

Disjunctive Normal Unsupervised LDA for P300-based Brain-Computer Interfaces

Majed Elwardy

Faculty of Engineering

and Natural Sciences,

Sabanci University, Istanbul, Turkey

Email: melwardy@sabanciuniv.edu

Tolga Taşdizen

Electrical and Computer

Engineering Department

University of Utah, USA

Email: tolga@sci.utah.edu

Müjdat Çetin

Faculty of Engineering

and Natural Sciences,

Sabanci University, Istanbul, Turkey

Email: mctetin@sabanciuniv.edu

Abstract—Can people use text-entry based brain-computer interface (BCI) systems and start a free spelling mode without any calibration session? Brain activities differ largely across people and across sessions for the same user. Thus, how can the text-entry system classify the desired character among the other characters in the P300-based BCI speller matrix? In this paper, we introduce a new unsupervised classifier for a P300-based BCI speller, which uses a disjunctive normal form representation to define an energy function involving a logistic sigmoid function for classification. Our proposed classifier updates the initialized random weights performing classification for the P300 signals from the recorded data exploiting the knowledge of the sequence of row/column highlights. To verify the effectiveness of the proposed method, we performed an experimental analysis on data from 7 healthy subjects, collected in our laboratory. We compare the proposed unsupervised method to a baseline supervised linear discriminant analysis (LDA) classifier and demonstrate its effectiveness.

Keywords—Brain-computer interface, P300 Speller, calibration session, unsupervised classifier, LDA

I. INTRODUCTION

A significant number of individuals suffer from losing all voluntary muscle control due to amyotrophic lateral sclerosis (ALS), traumatic brain injuries or spinal cord injuries [1]. Although the motor pathway is lost, neuronal activity of the brain still works in many of these cases. A brain-computer interface (BCI) aims to establish a direct communication channel between the brain and a computer or machine so disabled individuals can interact with the real-world [2]. Studies over last two decades have shown that the electroencephalogram (EEG) measured through the scalp can be used as the cornerstone for BCI [3]. Besides decoding a user's intent signals, it can be used to provide input signals in many applications including text entry [4], robotic arm control [5], cursor control [6].

The P300 speller is one of the most common BCI-based text-entry systems, which allows subjects to write text on the computer screen. Farwell and Donchin [4] demonstrated the first P300 speller paradigm which is also called the oddball paradigm. P300 is an event-related potential (ERP) elicited in the brain as a response to a visual or auditory stimulus. It is a positive deflection measured around the parietal lobe,

nearly 300 ms after the occurrence of the attended stimulus [7]. The system allows people to spell words and numbers by focusing on the desired character in a matrix shown on the screen (see Fig. 1). When the desired character is highlighted, the subject attends to the unexpected stimuli and a P300 wave is generated. The character which the user intends to type can be inferred by the intersection of the detected P300 responses in the sequence of row/column highlights. EEG signals suffer from low signal to noise ratio (SNR) due to several factors including the variability in brain activities. Therefore, P300 spellers need several stimulus repetitions to increase the classification accuracy [8].

One of the common problems in BCIs is the calibration process. The brain signals vary across people and across sessions for the same user [9]. For this reason, supervised training methods based on calibration sessions involving labeled training data are usually used. Furthermore, the BCI system should be trained for a specific person. The downsides of having to use such sessions include the consumption of additional time and increased fatigue for the users. Furthermore, such sessions might have to be repeated to account for any non-stationary behavior of the brain signals over the course of system use.

The work in this paper provides a contribution towards addressing these problems by proposing a new unsupervised classifier for P300-based spellers. In this approach, the disjunctive normal form plays a role in forming an energy function, which allows to update the randomly initialized classifier weights by using the logistic sigmoid function for classification and by exploiting the knowledge of the sequence of row/column highlights [10]. The idea is that one round of row/column highlights in the speller matrix should evoke a P300 response only after two (one row and one column) of the highlights. Note that exploiting this fact does not require knowledge of the labels of the data, hence this idea can be a basis for unsupervised learning. There have been several pieces of work on unsupervised methods for P300-based BCI spellers. An unsupervised method was proposed by Lu et al. [11]. Although that unsupervised classifier has also been applied to P300 data, it still needs some labeled data to train a subject independent classifier which then goes through adaptation. Another recent unsupervised classification method, based on a Bayesian model, has been proposed by Kindermans et al. [12].

This work has been supported by a graduate fellowship from the Scientific and Technological Research Council of Turkey.

There also exist semi-supervised adaptation methods which involve supervised training followed by adaptation of the classifier with the incoming EEG data [13]. We evaluate our disjunctive normal unsupervised linear discriminant analysis (DNUL) approach on EEG data collected in our laboratory and demonstrate its effectiveness in unsupervised learning through a comparison with a supervised method. We also demonstrate the sequential learning/adaptation capability of our approach as test data are collected.

II. METHODS

The following sections will provide the details of our proposed unsupervised classification method based on the disjunctive normal form [10].

A. Model Architecture

Consider a two-class classification problem: $\mathbf{C} = \{0, 1\}$, for which we observe the data samples (x_1, x_2, \dots, x_n) where n is the number of samples. Let us assume one row/column flash among a full sequence of flashes comes from the class $\mathbf{C} = 1$ and all other $(n - 1)$ row/column flashes in that sequence come from the class $\mathbf{C} = 0$ where $\mathbf{C} = 1$ corresponds to row/column containing the target letter and $\mathbf{C} = 0$ corresponds to row/column not containing the target letter. Let $y_j = f(x_j)$ for $j \in \{1, \dots, n\}$ where $y \in \{0, 1\}$ and $f(x_j)$ is the classification function. Let us define the following Boolean indicator function, which we will call the one-vs-all function $g(\mathbf{y})$.

$$g(y_1, y_2, \dots, y_n) = \begin{cases} 1, & \text{if only one argument is 1;} \\ 0, & \text{otherwise.} \end{cases} \quad (1)$$

Any Boolean function can be written as a disjunction of conjunctions, also known as the disjunctive normal form [14].

$$\begin{aligned} E(x) &= g(y_1, y_2, \dots, y_n) \\ &= (y_1, y'_2, \dots, y'_n) \cup (y'_1, y_2, \dots, y'_n) \cup \dots \cup (y'_1, y'_2, \dots, y_n) \end{aligned} \quad (2)$$

Furthermore, allowing M repeated observations, we define:

$$E(x) = \sum_i^M g(f(x_{1i}), f(x_{2i}), \dots, f(x_{ni})) \quad (3)$$

where we can relax the function f so it has real valued outputs in $[0, 1]$ rather than binary. We perform such relaxation through a logistic sigmoid function, where β is a sensitivity parameter.

$$f(x_{ji}) = \frac{1}{1 + e^{-\beta(\sum_j^n w_{ij} x_j + b_{ij})}} \quad (4)$$

Using De Morgan's laws and products of conjunctions yields the following differentiable energy function [14].

$$E(x) = \sum_i^M \left(1 - \prod_j^n \underbrace{\left(1 - f(x_{ji}) \prod_{k \neq j}^n (1 - f(x_{ki})) \right)}_{Q_j} \right) \quad (5)$$

where M denotes the number of rounds for row/column highlights.

B. Model Initialization

Let us consider a P300 speller paradigm with $\Gamma = \{(x, L(x))\}$, where x denotes the data and $L(x)$ denotes the binary class label corresponding to x . Furthermore, let Γ^+ denote class $L = 1$ corresponding to the desired target and Γ^- denote class $L = 0$ corresponding to the non-desired target.

Since the model is designed to work in an unsupervised fashion, the labels for learning the model will not be available to the algorithm. We will use the disjunctive normal form-based energy function in (5) to classify the two classes without using any labels. The weights w_{ij} of the disjunctive normal unsupervised linear discriminant (DNUL) classifier are randomly initialized and the bias terms set to 1. Consider a speller matrix as in Fig. 1. We have 6x6 characters that means the target character needs a set of row/column intensifications (highlights) to cover the matrix. We call the set of intensifications covering the entire array a trial group. Therefore, $n = 6$ in our algorithm.

The sigmoid function in (4) takes the value 0.5 in the middle of the classification line between two classes. The goal is to design a classifier to put the data from the desired class in Γ^+ when $f(x_{ji}) \geq 0.5$ and data from non-desired class in Γ^- when $f(x_{ji}) < 0.5$ by optimizing the energy function in (5).

C. Model Optimization

In order to learn the DNUL classifier, we use gradient ascent to maximize the energy function by taking the partial derivatives of the energy function with respect to the weights. The gradient of the energy function in (5) is given by:

$$\begin{aligned} \frac{\partial E}{\partial w} &= - \sum_i^M \frac{\partial}{\partial w} \left(\prod_j^n Q_j \right) \\ &= - \sum_i^M \sum_j^n \left(\frac{\partial Q_j}{\partial w} \prod_{l \neq j}^n Q_l \right) \\ \frac{\partial Q_j}{\partial w} &= - \frac{\partial}{\partial w} f(x_{ji}) \prod_{k \neq j}^n (1 - f(x_{ki})) \\ &\quad - f(x_{ji}) \sum_{p \neq j}^n \left(- \frac{\partial}{\partial w} f(x_{pi}) \prod_{k \neq p, j}^n (1 - f(x_{ki})) \right) \end{aligned} \quad (6)$$

The model performs iterations till the DNUL classifier converges updating the weights at each iteration: (7) where α is the step size. The bias term is included in the weight vector.

$$w_{ij}^{new} = w_{ij} + \alpha \frac{\partial E}{\partial w_{ij}} \quad (7)$$

III. EXPERIMENTAL RESULTS

The proposed DNUL and supervised LDA classification techniques are evaluated in this section with a real P300-based speller dataset. In this study, 7 male healthy subjects performed offline spelling, whose ages are between 18 and 30. Only two of the subjects had prior BCI experience before. These datasets were recorded in our lab at Sabancı University [15]. Temporal EEG data was recorded from 12 active channels



Fig. 1. Interface of P300-based speller matrix used in this study.

during the experiment which were placed at Fp1, Fp2, Fz, Cz, Pz, Oz, P3, P4, Po7, and Po8 locations according to the international 10-20 system, in addition to the two auxiliary electrodes for reference. The data are sampled at 2048 Hz. The recorded data are bandpass filtered in 1-12 Hz and decimated by 64. The signals are divided into one-second epochs which are used as the feature vectors for classification. The 6×6 spelling matrix uses the most common stimulus type. The intensification covers the rows and columns of the matrix in a block-randomized fashion. Each intensification flashes exactly once with an inter-stimulus interval (ISI) of 125 ms; the intensification duration of 50 ms and the remaining 75 ms waiting for the next intensification. Each subject recorded two sessions: one for the training session and one for the test session. The training session involved spelling 14 characters forming 2 Turkish words. The test session involved spelling 26 characters forming 4 Turkish words. In this work, we split the test dataset into two versions, one with 14 characters and the other one with 26 characters as shown in Table I. The data were recorded with the BioSemi ActiView software. We used the data preprocessing methods described in detail in [16].

The DNUL classification model is one of the most challenging as it starts initially unlearned without using labels. In this case, there is no need for the training session, the approach just evaluates the model on the upcoming EEG data. The accuracies presented in this study refer to the spelled characters. Most systems, including ours, classify the individual intensifications and combine the outputs to predict the spelled character.

The number of trial groups for spelling a character was pre-defined, the maximum number of trial groups recorded in these datasets was 15. Our experiments are divided into two categories. The first four experiments involve offline analysis (batch mode) as shown in Fig. 2. The last experiment, depicted in Fig. 3, is designed to simulate online spelling (sequential mode) in order to evaluate the sequential adaptation process of the classifier.

The initialization parameters of the DNUL model is the same for all experiments. For each classifier, we perform 10 optimizations. For each optimization, we initialize 2 random-weight vectors drawn from normally distributed random numbers $\sim \mathcal{N}(0, 1)$, one with \mathbf{w} and one with $-\mathbf{w}$. In total, we have 20 classifiers and we pick the classifier with the highest energy function. The number of iterations is set to 500 and the step size is set to $\alpha = 0.2$. The sensitivity parameter $\beta = 0.1$ was

chosen empirically and is used for the whole dataset. We are working on a mechanism to set this parameter automatically based on data.

Batch mode analysis: To start, we compare our approach with the supervised LDA classifier. The first two experiments are carried out by averaging the EEG dataset with a chunk of 15 trial groups for supervised or unsupervised learning and then the classifier is re-evaluated on the sequence of trial groups starting from 1 to 15. Both classifiers in Fig. (2a) are evaluated on the test dataset (I). As a rule, the LDA always learns with the training data and the DNUL learns on the fly with the test data. The curves display classification accuracy as a function of the number of trial groups involved in each data sample used to test the classifiers. In Fig. (2b) both classifiers are evaluated on the test dataset (II). Note that LDA and DNUL both learn on 14 characters for the experiment in Fig. (2a), whereas DNUL uses (of course unlabeled) data from all 26 characters for the case in Fig. (2b). The other two experiments in Fig. (2c) and (2d) follow the same methodology, but apart from this, they have a different configuration. In these two experiments, the number of trial groups the LDA and DNUL techniques use for supervised and unsupervised learning respectively matches the number

TABLE I
SPELLED WORDS IN TRAINING AND TEST DATASETS

Dataset	Spelled characters	Characters
Training dataset	KALEM_YOLCULUK	14
Test dataset (I)	KITAP_MASA_AGL	14
Test dataset (II)	KITAP_MASA_AGLAMAK_SIKINTI	26

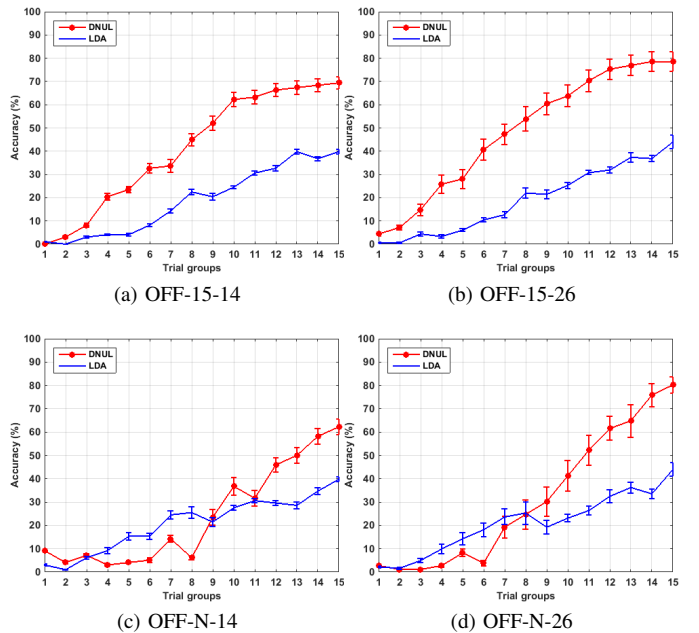


Fig. 2. Offline (Batch mode) analysis showing character classification accuracy over 7 subjects comparing DNUL with LDA. Error bars show 95% confidence intervals from the mean with sample size = 7.

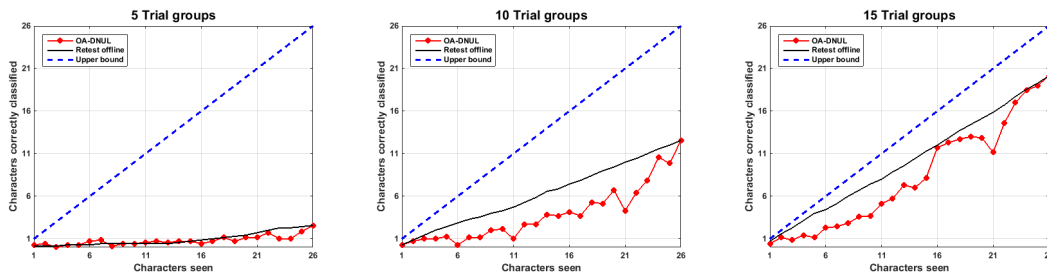


Fig. 3. Online spelling (sequential mode) showing the performance averaged over the 7 subjects using a different number of trial groups to predict a character. The horizontal axis represents the number of processed characters. The vertical axis represents the number of characters that were classified correctly. The dashed line is an upper bound showing the number of the seen characters.

of trial groups used for testing. These results demonstrate the unsupervised classification capability of DNUL. Interestingly, when the unlabeled data quality (through more repetitions) and quantity (through more characters) is sufficiently high, DNUL appears to provide better performance than supervised LDA trained on labeled data from a separate session. We speculate this might be due to the nonstationary nature of the EEG data across the sessions. The detailed accuracies for individual subjects corresponding to the experiments in Fig. 2 are shown in Table II.

Sequential mode analysis: This is a simulation to test the online adaptation process of DNUL. We design and update (adapt) the classifier after the data are received for each character and perform classification. We observe (see Fig. 3) that the classifier is improved as we receive more data. As expected, the classifier performs better if the data involve more trial groups. Finally, we also perform an offline "retest", that is we classify each previously seen character with the final classifier. This experiment demonstrates how DNUL can in principle be adapted and refined as more test data are received.

IV. CONCLUSION

In this paper, we have developed a novel unsupervised method for P300-based BCI speller systems, which allows us to run the classifier without using any calibration process and without any labeled data. Future work will include comparison with other unsupervised classification methods in BCI, such as [12]. It might also be possible to modify our energy function to incorporate additional terms to enforce various forms of clustering in the data.

TABLE II
PERFORMANCE VALUES FOR EACH SUBJECT OBTAINED WITH 15 TRIAL GROUPS FOR FIG. 2. (A) AND (D)

Subjects	Test dataset (I)		Test dataset (II)	
	LDA %	DNUL %	LDA %	DNUL %
S1	28.57	85.71	15.38	88.46
S2	28.57	42.85	34.62	69.23
S3	35.71	71.43	34.62	76.92
S4	35.71	92.86	61.54	88.46
S5	57.14	85.71	50	100
S6	42.86	21.43	50	42.31
S7	50	85.71	61.54	96.15
Average	39.8	69.39	43.96	80.22

REFERENCES

- [1] F. Nijboer, E. Sellers, J. Mellinger, M. Jordan, T. Matuz, A. Furdea, S. Halder, U. Mochty, D. Krusienski, T. Vaughan *et al.*, "A p300-based brain-computer interface for people with amyotrophic lateral sclerosis," *Clinical neurophysiology*, vol. 119, no. 8, pp. 1909–1916, 2008.
- [2] U. Hoffmann, J.-M. Vesin, T. Ebrahimi, and K. Diserens, "An efficient p300-based brain-computer interface for disabled subjects," *Journal of Neuroscience methods*, vol. 167, no. 1, pp. 115–125, 2008.
- [3] J.-J. Vidal, "Toward direct brain-computer communication," *Annual review of Biophysics and Bioengineering*, vol. 2, no. 1, pp. 157–180, 1973.
- [4] L. A. Farwell and E. Donchin, "Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials," *Electroencephalography and clinical Neurophysiology*, vol. 70, no. 6, pp. 510–523, 1988.
- [5] R. T. Lauer, P. H. Peckham, K. L. Kilgore, and W. J. Heetderks, "Applications of cortical signals to neuroprosthetic control: a critical review," *IEEE transactions on rehabilitation engineering: a publication of the IEEE Engineering in Medicine and Biology Society*, vol. 8, no. 2, pp. 205–208, 2000.
- [6] J. R. Wolpaw and D. J. McFarland, "Control of a two-dimensional movement signal by a noninvasive brain-computer interface in humans," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 101, no. 51, pp. 17 849–17 854, 2004.
- [7] T. W. Picton, "The p300 wave of the human event-related potential," *Journal of clinical neurophysiology*, vol. 9, no. 4, pp. 456–479, 1992.
- [8] U. Orhan, K. Hild, D. Erdogmus, B. Roark, B. Oken, and M. Fried-Oken, "Rsvp keyboard: An eeg based typing interface," in *Acoustics, Speech and Signal Processing (ICASSP), 2012 IEEE International Conference on*, March 2012, pp. 645–648.
- [9] J. R. Wolpaw and D. J. McFarland, "Multichannel eeg-based brain-computer communication," *Electroencephalography and clinical Neurophysiology*, vol. 90, no. 6, pp. 444–449, 1994.
- [10] M. Sajjadi, M. Seyedhosseini, and T. Tasdizen, "Disjunctive normal networks," *CoRR*, vol. abs/1412.8534, 2014.
- [11] S. Lu, C. Guan, and H. Zhang, "Unsupervised brain computer interface based on intersubject information and online adaptation," *Neural Systems and Rehabilitation Engineering, IEEE Transactions on*, vol. 17, no. 2, pp. 135–145, 2009.
- [12] P.-J. Kindermans, D. Verstraeten, and B. Schrauwen, "A bayesian model for exploiting application constraints to enable unsupervised training of a p300-based bci," *PLoS ONE*, vol. 7, no. 4, p. e33758, 04 2012.
- [13] D. J. McFarland, W. A. Sarnacki, and J. R. Wolpaw, "Should the parameters of a bci translation algorithm be continually adapted?" *Journal of neuroscience methods*, vol. 199, no. 1, pp. 103–107, 2011.
- [14] M. Hazewinkel, *Encyclopaedia of Mathematics: Volume 6*. Springer Science & Business Media, 2013.
- [15] C. Ulas and M. Cetin, "Incorporation of a language model into a brain computer interface based speller through hmms," in *Acoustics, Speech and Signal Processing (ICASSP), 2013 IEEE International Conference on*. IEEE, 2013, pp. 1138–1142.
- [16] A. Amcalar and M. Cetin, "Design, implementation and evaluation of a real-time p300-based brain-computer interface system," in *Pattern Recognition (ICPR), 2010 20th International Conference on*. IEEE, August 2010, pp. 117–120.