# Towards Predicting Task Performance from EEG Signals

[1,2,a]Michalis Papakostas, [1,2,b]Konstantinos Tsiakas, [2,c]Theodoros Giannakopoulos, [1,d]Fillia Makedon
[1]*Department of Computer Science, University of Texas at Arlington*
*Arlington, Texas*
[2]*Institute of Information and Telecommunications, NCSR "Demokritos"*
*Agia Paraskevi, Greece*
[a]*michalis.papakostas@mavs.uta.edu*
[b]*konstantinos.tsiakas@mavs.uta.edu*
[c]*tyiannak@gmail.com*
[d]*makedon@uta.edu*

***Abstract*—Smart wearable devices have lead to an increased need for processing and sharing large streams of physiological data in real-time. Modern Human-Machine Interaction (HMI) systems, especially applications designed for user training and assessment (e.g., educational or smart-rehabilitation systems), should be able to track and monitor those signals and adapt their parameters accordingly in order to optimally facilitate the special needs of each individual. Towards this end, we propose a passive Brain-Computer Interface (BCI), using a wireless non-intrusive EEG sensor under a robot assisted training task designed for cognitive assessment. As part of this ongoing work, we demonstrate our initial results on predicting user's task performance, from the EEG signals, before task completion. Our findings highlight the potentials of our hypotheses as we achieve a maximum accuracy rate equal to 74% when evaluated on 69 real subjects.**

***Keywords*-User Modeling; EEG; Behavior Monitoring; BCI; Machine Learning**

## I. INTRODUCTION

Monitoring user's performance and behavioral patterns during a task has been a very challenging research topic and has gained a lot of attention throughout the years. The main goal of such research is twofold. Firstly, monitoring user's performance and physiological signals using a dynamic framework can provide significant insights towards developing adaptable and personalized human-machine interaction and collaboration scenarios able to adjust their parameters on-the-fly. Secondly, such information can capture important user-behavior patterns, aiming to assist experts from other domains, such as education and smart rehabilitation, to improve the quality of their service [1], [2].

Moreover, recent technological advances both in terms of software and hardware, have allowed real-time access to different types of user's physiological signals in an unobtrusive way. Thus, creating breading ground for more sophisticated approaches to be applied. As related research has shown, passive BCIs are a very promising solution towards building safe and intelligent human-machine collaboration scenarios. [3].

Researchers in [4] proposed a passive BCI framework that adapts its behavior through interaction. Similarly, results showed in [5], [6] highlight the benefits of BCI frameworks for human-machine collaboration tasks in the domain of education and smart rehabilitation.

Inspired by the aforementioned research, we investigate the potentials of a BCI when applied on a robot assisted training system for working memory. Our experiments focus on predicting user's performance from the EEG data, before the user completes the task. According to our knowledge, this is the first effort that aims to directly predict user's performance on a specific task from the EEG data. Initial results indicate that there is a clear correlation between the EEG measurements and the final outcome of the task and that there are potential patterns able to capture certain cognitive behaviors across different users.

## II. THE SEQUENCE LEARNING TASK

Sequencing is the ability to arrange language, thoughts, information and actions in an effective order [7]. Extended research on the field of cognitive sciences has shown that sequence-learning tasks can be applied to evaluate human behaviors related to learning ability, short term memory and attention [8], [9].

Towards this direction, we developed the Sequence Learning (SL) task; a working memory task that evaluates the ability of a human to remember and repeat a sequence of items (e.g., letters, numbers, actions) [10]. For our experimental setup, we deploy the NAO[1] robot as a socially assistive robot that instructs, monitors and evaluates user's performance during the task. While performing the SL task, users have three buttons in front of them ("A", "B", "C") and the robot asks the user to repeat a given sequence of these letters by pressing the corresponding buttons. The game consists of four difficulty levels where, each level corresponds to a combination of 3,5,7 and 9 letters respectively.

[1]https://www.ald.softbankrobotics.com/en/cool-robots/nao

A complete session (human-robot interaction) consists of 25 turns/sequences. The level of each turn/sequence is decided randomly and all levels are equally distributed within a session. For the purposes of this research we considered a binary score at each turn, *success* or *fail*

## III. Data Collection

For the data collection, we recruited 69 CSE undergraduate and graduate students from the University of Texas at Arlington. Each user completed a single session of the SL task (25 turns/sequences). During the task, EEG signals were recorded using the Muse EEG headset[2], a low-cost and non-invasive EEG wearable device which, has been used previously for similar research purposes [11]. The Muse provides 4 channels of data; two coming from the forehead and two from behind the ears. The EEG signals were generated at a sampling rate of 220Hz. The device provides access to raw EEG signals as well as to a set of power spectral density measurements extracted from the raw data. The frequency bands provided by the device are $\delta$ (1-4 Hz), $\theta$ (5-8 Hz), $\alpha$ (9-13 Hz), $\beta$ (12-30 Hz) and $\gamma$ (30-50 Hz). Extensive details regarding the available data can be found at [10], [12]. At each turn of every session we store separately user's EEG captured during the *listening process* (robot pronounces a new sequence) from the EEG collected during the *acting-process* (user repeats the sequence by pressing the buttons). In the following Section, we describe our initial results on the task of predicting final user's performance (fail/success) at a single turn, using only the EEG from the listening process. In Figure-1, we illustrate the experimental setup.

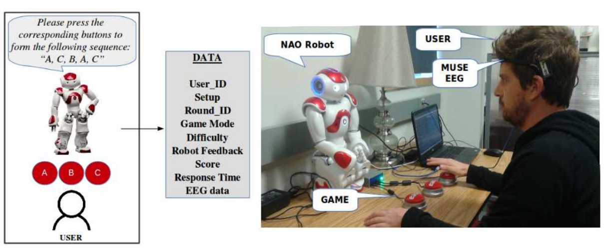


Figure 1. The Sequence Learning setup

The original data and details of the SL task along with the processed data and the code for the proposed work are available online[3][4].

## IV. Preliminary Results

As explained in the previous Section, we exploit the EEG signals captured during the *listening process*, to predict the final outcome of a single turn of the SL task. For validation purposes, we perform a 10-fold cross-validation across all users. At each fold, 80% of the users (55 subjects) were randomly picked for training, and the rest were used for testing. From each user, 25 interaction results were available, equal to the total number of turns/sequences played within a session. In total, we had 1375 training samples and 350 testing samples available at each fold. The distribution of the samples across the two classes always depended on the personal performance of the users picked each time for training. Across the 10 folds, the average prior-probabilities for success and fail in a single turn/sequence were 60% and 40% respectively.

### A. Feature Extraction

As discussed in Section-III, the Muse provides a set of frequency bands, extracted from the raw EEG in real-time through a digital signal processing component embedded in the device. For every frequency band, Muse estimates the absolute and relative band powers along with a band-power session score. According to Muse's documentation, the band session score is computed by comparing the current value of a band power to its history. Detailed information regarding the exact metrics and how they are estimated can be found at [12]. In total, for our experiments we exploited 15 different data streams, each coming from 4 different channels (see Section-III) thus, ending up with an initial feature representation of size equal to $4 \times 15 = 60$. More specifically, from every channel the following data streams were analyzed; $\delta$, $\theta$, $\alpha$, $\beta$ and $\gamma$ relative band powers, their respected absolute band powers and their session-score signals. From each of the 60 EEG feature-streams captured during the *listening process*, we extract the following statistical features:

- *standard deviation*
- *mean value*
- *maximum value*
- *minimum value*
- *spectral centroid*
  The center of gravity of the spectrum after applying FFT on the original signals.
- *spectral rolloff*
  The frequency below which 90% of the magnitude distribution of the spectrum is concentrated after applying FFT on the original signals.

The final feature vector representation consists of $60 * 6 = 360$ features, extracted from the EEG signals of a single subject and captured during the listening process, of a single turn/sequence of the SL task.

### B. Classification

For classification, we experimented with 5 different classification methods; SVMs , SVMS with an RBF kernel , Random Forests (RF), Extra Trees (ET) and Gradient Boosting (GB). For tuning, the *c* parameter of each classifier and for training each classification method the implementation described at [13] was applied. Before feeding the training data

[2] http://www.choosemuse.com/
[3] https://github.com/TsiakasK/sequence-learning
[4] https://github.com/MikeMpapa/EEG-Sequence-Learning

into the classifier, features are normalized to have $mean = 0$ and $std = 1$. In Table-I, we show the classification results. Since the two versions of SVM provided very similar results, we show only the linear-SVM evaluation as it was slightly superior. In all the cases, estimated time required for a single prediction was in the scale of *milliseconds*.

Table I
EEG CLASSIFICATION RESULTS

| | SVM | | GB | | RF | | ET | |
|---|---|---|---|---|---|---|---|---|
| | S | F | S | F | S | F | S | F |
| Prec | 0.75 | 0.48 | 0.81 | 0.56 | 0.89 | 0.24 | 0.91 | 0.2 |
| Rec | 0.69 | 0.55 | 0.78 | 0.6 | 0.69 | 0.54 | 0.69 | 0.54 |
| F1 | 0.72 | 0.51 | 0.79 | 0.58 | 0.78 | 0.33 | 0.78 | 0.29 |
| Acc | 0.65 | | **0.74** | | 0.67 | | 0.67 | |
| AVG F1 | 0.62 | | **0.69** | | 0.56 | | 0.54 | |

It is clear from the results that there is a significant statistical correlation between the EEG features and user's final performance. Despite the simplicity of the final features, the amount of captured information seems sufficient to provide a rough estimate with an average accuracy of 74% for the outcome of the task, when using a Gradient Boosting classifier.

## V. FUTURE WORK

These initial results indicate the potentials of the proposed approach. The most important aspect of this research is to investigate methods for building robust and adjustable user-models that can adapt to the behavior of each individual using a limited amount of new data. Towards this direction our future steps are to investigate additional features that can represent generic EEG patterns across different users, or user groups. Deep Learning approaches have the potential to significantly boost our results, since their superiority in many applications has proven their ability to capture invariant features. Moreover, the proposed problem seems that can be formulated better as a gradient optimization problem. To cross-validate our findings, we plan to evaluate our methods in other similar human-machine interaction tasks.

## VI. CONCLUSIONS

We propose a passive Brain-Computer Interface (BCI), using the Muse, a wireless non-intrusive EEG sensor under the scenario of the Sequence-learning task; a robot assisted training task designed for cognitive assessment. Our preliminary results highlight a clear correlation between user's brain activation and the actual outcome of the task, significantly before task completion. We evaluated our system on 69 real subjects following a user-independent modeling approach (each user was used either for training or for testing). Each interaction between the user and the system was represented by a feature vector of 360 statistical features extracted from the 60 available data streams captured at each timestamp by the Muse. Gradient Boosting classification provided the best classification results achieving a maximum accuracy of 74% with an average F1 of 69%


## ACKNOWLEDGMENT

This material is based upon work supported by the National Science Foundation under Grants 1565328, 1719031


## REFERENCES


[1] D. Szafir and B. Mutlu, "Pay attention!: designing adaptive agents that monitor and improve user engagement," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 2012, pp. 11–20.

[2] O. Barzilay and A. Wolf, "Adaptive rehabilitation games," *Journal of Electromyography and Kinesiology*, vol. 23, no. 1, pp. 182–189, 2013.

[3] Y. Wang and T.-P. Jung, "A collaborative brain-computer interface for improving human performance," *PloS one*, vol. 6, no. 5, p. e20422, 2011.

[4] Y.-T. Liu, N. R. Pal, A. R. Marathe, Y.-K. Wang, and C.-T. Lin, "Fuzzy decision-making fuser (fdmf) for integrating human-machine autonomous (hma) systems with adaptive evidence sources," *Frontiers in neuroscience*, vol. 11, p. 332, 2017.

[5] Q. Gao, L. Dou, A. N. Belkacem, and C. Chen, "Noninvasive electroencephalogram based control of a robotic arm for writing task using hybrid bci system," *BioMed Research International*, vol. 2017, 2017.

[6] A. Herweg, J. Gutzeit, S. Kleih, and A. Kübler, "Wheelchair control by elderly participants in a virtual environment with a brain-computer interface (bci) and tactile stimulation," *Biological psychology*, vol. 121, pp. 117–124, 2016.

[7] B. A. Clegg, G. J. DiGirolamo, and S. W. Keele, "Sequence learning," *Trends in cognitive sciences*, vol. 2, no. 8, pp. 275–281, 1998.

[8] S. E. Gathercole and A. D. Baddeley, *Working memory and language*. Psychology Press, 2014.

[9] K. M. Thomas, R. H. Hunt, N. Vizueta, T. Sommer, S. Durston, Y. Yang, and M. S. Worden, "Evidence of developmental differences in implicit sequence learning: an fmri study of children and adults," *Journal of cognitive neuroscience*, vol. 16, no. 8, pp. 1339–1351, 2004.

[10] K. Tsiakas, C. Abellanoza, M. Abujelala, M. Papakostas, T. Makada, and F. Makedon, "Towards designing a socially assistive robot for adaptive and personalized cognitive training."

[11] X. Liu, P.-N. Tan, L. Liu, and S. J. Simske, "Automated classification of eeg signals for predicting students cognitive state during learning," in *Proceedings of the International Conference on Web Intelligence*. ACM, 2017, pp. 442–450.

[12] "The muse," http://developer.choosemuse.com/research-tools/available-data.

[13] T. Giannakopoulos, "pyaudioanalysis: An open-source python library for audio signal analysis," *PloS one*, vol. 10, no. 12, 2015.