This is a post-peer-review, pre-copyedit version of an article published in IFMBE Proceedings book series (IFMBE, volume 76). The final authenticated version is available online at: https://doi.org/10.1007/978-3-030-31635-8_225

This is a post-peer-review, pre-copyedit version of an article published in IFMBE Proceedings book series (IFMBE, volume 76). The final authenticated version is available online at: **https://doi.org/10.1007/978-3-030-31635-8_225**

Slow cortical potential BCI classification using Sparse variational Bayesian logistic regression with automatic relevance determination

Aleksandar Miladinović[1], Miloš Ajčević[1,3,4], Piero Paolo Battaglini[2], Giulia Silveri[1], Gaia Ciacchi[2], Giulietta Morra[2], Joanna Jarmolowska[2], Agostino Accardo[1],

[1]Department of Engineering and Architecture, University of Trieste, Via A. Valerio 10 34127 Trieste, Italy
[2]Department of Life Sciences, University of Trieste, Via Fleming 22, 34127 Trieste, Italy
[3]Department of Medicine Surgery and Health Sciences, University of Trieste, Trieste, Italy
[4]NEUROFARBA Department, Neuroscience Section, University of Florence, Italy

Abstract. Detecting P300 slow-cortical ERPs poses a considerable challenge in signal processing due to the complex and non-stationary characteristics of a single-trial EEG signal. EEG-based neurofeedback training is a possible strategy to improve the social abilities in Autism-Spectrum Disorder (ASD) subjects. This paper presents a BCI P300 ERPs based protocol optimization used for the enhancement of joint-attention skills in ASD subjects, using a robust logistic regression with Automatic Relevance Determination based on full Variational Bayesian inference (VB-ARD). The performance of the proposed approach was investigated utilizing the IFMBE 2019 Scientific Challenge Competition dataset, which consisted of 15 ASD subjects who underwent a total of 7 BCI sessions spread over 4 months. The results showed that the proposed VB-ARD approach eliminates irrelevant channels and features effectively, producing a robust sparse model with 81.5 ± 0.12 % accuracy in relatively short modeling computational time 19.3 ± 1.4 sec, and it outperforms the standard regularized logistic regression in terms of accuracy and speed needed to produce the BCI model. This paper demonstrated the effectiveness of the probabilistic approach using Bayesian inference for the production of a robust BCI model. Considering the good classification accuracy over sessions and fast modeling time the proposed method could be a useful tool used for the BCI based protocol for the improvement of joint-attention ability in ASD subjects.

Keywords: BCI, Modeling, P300, ERP, Variational Bayesian Inference, Automatic Relevance Determination, Autism-Spectrum Disorder.

This is a post-peer-review, pre-copyedit version of an article published in IFMBE Proceedings book series (IFMBE, volume 76). The final authenticated version is available online at: https://doi.org/10.1007/978-3-030-31635-8_225

1 Introduction

Autism Spectrum Disorder (ASD) is a set of pervasive and sustained neurodevelopmental conditions characterized by persistent deficits in social interaction, alongside restricted, repetitive patterns of behavior, interests, or activities [1]. Several studies found that using EEG-based neurofeedback training is a possible way to improve the social skills in ASD subjects [2]. Electroencephalography (EEG) based brain-computer interface (BCI) is a system that acquires brain signals and provides feedback according to the performance of the participant, and it has been used, not only for the attention improvement [3], but also for motor neurorehabilitation [4–8]. In addition, the employment of Virtual Reality (VR) improves the effectiveness of BCI in general, and also the applicability of these skills in real life [9–11].

A novel social attention training paradigm (VR P300-based BCI paradigm) based on P300 ERPs for the enhancement of joint-attention skills was proposed in the paper [12, 13]. A P300 is a huge positive voltage with a peak latency around 300 ms after the appearance of a cognitive attended rare stimulus, and it can be used to monitor the attentional performance of ASD subjects.

In the paper [12, 13] the P300-ERP based paradigm was used to acquire the data, in the VR immersive environment where the subjects had to follow a non-verbal social agent cue performed by an avatar. The avatar moved its head toward a target object signaling to the participants where to focus on a so-called joint attention task [14]. Then the participants received feedback based on the P300 acquired signal allowing them to adjust and improve this social attentional ability. Therefore, the BCI system allows to detect the target object toward which the participants were focusing their attention with the highest possible accuracy. Different machine learning methods can be applied to produce a model which adequately provide an output that indicates the subject's mental state. However, without pre-processing, a high number of channels and features can lead to overfitting. To overcome this challenge, a technique that eliminates irrelevant channels and features is preferable.

Consequently, in this paper, we investigated a technique based on variational Bayes automatic relevance determination that is able to prune irrelevant inputs,

This is a post-peer-review, pre-copyedit version of an article published in IFMBE Proceedings book series (IFMBE, volume 76). The final authenticated version is available online at: ***https://doi.org/10.1007/978-3-030-31635-8_225***

together with a logistic regression approach with a standard regularization method to create an accurate and robust BCI model based on P300 ERPs.

2 MATERIALS AND METHODS

2.1 The EEG dataset

The proposed approach was evaluated utilizing the IFMBE 2019 Scientific Challenge Competition dataset [12, 13], which was acquired from 15 subjects who underwent a total of 7 sessions spread over 4 months: four in the first week and then once a month. This dataset represents the complete EEG recordings of a feasibility clinical trial (clinical-trial ID: NCT02445625 — clinicaltrials.gov). The recording was performed using amplifier g.Nautilus (gTEC, Austria) coupled with a VR system (Oculus Rift Development Kit 2 headset); the data were acquired from 8 dry electrodes positioned on the standard locations - C3, Cz, C4, CPz, P3, Pz, P4, POz, with the reference placed on the right ear and the ground electrode placed at AFz at 250Hz sampling rate. The participants were immersed in a virtual environment, a bedroom with a series of objects that were used as targets. The paradigm was subdivided into two parts: (1) calibration phase, where the object target was explicit; (2) online phase, where the participants had to participate in the joint attention task after which they received the feedback according to their performance.

For each session, the training data consisted of 20 blocks (with 10 runs each) while the online phase of 50 blocks (with a variable number of runs each) was used as a test set. In each run, 8 objects flashing in random order were presented in the virtual scene (1. a wooden plane hanging from the ceiling; 2. a printer on a shelf; 3. a corkboard on the wall; 4. a laptop on a table; 5. a ball on the ground; 6. a radio on top of a dresser; 7. a picture on the wall; 8. books on a shelf). In both phases, the participants had to mentally count how many times the target object flashed creating an EEG response later used for BCI.

On the basis of the training dataset, it is possible to build a BCI model capable to identify, with the highest possible accuracy, which target object the participant was focusing on during the online phase.

This is a post-peer-review, pre-copyedit version of an article published in IFMBE Proceedings book series (IFMBE, volume 76). The final authenticated version is available online at: https://doi.org/10.1007/978-3-030-31635-8_225

2.2 Data processing and BCI modeling

The data has been analyzed using Matlab (Mathworks version R2018b) and BCILAB framework [15]. The dataset had been already pre-filtered from 2 to 30Hz, therefore no additional frequency filter is applied by our approach at this stage. The window means approach [16], that has been widely applied to event-related slow-changing brain dynamics, such as, the perception of self-induced errors [17], machine-induced errors and/or surprisal [18, 19], prediction of movement intent [20], or (c)overt attention [21], was used. The paradigm consisted of computing 50ms windows without overlap on the segment from 100ms to 1000ms from stimulus onset, producing 18 features per channel used in the following machine learning stage. Since the dimensionality of the feature space was high (number of channels * number of windows), the robust classifiers with advance irrelevant feature pruning had to be applied. At the same time, it is important that the BCI model using sparse features can be produced in a reasonable time (i.e. at most within 20min after a training session), so preferable approach would be one without a priori unknown parameters obtainable using time-consuming repetitive cross-validation based optimization procedures. Consequently, in this work we investigated the use of Sparse variational Bayesian logistic regression with automatic relevance determination (VB-ARD) [22], to classify ERPs in ASD subjects and compared it to the logistic regression approach that utilizes standard L1-norm regularization strategy by multiple cross-validations. Briefly, Automatic relevance determination (ARD) consists of assigning an individual hyper-prior to each regression coefficient separately and determines the relevance of each of them throughout the optimization of the feature-related weights in order to effectively eliminate irrelevant ones. This approach conversely to [23–25] does not tune hyper-parameters by maximizing the marginal likelihood but applies the full Bayesian treatment, and finds the ARD hyper-posteriors by variational Bayesian inference [22]. Therefore, functions associated with these weights play no role in the predictions made by the model and are effectively pruned out, resulting in a feature sparse model. Furthermore, the fact that the approach does not compute only a point estimate of the weights, but a full posterior distribution, enables us to use more

This is a post-peer-review, pre-copyedit version of an article published in IFMBE Proceedings book series (IFMBE, volume 76). The final authenticated version is available online at: https://doi.org/10.1007/978-3-030-31635-8_225

advanced post-processing of the model and outputs. The Bayesian framework enables optimal sparsity and enables the number of features and channels to be estimated automatically from the training set without requiring time-consuming cross-validation. The VB-ARD model was compared with an L1-norm logistic regression (L1-LOG) model regularized by chronological 5-fold cross-validation techniques with 5 indices safety margin.

2.3 Postprocessing and Classification

The resulting regression binary output with corresponding class belonging probabilities provides information for the existence or non-existence of the target ERPs in each run. Finally, the mean probability of all runs in the block had been calculated for each label, and the label with the highest probability has been chosen as the output. In that way, the result averaging has been performed in the post-processing stage.

2.4 Performance measure and Statistical analysis

The performance of two approaches was measured in terms of model accuracy on each subject and in all 7 sessions and computation time on desktop PC with Intel Core i7 CPU @ 2.67 GHz with 16 GB RAM and Windows 10 OS was measured. A paired t-test approach was used to assess the differences in model accuracy and model computation time between VB-ARD and L1-LOG methods.

3 Results

Mean ± 1SD model accuracy and model computation time for each subject for VB-ARD and L1-LOG approaches are reported in Table 1. In general, both approaches showed high accuracy, however, VB-ARD produced model presented statistically higher accuracy compared to L1-LOG (81.5 ± 0.12 vs 79.0 ± 0.14 %; $p=0.0014$) and extremely lower model computation time (19.3 ± 1.4 vs 643.3 ± 83.0 sec; $p<0.0001$).

This is a post-peer-review, pre-copyedit version of an article published in IFMBE Proceedings book series (IFMBE, volume 76). The final authenticated version is available online at: ***https://doi.org/10.1007/978-3-030-31635-8_225***

The scatter plot for each subject for the accuracy of both approaches was reported in Figure 1. The topoplot of BCI models calculated with two approaches on the single subject are reported in Figure 2. It can be observed that VB-ARD approach conversely to the L1-LOG for different time windows produced a sparse model by assigning weights close to zero to less relevant channels and emphasizing the relevant ones.

Table 1. Mean ± 1SD values of model accuracy and model computation time for each subject for VB-ARD and L1-LOG approaches.

Subject	**VB-ARD**		**L1-LOG**	
	Accuracy Mean ± SD (%)	Comp. time Mean ± SD (sec)	Accuracy Mean ± SD (%)	Comp. Time Mean ± SD (sec)
1	71 ± 15	19.44 ± 1.87	58 ± 24	730.62 ± 131.04
2	89 ± 7	19.94 ± 0.86	91 ± 6	633.1 ± 93.26
3	77 ± 17	18.2 ± 1.55	77 ± 16	632.29 ± 82.35
4	89 ± 8	20.02 ± 0.75	87 ± 6	679.38 ± 100.91
5	86 ± 12	19.58 ± 0.51	83 ± 14	661 ± 72.41
6	82 ± 6	18.95 ± 1.58	79 ± 8	584.59 ± 18.4
7	77 ± 7	19.68 ± 0.97	75 ± 4	673.84 ± 41.44
8	87 ± 14	19.39 ± 0.9	85 ± 14	617.24 ± 47.56
9	78 ± 9	17.31 ± 0.7	72 ± 13	572.8 ± 32.27
10	86 ± 8	20.05 ± 1.12	84 ± 8	604.5 ± 61.67
11	85 ± 8	20.46 ± 1.23	85 ± 9	609.31 ± 52.03
12	82 ± 13	18.09 ± 1.4	81 ± 17	600.5 ± 59.64
13	74 ± 17	20.05 ± 0.85	71 ± 19	703.2 ± 92.13
14	73 ± 12	18.24 ± 1.43	73 ± 13	652.99 ± 87.72
15	87 ± 6	19.91 ± 1.61	87 ± 7	694.73 ± 66.3

This is a post-peer-review, pre-copyedit version of an article published in IFMBE Proceedings book series (IFMBE, volume 76). The final authenticated version is available online at: ***https://doi.org/10.1007/978-3-030-31635-8_225***

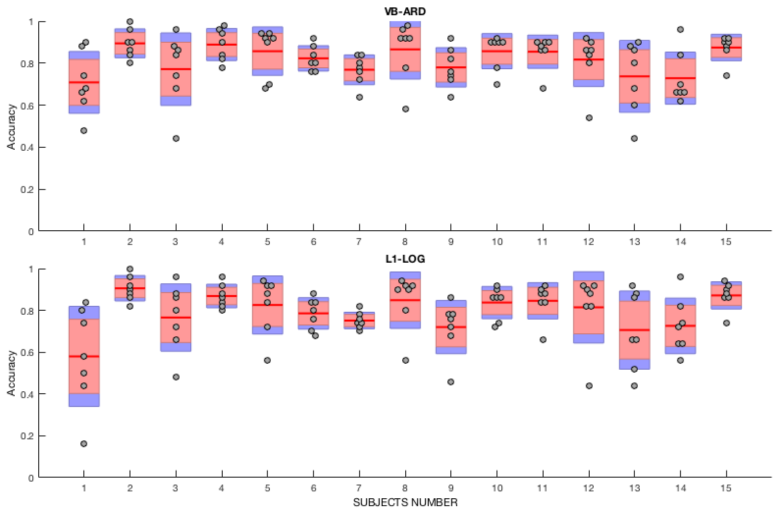
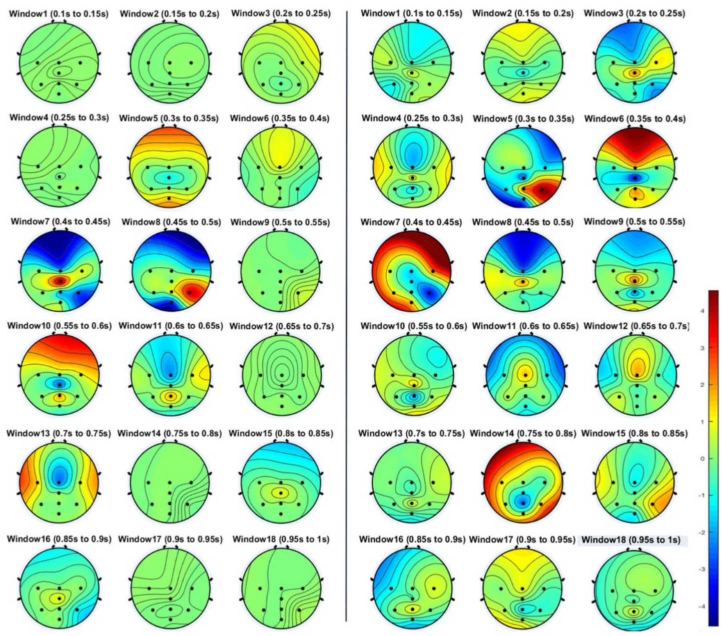


Fig. 1. Plot of accuracy for each subject and each session (points). Points are layed over a 1.96 SEM (95% confidence interval) in red and a 1 SD in blue. Upper panel - VB-ARD model results; Lower panel - L1-LOG model results.



This is a post-peer-review, pre-copyedit version of an article published in IFMBE Proceedings book series (IFMBE, volume 76). The final authenticated version is available online at: ***https://doi.org/10.1007/978-3-030-31635-8_225***

Fig. 2. BCI models on one subject. Left panel - Space feature model produced with VB-ARD approach. Right panel - Model produced with L1-LOG method.

4 Discussion

In order to perform a P300 BCI protocol for the improvement of joint-attention skills in Autism-Spectrum Disorder (ASD) subjects, an appropriate approach for the BCI classification is needed. Detecting P300 slow-cortical ERPs poses a considerable challenge to signal to process due to the complex and non-stationary characteristics of a single-trial EEG signal.

The main finding of this study is the high performance of a BCI protocol optimization using a robust logistic regression with automatic relevance determination based on full variational Bayesian inference. Our results showed that the proposed VB-ARD approach eliminates irrelevant channels and features effectively, producing a robust sparse model with 81.5 ± 0.12 % accuracy in relatively short modeling computational time 19.3 ± 1.4 sec, and it outperforms the standard regularized logistic regression in terms of accuracy and speed needed to produce the BCI model.

Comparison of two BCI models estimated with two different methods revealed that as expected VB-ARD in comparison to L1-LOG approach selects only the relevant time windows and channels that can be also used to study and evaluate the neurophysiological plausibility of produced model. Considering the parsimony of the calculated model we can expect that it represents rather generalized solution resulting in higher robustness, that have to be evaluated in the future studies.

In conclusion, we have presented a BCI computational method based on Bayesian variational inference for P300 ERP detection, which is able to extract and effectively individualize the informative features for binary regression/classification problem of target and non-target objects in the joint attention experiment in ASD subjects. As the approach provides a full posterior distribution it is expected that more advanced post-processing may further improve the performance. Finally, our approach does not average waveforms of every single trial in the pre-processing step but directly extract features with their corresponding class belonging probabilities, so the number of needed runs in the online phase can be adjusted based on so-far estimated discriminatory properties.

Conflict of Interest Statement

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Acknowledgment

The authors would like to thank the organizers (IFMBE Scientific Challenge; http://www.medicon2019.org/scientific-challenge/) and the providers of the BCI Competition Dataset [12, 13].

A. Miladinović is supported by European Social Fund (ESF) - FVG

Work partially supported by Master in Clinical Engineering, University of Trieste

References

1. American Psychiatric Association: Diagnostic and Statistical Manual of Mental Disorders (DSM). CoDAS. (2013).
2. Friedrich, E.V.C., Suttie, N., Sivanathan, A., Lim, T., Louchart, S., Pineda, J.A.: Brainâ€“computer interface game applications for combined neurofeedback and biofeedback treatment for children on the autism spectrum. Front. Neuroeng. (2014). https://doi.org/10.3389/fneng.2014.00021.
3. Rohani, D.A., Sorensen, H.B.D., Puthusserypady, S.: Brain-computer interface using P300 and virtual reality: A gaming approach for treating ADHD. In: 2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBC 2014 (2014). https://doi.org/10.1109/EMBC.2014.6944403.
4. van Dokkum, L.E.H., Ward, T., Laffont, I.: Brain computer interfaces for neurorehabilitation-its current status as a rehabilitation strategy post-stroke. Ann. Phys. Rehabil. Med. (2015). https://doi.org/10.1016/j.rehab.2014.09.016.
5. Miladinović, A., Barabaro, A., Eddi, V., Ajčević, M., Accardo, A., Battaglini, P.P., Jarmolowska, J.: Combined and singular effects of Action Observation and Motor Imagery paradigms on resting-state sensorimotor rhythms: *XV* Mediterranean Conference on Medical and Biological Engineering and Computing 2019 - *in press*
6. Ang, K.K., Guan, C.: Brain-computer interface for neurorehabilitation of upper limb after stroke. Proc. IEEE. (2015). https://doi.org/10.1109/JPROC.2015.2415800.
7. Tang, N., Guan, C., Ang, K.K., Phua, K.S., Chew, E.: Motor imagery-assisted brain-computer interface for gait retraining in neurorehabilitation in chronic stroke. Ann. Phys. Rehabil. Med. (2018). https://doi.org/10.1016/j.rehab.2018.05.431.
8. Vourvopoulos, A., Pardo, O.M., Lefebvre, S., Neureither, M., Saldana, D., Jahng, E.,

This is a post-peer-review, pre-copyedit version of an article published in IFMBE Proceedings book series (IFMBE, volume 76). The final authenticated version is available online at: https://doi.org/10.1007/978-3-030-31635-8_225

Liew, S.-L.: Effects of a Brain-Computer Interface With Virtual Reality (VR) Neurofeedback: A Pilot Study in Chronic Stroke Patients. Front. Hum. Neurosci. 13, (2019). https://doi.org/10.3389/fnhum.2019.00210.
9. Wainer, A.L., Ingersoll, B.R.: The use of innovative computer technology for teaching social communication to individuals with autism spectrum disorders, (2011). https://doi.org/10.1016/j.rasd.2010.08.002.
10. Bekele, E., Crittendon, J., Zheng, Z., Swanson, A., Weitlauf, A., Warren, Z., Sarkar, N.: Assessing the utility of a virtual environment for enhancing facial affect recognition in adolescents with autism. J. Autism Dev. Disord. (2014). https://doi.org/10.1007/s10803-014-2035-8.
11. Georgescu, A.L., Kuzmanovic, B., Roth, D., Bente, G., Vogeley, K.: The use of virtual characters to assess and train non-verbal communication in high-functioning autism. Front. Hum. Neurosci. (2014). https://doi.org/10.3389/fnhum.2014.00807.
12. Amaral, C.P., Simões, M.A., Mouga, S., Andrade, J., Castelo-Branco, M.: A novel Brain Computer Interface for classification of social joint attention in autism and comparison of 3 experimental setups: A feasibility study. J. Neurosci. Methods. (2017). https://doi.org/10.1016/j.jneumeth.2017.07.029.
13. Amaral, C., Mouga, S., Simões, M., Pereira, H.C., Bernardino, I., Quental, H., Playle, R., McNamara, R., Oliveira, G., Castelo-Branco, M.: A feasibility clinical trial to improve social attention in Autistic Spectrum Disorder (ASD) using a brain computer interface. Front. Neurosci. (2018). https://doi.org/10.3389/fnins.2018.00477.
14. Bakeman, R., Adamson, L.B.: Coordinating attention to people and objects in mother-infant and peer-infant interaction. Child Dev. (1984).
15. Kothe, C.A., Makeig, S.: BCILAB: a platform for brain–computer interface development. J. Neural Eng. 10, 056014 (2013). https://doi.org/10.1088/1741-2560/10/5/056014.
16. Blankertz, B., Lemm, S., Treder, M., Haufe, S., Müller, K.R.: Single-trial analysis and classification of ERP components - A tutorial. Neuroimage. (2011). https://doi.org/10.1016/j.neuroimage.2010.06.048.
17. Blankertz, B., Schäfer, C., Dornhege, G., Curio, G.: Single Trial Detection of EEG Error Potentials: A Tool for Increasing BCI Transmission Rates. Presented at the (2002). https://doi.org/10.1007/3-540-46084-5_184.
18. Ferrez, P.W., Del R. Millán, J.: Error-related EEG potentials generated during simulated brain-computer interaction. IEEE Trans. Biomed. Eng. (2008). https://doi.org/10.1109/TBME.2007.908083.
19. Zander, T.O., Kothe, C., Welke, S., Roetting, M.: Utilizing secondary input from passive brain-computer interfaces for enhancing human-machine interaction. In: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) (2009). https://doi.org/10.1007/978-3-642-02812-0_86.
20. Blankertz, B., Curio, G., Müller, K.-R.: Classifying single trial EEG: Towards brain computer interfacing. Adv. Neural Inf. Process. Syst. Conf. (2002). https://doi.org/10.1.1.19.8038.
21. Treder, M.S., Blankertz, B.: (C)overt attention and visual speller design in an ERP-based brain-computer interface. Behav. Brain Funct. (2010). https://doi.org/10.1186/1744-9081-6-28.
22. Drugowitsch, J.: Variational Bayesian inference for linear and logistic regression. ArXiv e-prints. (2013).
23. MacKay, D.J.C., Systems, N.: Bayesian Interpolation. Neural Comput. (1992). https://doi.org/10.1162/neco.1992.4.3.415.
24. Neal, R.M.: Bayesian Learning for Neural Networks. (1996). https://doi.org/10.1007/978-1-4612-0745-0.

25. Tipping, M.E.: Sparse Bayesian Learning and the Relevance Vector Machine. J. Mach. Learn. Res. (2001). https://doi.org/10.1162/15324430152748236.