Comparison of Two Different Classifiers for Mental Tasks-Based Brain-Computer Interface: MLP Neural Networks vs. Fuzzy Logic

Giovanni Saggio[1#*], Pietro Cavallo[#], Alessio Ferretti[#], Francesco Garzoli[#], Lucia Rita Quitadamo[§&], Maria Grazia Marciani[§], Franco Giannini[#*], Luigi Bianchi[§*&]

[#]*Dept. of Electronic Engineering, "Tor Vergata" University, via del Politecnico 1, 00133 Rome, Italy*

[§] *Dept. of Neuroscience, "Tor Vergata" University, via Montpellier 1, 00133 Rome, Italy*

[*]*Centro di Biomedicina Spaziale, "Tor Vergata" University, via del Politecnico 1, 00133 Rome, Italy*

[&] *Dept. of Neurophysiology, Fondazione S. Lucia, IRCCS, via Ardeatina 306, 00179 Rome, Italy*

[1]*saggio@uniroma2.it*

Abstract

This study is devoted to the classification of four-class mental tasks data for a Brain-Computer Interface protocol. In such view we adopted Multi Layer Perceptron Neural Network (MLP) and Fuzzy C-means analysis for classifying: left and right hand movement imagination, mental subtraction operation and mental recitation of a nursery rhyme.

Five subjects participated to the experiment in two sessions recorded in distinct days. Different parameters were considered for the evaluation of the performances of the two classifiers: accuracy, that is, percentage of correct classifications, training time and size of the training dataset. The results show that even if the accuracies of the two classifiers are quite similar, the MLP classifier needs a smaller training set to reach them with respect to the Fuzzy one. This leads to the preference of MLP for the classification of mental tasks in Brain Computer Interface protocols.

1. Introduction

Brain Computer Interface (BCI) systems have been developed as assistive devices for disabled people that have lost the control on their body after dramatic events, such as strokes, spinal traumas, degenerative diseases, etc. [1].

These people can no more communicate towards the external world and so lie into an isolation condition.

BCI systems try to furnish these people an alternative communication channel, by translating some of their brain signals into commands for piloting an external device that can be a wheelchair, a robotic arm, a Web surfer, a cursor on a screen, etc. This is obtained in the following way: the brain signals are acquired and then processed to extract some features of interest from them; these features are then classified and encoded into semantic symbols that are finally mapped into the output commands.

Some BCI systems can be driven by means of mental tasks [2],[3], in the sense that the user of the system mentally imagines to perform some particular tasks that are then recognized by a classifier and used to pilot the output peripheral. For example, as in the case of the experimental protocol described in this paper, the subject is asked to imagine right and left hand movements, to perform mental calculation and to mentally recite a nursery rhyme. These mental tasks, after the classification, can be associated, for example, to the 2D movement of a cursor on a screen thus implementing the selection of particular targets and so the communication.

Two different classifiers are here presented for the classification of four-class mental tasks data: Multi Layer Perceptron Neural Network and Fuzzy C-means analysis.

Since the 4 chosen tasks are uncommon in literature, it is not certain yet which classifier is better. So we chose MLP and Fuzzy Logic because the former is already consolidated in BCI, and the latter, despite it is a fairly new methodology in this field, performs a pretty well spatial separation.

Our purpose was to compare these methods and to find which better satisfies the requirements: accuracy, training time and size of the training dataset.

2. Methods

978-1-4244-4439-7/09/$25.00 ©2009 IEEE

This work was partially supported by the DCMC Project of the Italian Space Agency. This paper only reflects the authors' views and funding agencies are not liable for any use that may be made of the information contained herein.

2.1 The experimental protocol

Five subjects participated to the experiments. The EEG was recorded using 61 Ag-AgCl scalp electrodes located according to the International 10-20 system. The signals were 256 Hz sampled and bandpass-filtered between 0.5 Hz and 128 Hz. The sensitivity of the amplifier was set to 4 mV.

The experimental protocol consisted of four different imagery tasks, left and right hand movement imagination, mental subtraction operation and mental recitation of a nursery rhyme. Two sessions on distinct days were recorded for each subject. Each session consists of 200 trials (50 for each of the four possible tasks).

The subjects sat in a comfortable armchair in front of a computer screen. For every trial a text, indicating the task to perform, appeared on the black screen for 3 secs. The inter trial interval (ITI) was set to 1 sec.

Based on our past experiences, to reduce the computational workload, only 10 among the 61 electrodes (the green highlighted electrodes in fig. 1) were considered for the analysis and frequencies in the range of 8Hz and 13Hz (corresponding to α band [4]) were selected for every electrode. For a single trial, the relative powers of the signals in the above-said band, for each electrode, were computed, and constituted the features vector fed to the two different classifiers.

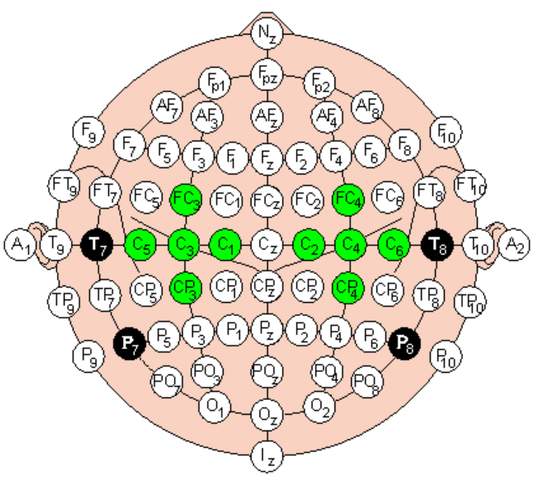

Fig. 1. The 10-20 standard with the ten considered electrodes highlighted in green.

2.2 Artificial Neural Networks

An Artificial Neural Network (ANN) is a computational model inspired by the way biological nervous systems, such as the brain, process information. It is composed of a large number of interconnected processing elements (artificial neurons) working to solve specific problems. An ANN is configured for a specific application, such as pattern recognition or data classification, through a learning process. As well as biological systems, learning involves adjustments to the connections that exist between the artificial neurones.

We used a Multi-Layer Perceptron (MLP) Neural Network made of four layers: one input, two hidden and one output. Each neuron is connected with a certain weight to every other neuron in the previous layer. At each time step, the input is propagated through layers. The input layer has 10 neurons, one for each considered electrode. For each trial input neurons receive the relative powers of electrodes normalized to an average of 0.5 value to make measures comparable between each other. At this point information is fed to the first hidden layer through weighted connections. Each hidden layer has 20 neurons. Excepted for the input layer all the neurons have a sigmoid activation function $fs(x)=\frac{1}{1+e^{-6(x-0.5)}}$ scaled in the range of 0 to 1.

Sigmoid was preferred due to its independent and fundamental space division properties [5],[6] as it models the frequency of action potentials of biological neurons in the brain.

The output layer has 4 neurons, one for each mental task to be recognized. In case of a successful classification the output of the neuron corresponding to the classified task tends to 1 whereas other outputs tend to 0.

Every neuron, except for the input layer, was initialized with a random weight in the range of $-1/\sqrt{n}$ to $+1/\sqrt{n}$, where n is the number of neurons connected by means of that weight [7],[8]. As commonly done it was assigned to the input layer a constant weight of 1.

After the output presentation a learning rule was applied. We used a supervised learning method [9] called backpropagation [10],[11]. The backpropagation calculates the mean-squared error between actual and expected output. The error value is then propagated backwards through the network, and small changes are made to the weights in each layer. The weight changes are calculated in order to reduce the error signal. The whole process was then repeated for each trial, and the cycle was reiterated until the overall error value drops below some pre-determined threshold.

2.3 Fuzzy Logic

Fuzzy logic arises as a method to formalize real-world concepts that cannot be categorically identified as true or false, but that may have some degree of truth. The fuzzy logic has particularly effectiveness in applications of information extraction and interpretation.

One of the hallmarks of fuzzy logic is that it allows nonlinear input/output relationships to be expressed by a set of qualitative "if-then" rules. Fuzzy rules provide a powerful framework for capturing and explaining the input/output data behavior.

Extracting fuzzy rules for pattern classification can be viewed as the problem of partitioning the input space into appropriate fuzzy clusters: groups of trials with similar structural characteristics [12].

This is made by applying the algorithm Fuzzy C-Means (FCM) on each n-dimensional vector of trials containing the relative powers of each electrode considered. The FCM is a clustering algorithm based on optimizing an objective function [13]. Given a set of elements $X = \{x_1, \ldots, x_n\} \subset \Re^p$, the aim of fuzzy clustering is to determine the prototypes in the way that the objective function $J(X,U,v) = \sum_{i=1}^{c} \sum_{k=1}^{n} (u_{ik})^m d^2(x_k, v_i)$ is minimized [13],[14],[15], where $u_{ik} \in [0,1]$ stands for the membership degree of x_k to the cluster *i*, and $d(x_k, v_i)$ is the Euclidean distance between x_k and the cluster *i*, represented by the so called prototype v_i; The apices c is the number of clusters. The choice of the value of the parameter *c* varies from case to case. For example, through many tests, the best classification for right hand thought was obtained with c equal to 8, whereas for left hand thought was obtained with 6. No theoretical foundations are yet available for the optimal choice of the parameter of the exponent m which was empirically set to 2.7 [12], [16].

This algorithm was applied for each mental task, such that they are represented by a set of clusters.

FCM is an iterative process in which each cluster $c_i (i = 1,2,\ldots,C)$ is regarded as a fuzzy set.

To deduce the Fuzzy rules from clusters, it is necessary to write a membership functions for each of them. We decided to use a triangular membership function as a best choice for adequately represents the clusters. This was made by projecting on each i-th axis, the i-th coordinate of the prototype and the two data points (trials) that are most distant from the prototype. We assigned the minimum membership function value (0) to the projected trials and the maximum value of 1 to the center of the cluster (prototype). In this way a Mamdani type fuzzy controller was implemented [17],[18].

3. Results And Discussion

In Table I, the accuracies achieved by the neural network applied to the 4 tasks are shown: the average percentage of correct classifications is around **78%**.

TABLE I
Multilayer Perceptron Classification

Subject	Percentage of correct classifications
E. T.	81%
P. C.	79.6%
F. G.	78%
D. P.	76.8%
F. C.	76.7%
Average percentage	78.4%

Percentage of correct classifications of the four mental tasks, using five examples per task to train the Neural Network.

We had better results by applying this method to the classification of the two motor tasks (left and right hand movement imagination) and of one motor towards one non-motor tasks (e.g. right hand and subtraction). In those cases, in fact, we achieved an average accuracy of 90%.

Also we empirically found that the best compromise between speed and correct classifications percentage is given by considering **5 trials** per task in the training phase.

The results obtained with the fuzzy logic in classifying each task are on an average of **78%** of correct classifications, with a peak of 82% (table II).

TABLE II
Fuzzy Logic Classification

Subject	PERCENTAGE OF CORRECT CLASSIFICATIONS
E. T.	82%
P. C.	79.1%
F. G.	77%
D. P.	76.2%
F. C.	77.1%
Average percentage	78,2%

Percentage of correct classifications of the four mental tasks, with the Fuzzy Logic analysis, using forty examples per task to make clusters.

The FCM algorithm has the advantage of obtaining a good compromise between the effectiveness of aggregation and the computational cost. In fact, each task required an average of 20 iterations to reach the optimum. Unfortunately this algorithm needs **40 training trials** in the clustering step to get a high percentage of classification.

It is to be noticed that, in literature, the considered tasks are not easy to classify; so the obtained accuracies are fairly good.

Both classifiers achieve the same mean percentage of correct classifications (fig. 2). However if we limited our analysis to this result it would seem that these classifiers can be used indifferently.

But it was fundamental to evaluate a focal parameter: the number of examples (in fig. 2 represented on the x axis) needed to reach the highest percentage.

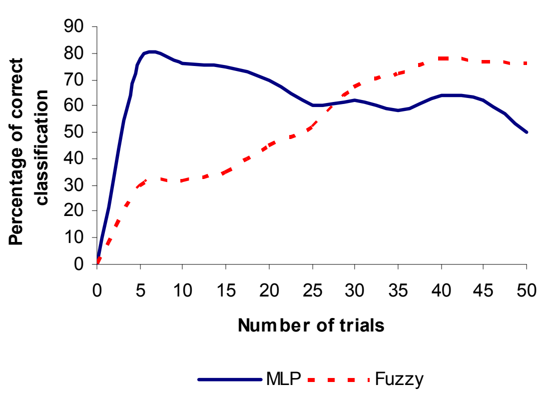

Fig. 2. Percentage of correct classifications as a function of the number of trials in the training set.

In fact, MLP uses a training set constituted by five trials per task respect to the Fuzzy classifier. As previously mentioned, it takes 4 seconds to record a trial, being spent 3 secs for performing the mental task and 1 sec for the ITI. This leads to a recording session of 80 secs for training the MLP (5 trials x 4 tasks x 4 secs) and 640 secs for the Fuzzy clustering (40 trials x 4 tasks x 4 secs) and so **a reduction of 8 times in the training of the former**.

This is critical because the training phase should be performed every time the patient that uses the BCI-system changed (BCI machines are ad-personam systems), and also if the system is reused by a different patient (in the replacement of the helmet the electrodes position can change).

Thus the training phase is done an enormous number of times and therefore it is essential for this stage to be as fast as possible.

In conclusion, it can be deduced that, as it needs a small training set, MLP is preferable with respect to the FCM algorithm, for the classification of mental tasks in BCI protocols.

4. Conclusion

Here it is presented a comparison among two different classifiers for the classifications of four-class mental tasks for BCI protocols: Multi Layer Perceptron Neural Network and Fuzzy C-means analysis. The results of the classifications show that both classifiers achieved the same mean accuracy; however it is evident that the MLP neural network needs a reduced number of trials for training purposes, having the advantage in the reduction of the recording session up to 8 times lower than Fuzzy analysis. This is of fundamental importance because, as we already said, training is a critical step since it is done a huge number of times in a BCI-system life. And so, the less a classifier is time-demanding in the training phase, the more the patient is comfortable with it, being part of the communication load dramatically reduced.

5. References

[1] J. R. Wolpaw, N. Birbaumer. D. J McFarland, G. Pfurtscheller and T. M. Vaughan "Brain-computer interfaces for communication and control," Clinical Neurophysiology, vol. 113, no. 6, 2002, pp. 767-791.

[2] A. Schogl, F. Lee, H. Bischof and G. Pfurtscheller, "Characterization of four-class motor imagery EEG data for the BCI-competition 2005", Journal of Neural Engineering, 2005, vol. 2, L14-L22.

[3] N. J. Huan and R. Palaniappan, "Neural network classification of autoregressive features from electroencephalogram signals for brain–computer interface design", Journal of Neural Engineering, 2004, vol. 1, 142-150.

[4] E. Kandel, J. Schwartz, and T. Jessell, "Principles of Neural Science" USA: McGraw Hill, 2000.

[5] S. Cammarata, “Reti neuronali, Dal perceptron alle reti caotiche e neuro-fuzzy”, Etas 1997

[6] K. Hara and K. Nakayamma, “Comparison of activation functions in multilayer neural network for pattern classification” IEEE World Congress on Computational Intelligence., 1994, vol. 5, pp. 2997-3002.

[7] C. Hernàndez-Espinosa and M. Fernandez Redondo, “Multilayer Feedforward Weight Initialization” European Symposium of Artificial Neural Networks 2001, pp. 119-124.

[8] H. Lari-Najafi, M. Nasiruddin and T. Samad “Effect of initial weights on back-propagation and its variations” 1989, vol.1, pp. 218-219

[9] L. G. Allred and G. E. Kelly, “Supervised learning techniques for backpropagation networks” IJCNN International Joint Conference on Neural Networks, 1990, vol. 1, pp. 721-728

[10] R. Hecht-Nielsen “Theory of the backpropagation neural network” IJCNN International Joint Conference on Neural Networks, 1989, vol.1, pp. 593-605

[11] C. Y. Chen and C. J. Hwang “A multi-level backpropagation network for pattern recognition systems”

[12] L. Mikhailov, A. Lekova, F. Fischer and H. A. Nour Eldin, “Method for fuzzy rules extraction from numerical data”, IEEE International Symposium on Intelligent Control, 1997.

[13] J. Abonyi, R. Babuska and F. Szeifert, “Modified Gath-Geva Fuzzy Clustering for Identification of Takagi-Sugeno Fuzzy Models”, IEEE transactions on systems, man, and cybernetics, 2002.

[14] M. Menard, “Extension of the objective functions in fuzzy clustering”, Fuzzy Systems, 2002.

[15] J.C. Bezdek, “Pattern Recognition with Fuzzy Objective Function Algorithms”, Plenum Press, 1981.

[16] L.B. Romdhane, B. Ayeb and S. Wang, “An improved scheme for the fuzzifier in fuzzy clustering”, Neural Networks for Signal Processing Proceedings of the 1997 IEEE Workshop, 1997, pp. 336-344.

[17] M. Sugeno and T. Yasukawa, “A Fuzzy Logic Based Approach to Qualitative Modeling” , IEEE Transactions on Fuzzy System, 1993, pp. 7-31.

[18] K. W. Wong, D. Tikk, T.D. Gedeon and L.T. Koczy, “Fuzzy rule interpolation for multidimensional input spaces with applications: a case study”, IEEE transactions on fuzzy systems, 2005, vol. 13, no. 6.