A Hierarchical BCI System Able to Discriminate between Non Intentional Control State and Four Intentional Control Activities

Julio Abascal, Andoni Arruti, José I. Martín and Javier Muguerza
Department of Computer Architecture and Technology, University of the Basque Country (UPV/EHU)
Manuel de Lardizabal 1, 20018 Donostia, Spain

Keywords: Brain-Computer Interface (BCI), Non Intentional Patterns Detection, Electroencephalogram (EEG), Clustering, Supervised Learning.

Abstract: This paper presents a two-level hierarchical approach to recognising intentional and non intentional mental tasks on a brain-computer interface. A clustering process is performed at the first recognition level in order to differentiate Non intentional Control state (NC) patterns from Intentional Control (IC) patterns. At the second level, the IC detected patterns are classified by means of supervised learning techniques, applied to the type of movement (left hand, right hand, tongue or foot imagery movement). The objective is to achieve high correct movement recognition scores, with a low percentage of wrong decisions (that is, low false positive rates), to avoid user frustration. Offline evaluation of the proposed prototype shows 84.5% accuracy, with a 6.7% false positive rate.

1 INTRODUCTION

Brain-Computer Interfaces (BCI) based on Electroencephalography (EEG) enable users to command computers just by measuring EEG signals associated with brain activity (Wolpaw et al., 2002). This kind of BCI requires a system to identify user brain activity patterns that are later translated into commands (Lotte et al., 2007).

Most BCI systems are based on synchronous protocols where the subject must follow a fixed repetitive scheme to switch from one mental task to the next (Pfurtscheller and Neuper, 2001) (Wolpaw et al., 2002). In synchronous BCI systems, the EEG phenomena to be recognized are time-locked to diverse cues. A trial typically lasts from 4 to 10 s or more. In contrast, in asynchronous BCI systems the subject makes voluntary, self-paced decisions on when to stop performing a mental task and when to start the next one (Nooh, Yunus and Daud, 2011). Designing an asynchronous BCI system requires continuous analysis of EEG signals. This analysis should determine whether the user is in an Intentional Control (IC) state, that is, if (s)he is producing one of the brain activity patterns used to control the BCI, or if (s)he is in a Non Control (NC) state. Finally, if the user is in an IC state, the system also has to determine which kind of brain activity pattern is being produced. Therefore, to deal with asynchronous problems, it is necessary to be able to differentiate between known and unknown activity patterns.

This paper presents the preliminary results of a study dealing with the problem of classifying patterns between the different types of IC states after an NC state discarding process. The approach has a two level hierarchical structure. The first level determines whether an activity pattern is present or not by applying a clustering process. The second level detects which of four mental tasks (left hand, right hand, tongue and foot imaginary movements) has been produced by the user. For this level a supervised classifier based on Support Vector Machine paradigm is proposed.

The remainder of the paper is organized as follows. Section 2 explains the experimental protocol used, the data acquisition process and the pre-processing carried out. Section 3 presents the proposed system to classify the EEG signals into the four types of imaginary movements considered, discarding the NC states. Results are presented and discussed in Sections 4 and 5. Finally, some conclusions and references are presented.

Abascal J., Arruti A., I. Martín J. and Muguerza J..
A Hierarchical BCI System Able to Discriminate between Non Intentional Control State and Four Intentional Control Activities.
DOI: 10.5220/0004723000910097
In *Proceedings of the International Conference on Physiological Computing Systems* (PhyCS-2014), pages 91-97
ISBN: 978-989-758-006-2

2 EXPERIMENTAL PROTOCOL

For our tests, we used the IIIa dataset from the BCI competition III (Blankertz et al., 2006). It contains data from 3 subjects: K3b, K6b and L1b, collected as follows (Schlögl et al, 2005) (see Figure 1). Each subject, sitting in front of a computer, was asked to perform imaginary movements of the left hand, right hand, tongue or foot during a specified time interval according to a cue. The order of cues was random. 60 electrodes were placed on the subject's scalp recording a signal sampled at 250 Hz and filtered between 1 and 50 Hz using a Notch filter. Each trial started with a blank screen. At t = 2s, a beep was generated and a cross "+" was shown to attract the subject's attention. At t = 3s an arrow pointing to the left, right, up or down was shown for 1s and the subject was asked to try one of four imaginary movements until the cross disappeared at t = 7s. This was followed by a 2s break, and then the next trial began. The dataset contains 360 instances (cases) for subject K3b, 240 for K6b and 240 for L1b. Each instance was labelled as belonging to one of the four classes. Each dataset contains a balanced distribution of the classes.

Comparing the subjects shows that K3b presents the highest accuracy and K6b the lowest (Lee et al., 2005). This is attributed to the different amount of BCI training received by the subjects. K3b was the most experienced, L1b had less experience and K6b was a beginner. In this work, we have used the data corresponding to subjects with extreme skills: K3b and K6b.

This dataset was designed for a synchronous BCI system, where the subject is aware of the time period to imagine the proposed activity. Therefore, this data was collected while the subjects were performing a cue-based (synchronized) task. A summary of the results obtained in this synchronous data logging exercise can be seen in (AlZoubi, Koprinska and Calvo, 2008).

In this paper, we extended this experimental data by adding data associated to NC states to the dataset. We associated the first three seconds (t = 0-3s) of each trial with no intentional activity. Therefore, for subject K3b we get 720 instances: 360 instances that indicate one of the four different intentional activities (IC state), corresponding to processing 4-7 seconds of each trial, and another 360 instances with non intentional brain activity (a new class: NC state), corresponding to the data obtained from 0-3 seconds of each trial. We processed the data from subject K6b in the same way, and we obtained 480 instances: 240 instances of intentional activity and 240 instances of non intentional activity.

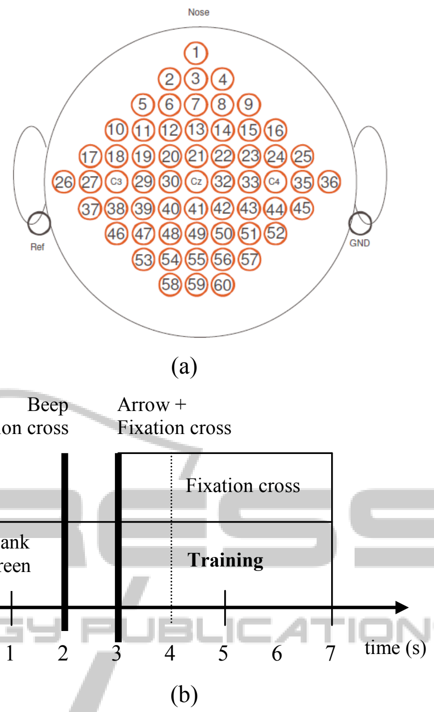

Figure 1: (a) EEG electrode placement, and (b) data acquisition paradigm (BCI Competition III).

We applied the same pre-processing techniques applied by AlZoubi et al. in (AlZoubi, Koprinska and Calvo, 2008). Firstly, we applied the Common Spatial Patterns (CSP) method (Müller-Gerking, Pfurtscheller and Flyvbjerg, 1999) to the raw EEG data. The result of applying CSP to the original 60 signals is a new set of 60 signals sorted by their ability to predict class. We selected the first 5 projections, and then we applied 3 frequency band filters (for 8-12 Hz, 21-20 Hz and 20-30 Hz). Finally, we extracted 7 features: max, min and mean voltage values, voltage range, number of samples above zero volts, zero voltage crossing rate and average signal power. This process gives 525 [5x5x3x7] (5 classes, 5 projections, 3 filters and 7 features) discrete numeric values for each case of the dataset.

After pre-processing the data corresponding to each subject, we split all the data into three sets: a training dataset for the clustering process at the first level (to detect the IC states), another training dataset for learning a supervised classifier at the second level (to distinguish between the four types of mental tasks) and a test dataset for evaluating the system's performance. Therefore, we created three datasets randomly for each subject, containing 240 cases for subject K3b (120 correspond to NC state and 120 to IC state, namely 30 cases for each imaginary movement), and 160 cases (80 for NC state and 80 for IC state: 20 cases for each imaginary movement) for subject Kb6.

One of the most important problems in machine learning is the need to deal with high numbers of dimensions. This problem is known as the curse of dimensionality: small numbers of training instances but highly dimensional. In these cases, it is necessary to simplify the learning algorithm by reducing the dimensionality before starting the learning process. This can be done by selecting the problem's most informative features and discarding the most irrelevant and redundant features. In this work we applied the Correlation-Based Feature Selection (CFS) method (Hall, 2000) which is the same feature selection method used by AlZoubi et al. in (AlZoubi, Koprinska and Calvo, 2008). This method bases its selection on searching for features that are highly correlated with a specific class variable and least correlated with the other variables. We used the implementation provided by Weka data-mining platform (Witten and Frank, 2005). We used the best first (greedy) search option starting with an empty set of features and adding new features. It is important to note that feature selection was only performed using training data, and test data was not used in any way during feature selection. As a result, 45 and 40 features were selected for K3b and K6b, respectively.

3 PROPOSED HIERARCHICAL SYSTEM: TRAINING PHASE

As mentioned in the Introduction, the system proposed in this work has a hierarchical structure that can be seen in Figure 2. The first level determines the presence or absence of intentional activity in the EEG signal, applying clustering techniques. The second level determines whether the detected intentional activity is a left hand, right hand, tongue or foot imaginary movement.

3.1 First Level: Detecting IC Activity

We used the K-means algorithm (Weka implementation) with the Euclidean distance for the first level of the system. One key issue in this phase is how good the system is at rejecting the NC state. Hence, to design this level we transformed the five-class dataset into a two-class dataset.

It is critical to minimize the False Positive Rate of the NC class (FPR = FP / (FP+TN)). FPR depends on False Positives (FP, acceptance of an NC state as an IC state) and on True Negatives (TN, rejection of a true NC state). A high FPR tends to cause excessive user frustration making the resulting BCI unusable (e.g. if the system is used to control a wheelchair, an FP would imply undesirable chair movements implying high risk for the user). According to the literature, we have selected a maximum threshold of 10% for the system's first level FPR. This value is similar to the FPR used in (Lotte, Mouchère and Lécuyer, 2008) (Scherer et al., 2008). After the clustering process, each cluster is labelled as belonging to one of the two classes, NC or IC, taking into account that the FPR (for the NC class) must be less than the selected threshold (10%). This has been done by establishing a minimum number of IC class patterns for the generated clusters (IC-threshold) to be labelled as an IC cluster.

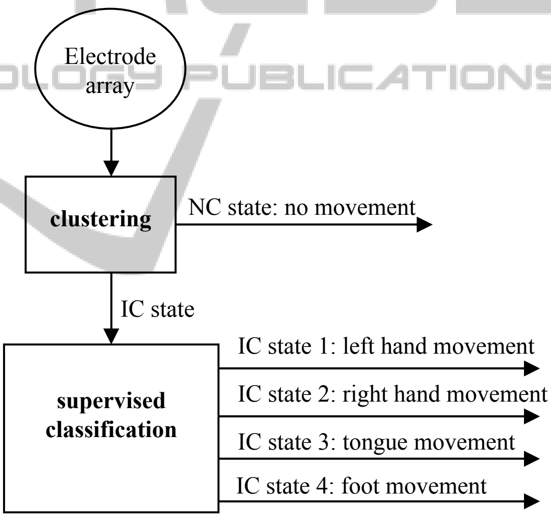

Figure 2: Structure of the proposed BCI system.

Therefore, for this first stage of the system, we had to select the IC-threshold and the K value for the K-means algorithm. We analysed 10 different values for the parameter K: 5, 10, 15, 20, 25, 30, 35, 40, 45 and 50. This estimation was made by applying a 10-fold cross-validation methodology using the first set of training data with patterns from the two-class system: NC state and IC state (grouping all patterns of each imaginary movement). Only clusters exceeding the IC-threshold were labelled as IC class. We calculated the nearest cluster for each unused pattern of the cross-validation fold using average linkage distance. As shown in Table 1, to maintain the FPR level under 10%, a 70% IC-threshold was necessary. The best results were obtained with K =

Table 1: FPR (NC class) and accuracy (10-CV) depending on the value of the K parameter and the IC-threshold. Grey shading for FPR under 10%, * shows the highest accuracy, and bold highlights the best option for each subject.

IC-threshold	10-CV		K value									
			5	10	15	20	25	30	35	40	45	50
90%	K3b	FPR (%)	0.0	5.8	6.7	4.2	6.7	4.2	5	5.8	4.2	5.8
		Accuracy (%)	51.7	62.1	62.9	65.8	65	70	70.8	67.1	72.1*	70.8
	K6b	FPR (%)	0.0	6.3	5.0	5.0	8.8	7.5	10.0	13.8	13.8	13.8
		Accuracy (%)	50.0	56.9	63.8	70.0*	64.4	61.3	59.4	63.8	66.3	66.9
80%	K3b	FPR (%)	10.8	12.5	10.0	8.3	10	8.3	8.3	8.3	6.7	10
		Accuracy (%)	74.6	75.0	74.6	77.9	72.9	80.4	79.2	79.2	81.3*	82.9
	K6b	FPR (%)	7.5	11.3	6.3	6.3	10.0	12.5	12.5	18.8	16.3	16.3
		Accuracy (%)	67.5	68.8	73.8*	73.8*	72.5	70.0	68.1	73.1	73.1	73.8
70%	K3b	FPR (%)	10.8	12.5	10.8	11.7	15.8	**8.3**	10.8	13.3	12.5	13.3
		Accuracy (%)	75.4	79.6	75.4	80.0	76.3	**83.8**	80.4	79.2	80.4	81.3
	K6b	FPR (%)	15.0	16.3	**7.5**	11.3	16.3	13.8	13.8	20.0	16.3	16.3
		Accuracy (%)	78.1	76.3	**77.5**	76.3	74.4	76.9	70.0	73.8	731	73.8
60%	K3b	FPR (%)	15.8	19.2	10.8	14.2	19.2	15.8	18.3	20.0	15.8	19.2
		Accuracy (%)	75.8	79.6	76.7	82.5	79.2	83.3	80.0	78.8	81.7	82.1
	K6b	FPR (%)	18.8	17.5	10.0	13.8	20.0	18.8	22.5	20.0	20.0	20.0
		Accuracy (%)	81.9	79.4	80.0	78.1	75.0	80.0	75.0	75.6	76.3	74.4

Table 2: Accuracy (10-CV) of the classifiers.

10-CV Accuracy (%)	1R	DT	1-NN	5-NN	NB	RBF	SVM	LR	AdaB	Bag	RF
K3b	45.8	60.8	77.5	82.5	80.8	77.5	**84.2**	74.2	72.5	74.2	77.5
K6b	31.3	58.8	53.8	58.8	60.0	63.8	**65.0**	53.8	62.5	56.3	56.3

30 for subject Kb3, and K = 15 for subject K6b. For these K values, the FPR (NC class) was 8.3% and 7.5%, respectively.

3.2 Second Level: Classifying the Type of Imaginary Movement

We used supervised learning algorithms to implement this level. From the great variety of algorithms that have been applied in BCI systems (Lotte et al., 2007), we selected 11 algorithms: 1R rule, Decision Tree (DT), k-NN (1-NN and 5-NN), Naive Bayes (NB), Radial-bases Network (RBF), Support Vector Machine (SVM), Logistic Regression (LR), Ada Boost (AdaB, combining 10 decision trees), Bagging (Bag, combining 10 decision trees) and Random Forest (RF). We chose these algorithms because they represent different paradigms (rule-based, tree-based, distance-based, probabilistic, function-based, ensemble of classifiers) and they are state of the art in data mining. We use their Weka (Witten and Frank, 2005) implementation by applying the default values for the parameters.

All classifiers were trained using the second training set (120 cases, 30 cases of each imagery mental task). The best algorithm was estimated by applying a 10-fold cross-validation methodology. Table 2 shows the accuracy achieved by each classifier for both subjects. The best overall classifier was the SVM algorithm with an accuracy of 84.2% for subject K3b and 65.0% for subject K6b.

Summarizing, the proposed hierarchical BCI system consists of a first level that differentiates between NC and IC states, based on clustering techniques; and a second level, based on a SVM classifier, that discriminates between the four types of mental tasks considered. The optimal number of first level clusters is different for each subject: 30 for subject K3b and 15 for subject K6b.

4 EXPERIMENTAL RESULTS: EXPLOITATION PHASE

Once the system has been trained, it can be used for classifying new EEG patterns, so far unknown to the system. The newly-designed system's performance was tested using the test dataset. As previously

Table 3: Classification performance for the first level using the test dataset.

	Confusion Matrix			FPR (%)	Accuracy (%)
K3b (K-means, K=30)		IC-estimated	NC-estimated	6.7	81.7
	IC-real	84	36		
	NC-real	8	112		
K6b (K-means, K=15)		IC-estimated	NC-estimated	2.5	80.0
	IC-real	50	30		
	NC-real	2	78		

Table 4: Classification performance for the second level with the test dataset.

	Confusion Matrix					Accuracy (%)
K3b (SVM)		Left-estimated	Right-estimated	Tongue-estimated	Foot-estimated	84.5
	Left-real	6	7	0	2	
	Right-real	2	26	0	0	
	Tongue-real	1	0	18	1	
	Foot-real	0	0	0	21	
K6b (SVM)		Left-estimated	Right-estimated	Tongue-estimated	Foot-estimated	64.0
	Left-real	2	9	0	0	
	Right-real	1	11	1	0	
	Tongue-real	3	3	1	0	
	Foot-real	0	0	1	18	

explained, the test set for subject K3b comprises 240 cases (120 corresponding to NC state and 120 cases for IC state), while the test set for subject K6b consists of 160 cases (80 of each state). Considering only the IC patterns, subject K3b's test set consists of 30 cases of each of the 4 types of mental tasks, whereas, subject K6b's test set consists of 20 cases of each type of movement.

Table 3 summarizes the results obtained for the first level of the system when the new patterns were processed. This table shows the confusion matrix obtained, as well as the FPR (NC class) and the accuracy of this first level. There are 8 cases of NC patterns misclassified into clusters labelled with some kind of movement (i.e. intentional activity) leading to 6.7% FPR (NC class) for subject K3b. 84 cases (from the initial 120) corresponding to activity patterns (IC state) will be classified in the second level of the system. On the other hand, they are only 2 cases of misclassified NC patterns; yielding a 2.5% FPR for subject K6b, 50 cases (from the initial 80) correspond to activity patterns.

Table 4 shows the confusion matrix for the second level of the system. The patterns are classified in the 4 possible imaginary movements (right, left, tongue, foot) using a Support Vector Machine (SVM) classifier. The accuracies obtained from the two subjects are 84.5% and 64.0%, respectively.

Analysing the confusion matrix shows that it is more difficult for both subjects to detect the left hand movement. The number of patterns reaching the second level of the system for this movement is clearly lower than for all other movements, and, there is a bias in the system that classifies these patterns as belonging to the right imaginary movement. In general, subject K6b obtained worse results and presented specific difficulties with the tongue movement (only 7 patterns reached the second level, and, only one was correctly classified).

Analysing the overall system's performance in terms of classifying the five different patterns (NC class + 4 imaginary movements), the accuracy for subject K3b is 76.3% and 68.8% for the other subject. The differences in results obtained for each subject confirmed the description of subject K6b as a less trained (beginner) user, and, as a consequence, the system had greater difficulty dealing with this subject's EEG patterns.

Although it is difficult to compare the results obtained with other works, mainly because we have included a class for the Non Intentional Control state, we can say that the results obtained are similar to the work presented by AlZoubi et al. in (AlZoubi, Koprinska and Calvo, 2008). They obtained a 78.5% average result for both subjects, whereas in our case the accuracy was 72.6% (taking into account the difficulty of introducing the NC class).

Table 5: Accuracy (10-CV) of the classifiers (one-level system).

10-CV Accuracy (%)	1R	DT	1-NN	5-NN	NB	RBF	SVM	LR	AdaB	Bag	RF
K3b	40.4	63.9	75.8	79.6	75.7	77.1	**81.1**	69.9	76.0	72.5	73.3
K6b	37.0	50.0	44.0	46.5	52.0	50.5	**60.0**	47.5	56.0	57.0	56.0

5 COMPARISION WITH AN ONE-LEVEL SYSTEM

In order to test the validity of the proposed hierarchical system, we also implemented a one level system without the clustering phase. We analysed the performance of the same classifiers used in Section 2 to develop the second level classifier for the five classes (NC, left hand, right hand, tongue or foot). All the classifiers were now trained using the two previously defined training data sets: 300 cases (60 cases of each class) for subject K3b and 200 cases (40 cases of each class) for subject K6b. The best algorithm was estimated again by applying a 10-fold cross-validation methodology. Table 5 shows the accuracy achieved by each classifier for both subjects. The best overall classifier was again the SVM algorithm, with an accuracy of 81.1% for subject K3b and 60% for subject K6b.

The system performance was tested with the same test data set used for the hierarchical proposal: 240 cases for subject K3b and 160 cases for subject K6b. The obtained accuracy using the SVM classifier was lower than the one obtained with the two-level system: 74.5% for subject K3b and 66.9% for subject K6b. These results confirm that the two-level approach has higher performance than the one-level system.

6 CONCLUSIONS

In this paper we proposed a two-level hierarchical approach to recognise mental tasks including intentional and non intentional states on a brain-computer interface. At the first level, the proposal performs a clustering process in order to differentiate patterns of Non intentional Control state (NC) from patterns of Intentional Control (IC). At the second level, the IC detected patterns are classified by movement type (left hand, right hand, tongue or foot imaginary movement) by a supervised learning classifier.

After a pre-processing phase and reducing the number of dimensions of the problem, we applied the K-means algorithm for the first level of the system, obtaining the best results with K = 30 (accuracy of 81.7%) and K = 15 (accuracy of 80.0%) for subjects K3b and K6b, respectively, using BCI III Competition dataset IIIa. The best results obtained for the second level were achieved with the Support Vector Machine classifier with 84.5% and 64.0% overall accuracy, respectively. These results were obtained maintaining the False Positive Rate for the NC class under 10% (achieving 6.7% and 2.5% rates for the subjects participating in the experiment). The classification phase results encourage us to apply Support Vector Machine based algorithms in the clustering phase.

This work takes advantage of the good results obtained in synchronous experiments to apply them in a more realistic but more demanding asynchronous environment. In the asynchronous case, the data includes inactivity periods along with activity states. Our work firstly distinguishes activity from non-activity status. Once voluntary activity is found, we detect the type of virtual movement associated with it. In this way, our research proposes a step forward towards practical asynchronous detection.

Nevertheless, this work did not deal with the problem of detecting IC states in a continuous EEG signal (asynchronous or self-paced BCI). We just introduced the NC state as a new class. This may explain why our results are generally better than presented in other works (Scherer et al., 2008) (Satti, Coyle and Prasad, 2009). We plan to apply our proposal to the asynchronous BCI problem in order to be able to compare both approaches.

ACKNOWLEDGEMENTS

This work was funded by the University of the Basque Country (Aldapa, GIU10/02), by the Science and Education Department of the Spanish Government (ModelAccess project, TIN2010-15549), and by the Department of Education, Universities and Research of the Basque Government (IT395-10 research group grant).

REFERENCES

Wolpaw, J., Birbaumer, N., McFarland, D., Pfurtscheller, G., Vaughan T., 2002. Brain-computer interfaces for communications and control. *Clinical Neurophysiology*, 113 (6): 767-791.

Lotte, F., Confgedo, M., Lécuyer, A., Lamarche, F., Arnaldi, B., 2007. A review of classification algorithms for eeg-based brain-computer interfaces. *Journal of Neural Engineering*, 4:R1-R13.

Pfurtscheller, G., Neuper, C., 2001. Motor imagery and direct brain-computer communication. *Proc. IEEE*, vol. 89, pp.1123-1134.

Nooh, A. A., Yunus, J., Daud, S.M., 2011. A review of Asynchronous Electroencephalogram-based Brain Computer Interface Systems. *International Conference on Biomedical Engineering and Technology*, Singapore, vol. 11.

Blankertz, B., Müller, K. R., Krusienski, D., Schalk, G., Wolpaw, J. R., Schlögl, A., Pfurtscheller, G., Millan, J., Schröder, M., Birbaumer, N., 2006. The BCI2000 Competition III: Validating Alternative Approaches to Actual BCI Problems. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 14, no. 2, pp. 153-159.

Schlögl, A., Lee, F., Bischof, H., Pfurtscheller, G., 2005. Characterization of four-class motor imagery EEG data ffor the BCI competition 2005. *Journal of Neural Engineering*, 2: L14-L22.

Lee, F., Sherer, R., Leeb, R., Neuper, C., Bischof, H., Pfurtscheller, G., 2005. A Comparative Analysis of Multi-class EEG Classification for Brain-computer interface. *10th Computer Vision Winter Workshop (CVWW)*. Technical University of Graz, Austria.

AlZoubi, O., Koprinska, I., Calvo, R.A., 2008. Classification of Brain-computer Interface Data. *7th Australasian Data Mining Conference (AusDM)*, Adelaide (Australia), pp. 123-132.

Müller-Gerking, J., Pfurtscheller, G., Flyvbjerg H., 1999. Designing optimal spatial filters for single-trial EEG classification in a movement task. *Clinical Neurophysiology*, vol. 110, no. 5, pp. 787-798.

Hall, M., 2000. Correlation-based feature selection for discrete and numeric class machine learning. *17th International Conference on Machine Learning (ICML)*, 359-366, Morgan Kaufmann.

Witten, L. H., Frank, E., 2005. *Data Mining: Practical Machine Learning Tools and Techniques*. Morgan Kaufmann, San Francisco.

Lotte, F., Mouchère, H., Lécuyer, A., 2008. Pattern Rejection Strategies for the Design of Self-Paced EEG-based Brain-Computer Interfaces. *19th International Conference on Pattern Recognition*, Florida (USA), pp. 1-5.

Scherer, R., Lee, F., Schlögl, A., Leeb, R., Bischof, H., Pfurtscheller, G., 2008. Towards self-paced brain-computer communication: Navigation through virtual worlds. *IEEE Transactions on Biomedical Engineering*, vol. 55, no. 2, pp. 675-682.

Satti, A., Coyle, D., Prasad, G., 2009. Continuous EEG Classification for a Self-paced BCI. *4th International IEEE EMBS Conference on Neural Engineering*, Turkey, pp. 315-318.