

Studying the Use of Fuzzy Inference Systems for Motor Imagery Classification

Lotte Fabien, Lécuyer Anatole, Lamarche Fabrice and Arnaldi Bruno

Abstract—This paper studies the use of Fuzzy Inference Systems (FISs) for motor imagery classification in EEG-based Brain-Computer Interfaces (BCIs). The results of the four studies achieved are promising as, on the analysed data, the used FIS was efficient, interpretable, showed good capabilities of rejecting outliers and offered the possibility of using *a priori* knowledge.

Index Terms—Brain-Computer Interface, classification, Fuzzy Inference System, ElectroEncephaloGraphy, motor imagery

I. INTRODUCTION

A Brain-Computer Interface (BCI) is a new communication system that triggers an increasing interest within the scientific community [1]. BCIs enable people to interact with computers by producing different brain activity “patterns”. In order to obtain a usable and efficient BCI, these patterns, usually measured using ElectroEncephaloGraphy (EEG), must be identified automatically and accurately. To do so, most BCI systems rely on a classifier [2]. However, some researchers regret that the classifiers currently used for BCIs cannot be easily interpreted and that adding *a priori* knowledge to them remains a difficult task [3]. They suggest that alternative classification algorithms need to be explored [3].

Fuzzy Inference Systems (FISs) are fuzzy classifiers that can learn fuzzy “if-then” rules able to classify data [4]. They exhibit several interesting properties that may address the limitations aforementioned. First, they are known to be readable, which means we can interpret what they automatically learnt [4]. Then, they are extensible, i.e., *a priori* knowledge can be easily added to them under the form of “hand-made” fuzzy rules [4]. Last, FISs are universal approximators [5] suitable for the classification of non-stationary biomedical signals [6] [7], which can be interesting for EEG classification.

Therefore, in this paper, we study the use of an FIS for the classification of popular EEG signals: motor imagery. The following sections describe the FIS algorithm and the data used as well as the related studies achieved.

II. THE FIS OF CHIU

Among available FISs [5] [8], we chose the Chiu’s FIS (CFIS) [4]. Indeed, CFIS is robust to noise, which is fundamental when dealing with such noisy data as EEG signals. Moreover, according to Chiu, the CFIS is generally more efficient than neural networks. Finally, it is a clustering-based

FIS, making it suitable for dealing with small training sets [5]. With the CFIS, fuzzy “if-then” rules can be automatically extracted from data in three successive steps:

- a) *Clustering of training data.* A clustering algorithm known as “subtractive clustering” [4] is applied to the training data of each class separately. This algorithm enables to find automatically the number of clusters and their positions.
- b) *Generation of the fuzzy rules.* A fuzzy “if-then” rule is generated for each cluster found previously. For a cluster j , belonging to class Cl_i , the generated fuzzy rule is:

if X_1 is A_{j1} and ... and X_N is A_{jN} then class is Cl_i

 where N is the dimensionality of the data, X_k is the k^{th} element of a feature vector X and A_{jk} is a Gaussian fuzzy membership function:

$$A_{jk}(X_k) = e^{-\frac{1}{2}(\frac{X_k - x_{jk}}{\sigma_{jk}})^2} \quad (1)$$

where x_{jk} is the k^{th} element of the vector representing the center of the cluster, and σ_{jk} is a positive constant which is initially the same for all A_{jk} .

To increase accuracy, the membership functions can be “two-sided” Gaussians with a plateau and a different standard deviation on each side (Fig. 1) [4].

- c) *Fuzzy rule optimization.* Each membership function A_{jk} is tuned according to gradient descent formulas that use a classification error measure E and a learning rate λ [4]:

$$x_{jk} \Leftarrow x_{jk} - \lambda \frac{\partial E}{\partial x_{jk}} \quad \text{and} \quad \sigma_{jk} \Leftarrow \sigma_{jk} - \lambda \frac{\partial E}{\partial \sigma_{jk}} \quad (2)$$

Once trained, the CFIS can use its set of fuzzy rules to classify any new feature vector X . The class assigned to X corresponds to the class associated with the rule j for which its degree of fulfillment $\mu_j(X) = \prod_{k=1}^N A_{jk}(X_k)$ is the highest.

III. MOTOR IMAGERY EEG DATA

A. EEG data

In order to evaluate the CFIS, we used the motor imagery data set IIIb of “BCI competition III” [9]. This data set corresponds to three subjects who had to imagine left or right hand movements. The EEG signals were recorded by the Graz group [10] [11], using bipolar electrodes C3 and C4, and were filtered in 0.5-30 Hz. No additional preprocessing was performed before feature extraction.

B. Feature extraction method

The feature vectors were obtained by extracting Band Power (BP) features, as in [10], from both electrodes C3 and C4, in

a statistically optimal time window. Such features are popular and known to be efficient for motor imagery classification [10]. The most reactive frequency bands for each subject were selected offline using a statistical paired t-test which compared the BP means of both classes, for every 1 Hz frequency band between 1 Hz and 30 Hz. As expected, the optimal frequencies for discrimination were found in the α and β bands. The feature extraction method led to a four dimensionnal feature vector $[C3_\alpha, C3_\beta, C4_\alpha, C4_\beta]$ in which Cp_y is the BP value for eletrode Cp in the y band.

IV. FIRST STUDY: PERFORMANCES

We compared the performance of CFIS with that of three classifiers widely used in the BCI community which provided the best results in several BCI experiments [2]. The first classifier was a non-linear Support Vector Machine (SVM) using a Gaussian kernel. The second classifier was a Multi-Layer Perceptron (MLP) with one hidden layer and sigmoid activation functions. The third classifier was a Perceptron, i.e, a Linear Classifier (LC) [2]. The optimal values for the hyperparameters (e.g., regularization parameter C of the SVM) of all classifiers were chosen using 10-fold cross validation.

TABLE I
ACCURACY (%) AND MUTUAL INFORMATION (MI) OF CLASSIFIERS

Subject	CFIS	SVM	MLP	LC
1	86.7±1.6	86.8±0.0	86.6±0.3	84.1±0.9
2	74.7±1.5	75.9±0.5	75.5±0.1	71.8±1.8
3	75.7±0.6	75.4±0.5	74.6±0.1	72.7±2.0
Mean	79±1.2	79.4±0.3	78.9±0.2	76.2±1.6
1	0.49±0.07	0.37±0.12	0.63±0.03	0.45±0.03
2	0.17±0.00	0.20±0.00	0.29±0.02	0.19±0.00
3	0.26±0.01	0.24±0.03	0.29±0.04	0.20±0.03
Mean	0.30±0.03	0.27±0.05	0.40±0.03	0.28±0.02

All classifiers were trained on the provided training sets using the BP features described above. Table I displays the average accuracy and Mutual Information (MI) [11] obtained by each classifier on each subject's test set. In terms of accuracy, CFIS outperformed LC and reached similar results as SVM and MLP. Concerning MI, CFIS performed better than SVM and LC and was outperformed by MLP.

V. SECOND STUDY: READABILITY

The rules automatically extracted by the CFIS from the EEG data are displayed in Figure 1. Interestingly enough, only two fuzzy rules were extracted for each subject.

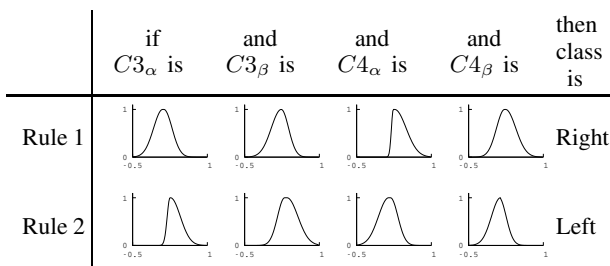


Fig. 1. Fuzzy rules automatically extracted by CFIS for subject 1.

The interpretation of these rules is that the power of electrode C3 in the α and β bands during imagined right hand movements is smaller than that during imagined left hand movements. A symmetric behaviour can be observed for electrode C4. In EEG research, this phenomenon is known as contralateral Event Related Desynchronisation (ERD) [10]. This shows that CFIS is actually readable and it could be used to automatically extract knowledge about the brain dynamics.

VI. THIRD STUDY: EXTENSIBILITY

A. Conception of "hand-made" fuzzy rules

It is possible to add Hand-Made Fuzzy Rules (HMFR) to a FIS as *a priori* knowledge. Typical *a priori* knowledge concerning hand motor imagery EEG data concerns the presence of contralateral ERD in the α and β bands [10]. A human expert could formalize this knowledge using simple rules:

Rule 1: if $C4_\alpha > C3_\alpha$ and $C4_\beta > C3_\beta$ then class is Right

Rule 2: if $C3_\alpha > C4_\alpha$ and $C3_\beta > C4_\beta$ then class is Left

To describe $(Cp_y - Cq_y) > 0$ using membership functions, we could use the following function:

$$h(x) = \begin{cases} 0 & x \leq 0 \\ 1 & x > 0 \end{cases}$$

However, h is a crisp, i.e., a non fuzzy function, which range is in $[0,1]$ and not in $\{0,1\}$. Therefore, we used a sigmoid function g to describe the same relationship as h in a "fuzzier" way. Besides, with λ being the slope of g , when $\lambda \rightarrow \pm\infty$, then $g \rightarrow \pm h$.

Finally, two HMFR using g as membership functions, can be designed to discriminate left or right imagined hand movements as displayed on Fig. 2.

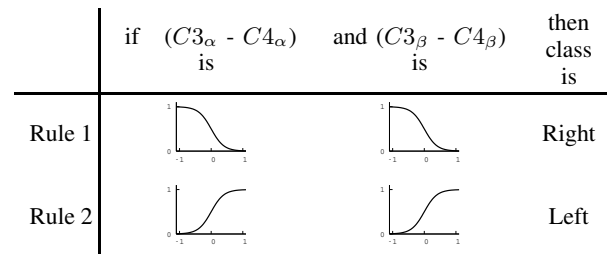


Fig. 2. Hand-made fuzzy rules used to classify motor imagery data

Using schemes of trials and errors on the training sets, the optimal value for λ was chosen to be ± 21500 in the four membership functions. It should be noted that such rules cannot be learnt by the CFIS as they describe relationships between features and not the properties of the features.

B. Performance

We computed the accuracy of the proposed HMFR, as well as the average accuracy of a CFIS that contains both automatically extracted rules and HMFR (see Table II).

These results show first that the HMFR accuracy is much higher than chance, which means that efficient HFMR using *a priori* knowledge on motor imagery can be designed easily and can classify imagined hand movements. Second, they seem to show that adding HMFR to fuzzy rules automatically extracted by a CFIS may be rewarding if the HMFR alone are already efficient (see results obtained for subject 1).

TABLE II
ACCURACY (%) OF TRAINED CFIS VERSUS FIS MADE OF HMFR

Subject	1	2	3
HMFR	87.4%	66.5%	72.6%
CFIS	86.7±1.6	74.7±1.5	75.7±0.6
CFIS+HMFR	88.1±0.4	72.4±1.0	75.6±0.6

VII. FOURTH STUDY: REJECTION OF OUTLIERS

A. Method

Outliers are commonly recorded during BCI experiments. They should not be classified, but rejected in order to reduce the error rate. To evaluate the CFIS outlier rejection capabilities, we generated artificial outliers by randomly placing new feature vectors at a large distance from the feature vectors of the test sets. This distance was selected randomly from 2 to 4 times the standard deviation of all feature vectors to ensure that outliers lied outside the pattern of each class. We added 25% of such outliers to each test set. A third class label was assigned to the outliers to ensure their classification would increase the error rate.

For the CFIS, a feature vector was rejected if the highest degree of fulfillment was smaller than a given threshold. For both SVM and LC, rejection occurred if the absolute value of their output was smaller than the given threshold. For the MLP, rejection was performed if the largest output value was smaller than the threshold. All classifiers were trained on the training set of each subject. Then, we computed the error-reject curves for these classifiers on each subject's test set with added outliers. These curves were computed by gradually increasing the value of the rejection thresholds and computing the error and reject rate for all these values.

B. Results

The error-reject curves for subjects 1 and 2 are displayed on Fig. 3. The curves for subject 3 show exactly the same trends but are not displayed here due to space limitations.

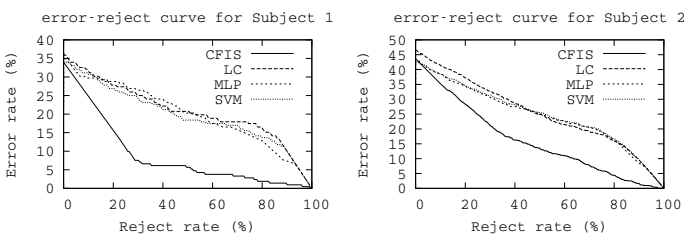


Fig. 3. The error-reject curves for each classifier, of subject 1 and 2.

The error-reject curves for MLP, LC and SVM suggests that these classifiers must reject a lot of feature vectors before reaching a low error rate which means they cannot make a clear distinction between outliers and regular vectors. On the contrary, the area below the CFIS curves is much smaller than the one of any other classifier. This shows that CFIS is able to identify and reject efficiently the outliers, which makes its error rate drop dramatically. These differences can be explained

by the fact that CFIS is a generative classifier which models explicitly the class boundaries using membership functions. On the contrary, MLP, Perceptron and SVM are discriminative classifiers which only model how to distinguish the classes. As such, they can only identify to which of the targeted classes the feature vector most likely belongs to, and not if it actually belongs to one of these classes or not.

CONCLUSION AND DISCUSSION

In this paper we have studied the use of a Fuzzy Inference System (FIS) for motor imagery classification in BCIs. We first studied the performance of the Chiu's FIS (CFIS) [4]. It reached similar results than the most popular classifiers used in BCIs. Second, we stressed the interpretability of the CFIS, which could be used for brain knowledge discovery. Third, we studied the possibility of adding *a priori* knowledge to the CFIS under the form of Hand-Made Fuzzy Rules (HMFR). Suitable HMFR were shown to improve the performance of a trained CFIS. Finally, the CFIS capabilities of rejecting outliers were assessed on artificial data, showing its superiority. Taken together, our results suggest that FIS classifiers are promising classifiers for BCIs as they address several issues raised in the community [3].

Future work could deal with the exploration of different FISs such as NEFCLASS [8], on different EEG data and with the comparison of FISs with other interpretable classifiers such as decision trees.

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, pp. 767–791, 2002.
- [2] F. Lotte, M. Congedo, A. Lécuyer, F. Lamarche, and B. Arnaldi, "A review of classification algorithms for eeg-based brain-computer interfaces," *Journal of Neural Engineering*, vol. 4, pp. R1–R13, 2007.
- [3] D. J. McFarland, C. W. Anderson, K.-R. Muller, A. Schlogl, and D. J. Krusienski, "Bci meeting 2005-workshop on bci signal processing: feature extraction and translation," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 14, no. 2, pp. 135 – 138, 2006.
- [4] S. L. Chiu, "An efficient method for extracting fuzzy classification rules from high dimensional data," *Journal of Advanced Computational Intelligence*, vol. 1, pp. 31–36, 1997.
- [5] S. Guillaume, "Designing fuzzy inference systems from data: Aninterpretability-oriented review," *IEEE Transactions on Fuzzy Systems*, vol. 9, no. 3, pp. 426–443, 2001.
- [6] F. Lotte, "The use of fuzzy inference systems for classification in eeg-based brain-computer interfaces," in *Proc. of the 3rd international Brain-Computer Interface workshop*, 2006, pp. 12–13.
- [7] P. Herman, G. Prasad, and T. M. McGinnity, "A fuzzy logic classifier design for enhancing bci performance," in *Proc. of the 3rd international Brain-Computer Interface workshop*, 2006, pp. 42–43.
- [8] D. Nauck, "Neuro-fuzzy systems: Review and prospects," in *In Proceedings of Fifth European Congress on Intelligent Techniques and Soft Computing (EUFIT'97)*, 1997, pp. 1044–1053.
- [9] B. Blankertz, K. R. Muller, D. J. Krusienski, G. Schalk, J. R. Wolpaw, A. Schlogl, G. Pfurtscheller, J. D. R. Millan, M. Schroder, and N. Birbaumer, "The bci competition iii: Validating alternative approaches to actual bci problems," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 14, no. 2, pp. 153–159, 2006.
- [10] G. Pfurtscheller and C. Neuper, "Motor imagery and direct brain-computer communication," *proceedings of the IEEE*, vol. 89, no. 7, pp. 1123–1134, 2001.
- [11] A. Schlogl, C. Neuper, and G. Pfurtscheller, "Estimating the mutual information of an eeg-based brain-computer interface," *Biomed Tech*, vol. 47, no. 1-2, pp. 3–8, 2002.