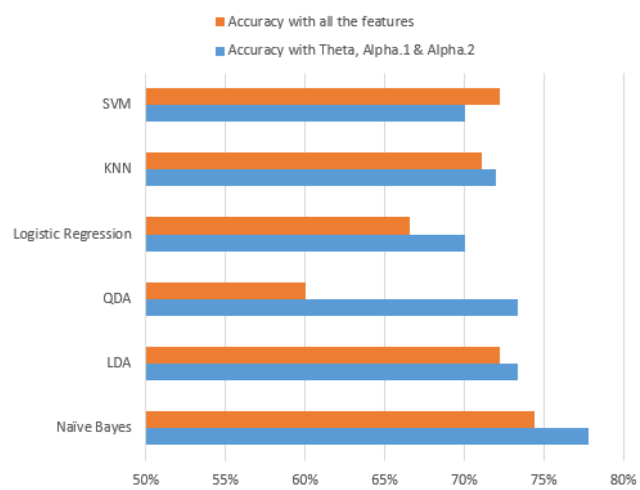

Figure 6. Classification accuracies with leave one student out CV

DISCUSSION

With the normalized data a remarkable improvement in cross validated classification rate from previous study on the same dataset was achieved. Normalization was required as the extreme values of each predictors for each subject is generally different. Normalizing with respect to each subject ensured that the data is scaled and is comparable among all the subjects.

It was noticed in the analysis that the state of confusion of the subject had great impact on Theta, Alpha.1 and Alpha.2 waves. Selecting only these three predictor variables increased the overall classification rates.

Boxplots of significant predictors across different subjects indicate a distinct division with respect to the state of confusion, as shown in Figure 4. We observe a strong relationship between the response (state of confusion) and predictors. However, this distinction level varies across each subject even when the normalization method we used. This indicates that the physiological response to a given situation is highly subjective.

Future Recommendations

The dataset used in this study has much fewer data points, this creates a suboptimal solution for advanced classification techniques.

As discussed earlier, the EEG data is subjective and user specific. A better normalization technique can be applied by using a baseline data for each subject and then evaluating the responses. Therefore, a more rigorous experiment should be done with more subjects, and a criterion needs to be developed to assess the baseline data for normalization purposes.

Instead of self-evaluation of the state of confusion, a performance assessment can be applied by asking students to undertake evaluation tests at the end of each video. This could lead to a better insight of their understanding of the taught content.

ACKNOWLEDGEMENTS

We would like to thank Mr. Haohang Wang, PhD. candidate at Carnegie Mellon University for sharing his experimental data.

REFERENCES

Ekandem, JI, Davis, T. A., Alvarez, I, James, M. T., Gilbert, J. E. (2012). Evaluating the ergonomics of BCI devices for research and experimentation. *Ergonomics Vol. 55, No. 5, May 2012, 592–598.*

James, G., Witten, D., Hastie, T., Tibshirani, R. (2013) An Introduction to Statistical Learning with applications in R. *Springer publication.*

Klimesch, W. (1999) EEG alpha and theta oscillations reflect cognitive and memory performance: a review and analysis, *Elsevier Science Brain Research Reviews 29. 169–195.*

Leon, J-A., Cano, J-M., Ibarrola, J. (2016) Are low cost Brain Computer Interface headsets ready for motor imagery applications? *Expert Systems with Applications 49 136–144.*

Milnik, V (2009) Instruction of Electrode Placement to the international 10–20-System. *Neurophysiologie-Labor.*

Nikolaeva, S. G., Roberts, D. N. (2016) Various Aspects of the Human Factor in Online Education: Quantitative Literature Review. *HFES, 2016.*

Park, J.-H., & Choi, H. J. (2009). Factors Influencing Adult Learners' Decision to Drop Out or Persist in Online Learning. *Educational Technology & Society, 12 (4), 207–217.*

Poltavski, D. V. (2015) The Use of Single-Electrode Wireless EEG in Biobehavioral Investigations. *Methods in Molecular Biology [1064-3745] vol: 1256 pg: 375 - 390.*

Scotti, S., Mauri, M., Cerutti, S., Mainardi, L., Villamira, M. (2005). Quantitative evaluation of distant student psychophysical responses during the e-learning processes. *IEEE Engineering in Medicine and Biology 27th Annual Conference.*

Wang, H., Li, Y., Hu, X., Yang, Y., Meng, Z., Chang, K. (2011) Using EEG to Improve Massive Open Online Courses Feedback Interaction. *CEUR workshop proceedings.*

Yang, D., Wen, M., Howley, I., Kraut, R., Rose, C. (2015). Exploring the effect of confusion in decision forums of Massive Open Online Courses. *L@S 2015 Learning.*

Zermeno, M., & Garza, L. (2016). Research analysis on MOOC course dropout and retention rates. *Turkish online journal of distance education, Volume 17.*