Towards Personalized Neural Networks for Epileptic Seizure Prediction

António Dourado, Ricardo Martins, João Duarte, Bruno Direito

Centro de Informática e Sistemas da Universidade de Coimbra
Departament of Informatics Engineering
Pólo II Universidade 3030-290 Coimbra
{dourado,brunodireito@dei.uc.pt }, {rfam, jvd@student.dei.uc.pt}

Abstract. Seizure prediction for untreatable epileptic patients, one of the major challenges of present neuroinformatics researchers, will allow a substantial improvement in their safety and quality of life. Neural networks, because of their plasticity and degrees of freedom, seem to be a good approach to consider the enormous variability of physiological systems. Several architectures and training algorithms are comparatively proposed in this work showing that it is possible to find an adequate network for one patient, but care must be taken to generalize to other patients. It is claimed that each patient will have his(her) own seizure prediction algorithms.

Keywords: Epilepsy, data mining**,** seizure prediction, classification, neural networks.

1 Introduction

About one third of epileptic people, meaning 0.4% of population, are not treatable by medication or ressective surgery [1]. At any time, anywhere, they can suffer from a seizure, "*like a bolt from the sky*", during some seconds or some minutes, seriously affecting their motoricity, perception, language, memory and conscious. It they could predict the seizures, their life would change substantially.

Seizure prediction has been the object of extensive and intensive research for the last 20 years. For an excellent review see for example [2][3]. More recently computational intelligence techniques have been identified as having a high potential [4][5] for seizure identification. Seizure prediction, the object of the present work, is a different problem from seizure identification. Prediction is much more harder than identification. However, from the clinical point of view, no significative practical advance has been verified: there is not any system usable by patients allowing them to predict a coming seizure and to take action to preserve his(her) safety and privacy, improving substantially his(her) social integration. This is probably because most of the researchers look for a general method and algorithm that would work for every patient. And although several authors propose methods to which they claim a high performance, the considered performance criteria is only partial, neglecting other parts of the problem that prevent them to be used in a clinical environment. Physiological systems, as every biological one, have a high variability, and, in the case of seizure prediction, it seems more advisable to look for a predictor well designed for each patient. Neural networks, by their diversity in architectures and training algorithms, have a high plasticity well suited for that purpose.

In the present study this problem is worked out as a classification task. The ElectroEncephaloGram (EEG) is the main electrical measure of the activity of the brain. It is supposed that epileptic seizures are an abrupt change in the electrical activity of the brain and that these changes are captured by the EEG. The challenge is then to process the EEG in such a way that four brain states can be identified: the normal state, the time interval preceding a seizure, the seizure itself, and the time interval for (re)normalization of the brain activity. This cannot be done directly with the EEG; instead some special features must be extracted from the EEG signal. These features must change as the brain evolves among these states and these changes, particularly during the pre-seizure period, may eventually lead the seizure prediction.

In the present work a set of features is extracted from the EEG signal. They quantify several characteristics about energy, time-frequency decomposition, nonlinear behavior, composing a 15 dimensional features space where classification is then to be done into the four classes (brain states): inter-ictal, pre-ictal, ictal, pos-ictal. These features have been considered by several authors with a high potential for the discrimination of the brain state with respect all or some of these four classes [6][7][8][9][10][11][12].

In this work two patients from Freiburg Database [13] are considered. They have been chosen by their different epileptic zones , one in frontal lobe, the other in temporal lobe.

Several architectures and training algorithms are comparatively used for seizure prediction in one and in the other. The performance criteria has three facets: specificity, sensitivity, overall classification rate. The results show that there are several architectures adequate for a patient but they do not work properly for the other patient.

In the next Paragraph 2 the data and the features used for the classification stage are presented. Then in Paragraph 3 the results obtained with several network architectures are discussed. Conclusions and future work are set in the last Paragraph 4.

2. The data and the set of features features for classification

The data used in this investigation has been collected from the epilepsy database of Freiburg Center for Data Analysis and Modeling (FDM) of Albert Ludwigs University of Freiburg [13]. Two patients have been selected, patient A with frontal lobe epilepsy and patient B with temporal lobe epilepsy. The intracranial recordings utilized were acquired using Neurofile NT digital video system with 128 channels, 256 Hz sampling rate, and a 16 bit analogue-to-digital converter.

Applying energy concepts, wavelet transform, nonlinear dynamics, 14 features have been extracted, listed in Table 1.

Intracranial EEG data is processed by the developed methods. The time interval between two consecutive computations of the 14 presented features is 5 seconds. One single chanel of the EEG, the focal one, is used. Other studies use more channels [14].

This section presents an overview of the methods which lead to this set of features. The methods were developed in Matlab and its toolboxes (including Neural Networks Toolbox) [15], and other freely available software, like the nonlinear time series analysis TSTOOL) [16].

Energy variation analysis is based on the algorithm presented in [7]. EEG signal is processed through two windows with different length to analyze energy patterns. The main objective is to confirm the increase of energy bursts in the periods that precede seizures. Accumulated energy was approximated by using moving averages of signal energy (using a short-term energy observation window versus a long-term energy observation window). A similar displacement was applied to both windows and both ended at the same time point. These features allow the observation of energy patterns before epileptic seizures.

Table 1. The 14 extracted features from EEG to be used in classification of the barin state.

Concept	**Feature**
Signal Energy	Accumulated energy
Signal Energy	Energy level
Signal Energy	Energy variation (short term energy)
Signal Energy	Energy variation (long term energy)
Wavelet Transform coeficient energy	short term energy band (0Hz – 12,5Hz)
Wavelet Transform coeficient energy	long term energy band (0Hz – 12,5Hz)
Wavelet Transform coeficient energy	short term energy band (12,5Hz – 25Hz)
Wavelet Transform coeficient energy	long term energy band (12,5Hz – 25Hz)
Wavelet Transform coeficient energy	short term energy band (25Hz – 50Hz)
Wavelet Transform coeficient energy	long term energy band (25Hz – 50Hz)
Wavelet Transform coeficient energy	short term energy band (50Hz – 100Hz)
Wavelet Transform coeficient energy	long term energy band (50Hz – 100Hz)
Nonlinear system dynamics	Correlation dimension
Nonlinear system dynamics	Max Lyapunov Exponent

Wavelet coefficients have been submitted to a similar energy analysis, allowing by this way the identification of variations in the different frequency bands that constitute the EEG signal. Based on the mechanism previously explained, the coefficients obtained by wavelet decomposition are processed and the accumulated energy of these series is determined. As before, accumulated energy was approximated by using moving averages of coefficients energy (using a short-term energy observation window versus a long-term energy observation window). The mother wavelet used in the presented study was daubechies-4; the decomposition was completed with four levels.

Nonlinear analysis faces the EEG as trajectories of a nonlinear system. Two nonlinear dynamic features, maximum Lyapunov exponent and correlation dimension through a sliding window, are computed using [15]. The construction of the attractor, after the determination of the parameters delay time and embedding dimension, allows the calculation of the maximum Lyapunov exponents and correlation dimension. The estimation of the maximum Lyapunov exponents consists in the quantification of the exponential growth of the average distance between two nearby trajectories of the attractor, through error approximation. Correlation dimension is determined by takens estimator method [15].

The joint analysis of the extracted features created a 14-dimension space which represents the EEG signal in several components (energy signal, frequency and system dynamics). The objective of the study is to investigate the eventual occurrence of hidden characteristics in data such that clusters can be discovered allowing an acceptable classification of EEG data into 4 classes:

- inter-ictal (normal EEG pattern)
- pre-ictal (two minutes prior to the seizure onset)
- ictal (the seizure onset)
- pos-ictal (two minutes subsequent to seizure end)

One cycle is composed of one series of these classes.

The overall approach is illustrated in Fig. 1.

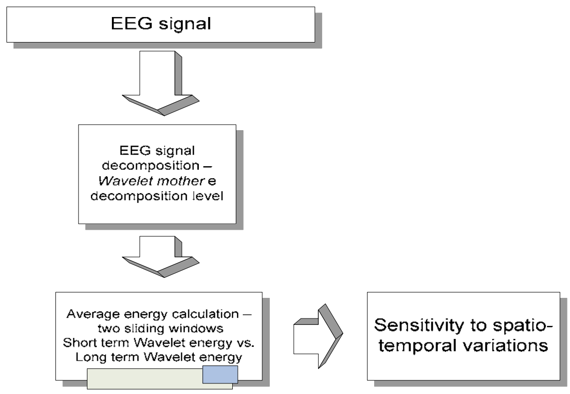

Figure 1. The approach EEG –features extraction-classification

3 The applied neural network architectures and its results

The data sets have the following characteristics: patient A- 2 cycles, 1821 points; patient B- 3 cycles, 2432 points. Data have been normalized feature by feature in [0 1].

The best found architectures and training algorithms

After an extensive experimentation, the following neural network structures have been applied and compared, because they have been found to be the most promising:

(i) Three layer logsig feedforward (FFNN): 14 neurons in the first layer, 56 in the second and 4 in the output layer, Fig. 2. The output layer numerical values are rounded to integers and it has been trained to classify each class accordingly to the following coding:

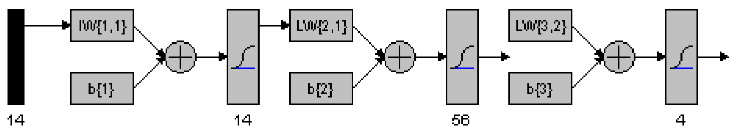

Fig. 2. The best architecture found patients A and B. Bias and weights are proper to each patient.

Class 1 (inter-ictal) : [1 0 0 0] Class 2 (pre-ictal): [0 1 0 0] Class 3 (ictal): [0 0 1 0] and Class 4 (pos-ictal) [0 0 0 1].Training was done using the Levenberg-Marquardt algorithm, better that the backpropagation one.

(ii) In order to catch the nonlinear dynamic nature of the signal, experiments have been made introducing a taped delay line in the network inputs (first layer), as implemented in the Matlab Neural Networks Toolbox. Delays of 1 and 2 have been experimented.

(iii) Radial Basis Function neural network (RBF) with variable size of the first (radial) layer and 4 linear neurons in the output layer. It was trained by the hybrid algorithm.

The performance criteria of the classifier

In seizure prediction (as in the general problem of medical diagnosis) there are four possible outcomes to a diagnosis operation:

- positive true (PT), when the diagnose is positive and the event has been confirmed,
- positive false (PF) when the diagnose is positive and the event has not been confirmed,
- negative true (NT), when the diagnose is negative and the event has been confirmed as not existing,
- negative false (NF) when the diagnose is negative abut the event has finally been confirmed as true.

For clinical applications any automatic diagnosis systems must give all the PT events and all the NT events. But it must also give zero PF and zero NF answers. Two performance indexes are defined:

Sensitivity: related to the Positive outcome, given by (1)

$$SENSIT = \frac{PV}{PV + NF} \tag{1}$$

Specificity: related to Negative outcome, given by (2)

$$SPECIF = \frac{NV}{NV + PF} \tag{2}$$

We can also define the overall index given by (3)

$$OVERALL = \frac{PV + NV}{PV + PF + NV + NF} = \frac{PV + NV}{ALL} \tag{3}$$

It is frequent that one author presents one of these indexes to measure the performance of a seizure prediction algorithm. However from a clinical judgment, only when the sensitivity and the specificity are **both** near to 1 can the algorithm be applied. A perfect system has both sensitivity and specificity equal to one. Probably that is why very few application of automatic diagnosis systems are really working today, although there is an extensive published literature on diagnosis algorithms with or high sensitivity or high specificity.

All the three indexes are used, as shown in the interface in Fig. 3.

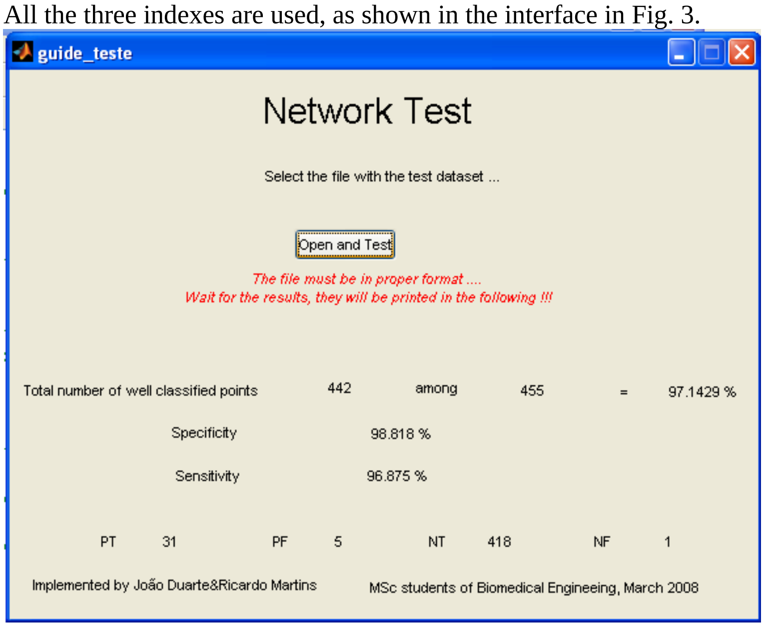

Figure 3. The interface for testing the networks. The networks, data sets and interface are available to download at eden.dei.uc.pt/~dourado/seizureprediction. It works under Matlab 2007b.

Results

Table 2 shows some results for patient A. The FFNN has been used with and without input delays. But, although in theory a better result could be expected with delays (considering the brain as a dynamic system), in fact these two networks have a much worse performance. RBF shows also a poor performance. If one only cares about specificity, then all four nets are very good. The FFNN with 2 delays shows an absolutely good specificity of 1 and an absolutely bad sensitivity of 0. This illustrates the fact that only one of these parameters is not a proper performance index.

Table 2. Some results for patient A wth FFNN and RBF in test training set. The training criteria has been SSE (sum of Squared Error), with Levenberg-Marquard algorithm. Data is normalized.The last line if for RBF

Input Delay (FFNN)	Size of test data set	Nº well classified	PT	PF	NT	NF	SENSIT	SPECIF
0	455	441	31	5	418	1	0.9688	0.9882
1	409	252	0	2	348	59	0	0.9943
2	409	203	0	0	350	59	0	1
RBF	455	391	14	1	422	18	0.4375	0.9976

Table 3 shows similar results for patient B and the same comments can be done. The RBF in last line has been trained and tested, in this case, with original, non normalized data. It shows a slightly better specificity but a much worse sensitivity.

Table 3. Some results for patient B wth FFNN and RBF in test training set. The training criteria has been SSE (sum of Squared Error), with Levenberg-Marquard algorithm. Data is normalized.The last 2 lines are for RBF: normalized and original data

Input Delay (FFNN)	Size of test data set	Nº well classified	PT	PF	NT	NF	SENSIT	SPECIF
0	608	594	55	4	544	5	0.9167	0.9967
1	409	252	0	2	348	59	0	0.9943
2	455	441	31	5	418	1	0.9688	0.9882
RBF	608	499	48	52	496	12	0.8	0.9054
RBF	608	465	1	0	548	59	0.0167	1

If the FFNN is trained simultouneously for the datasets of both patients, the training performance is rather poor. It is very hard, because of the different patients and different types of epilepsy, to find a network that, with the same weights and bias, works well for both. Of course one can always increase the dimension and improve training, until probably overtraining, loosing the generalization capability of the network. From a pratical clinic use, for example in ambulatory, where a patient transports with him some alarmig device forecasting the eminent coming of a seizure, the need for a personalized neural network is not a serious problem.

Table 4. Case of joining the data sets of both patients (training results) with Levenberg-Marquardt algorithm.

Input Delay (FFNN)	Size of training data set	Nº well classified	PT	PF	NT	NF	SENSIT	SPECIF
0	1064	912	18	12	958	75	0.20	0.98

Testing the network A in patient B, or network B in patient A, gives the results presented in table 5. The degradation of performance is evident.

Table 5. Testing patient A network into patient B and vice-versa

Case	Size of test dataset	Nº well classified	PT	PF	NT	NF	SENSIT	SPECIF
A in B	608	406	0	35	513	60	0	0.9361
B in A	455	265	2	69	354	30	0.625	0.8368

4. Conclusions

There is still a long way to set extensive guidelines for building seizure predictors for epileptic patients. However the shown results evidence two simple principles: (i) there is no general predictor good for all patients, and (ii) the predictor of one patient is not acceptable for other patient. This has as consequence that each patient must be the object of a personalized study, using as much data as possible, following its behavior and training it permanently. Neural networks have a high plasticity that can be profitably used for this purpose. However, other techniques should also be studied, such as support vector machines (SVM) that may have an important role in constructing nonlinear boundaries in the high dimensional features space, resulting eventually in better classification among the four classes in the context of seizure prediction.

References

[1] T. Browne, G. Holmes, *Handbook of epilepsy*, Lippincott Williams & Wilkins, 2000.

[2] Review 1 F. Mormann, R.G. Andrzejak , C.E. Elger ,K. Lehnertz , "Seizure prediction: The long and winding road", Brain 130:314-333, 2007

[3] Review 2 B. Schelter, M. Winterhalder, G. Feldwisch , H. Drentrup, J. Wohlmuth, J. Nawrath, A. Brandt, A. Schulze-Bonhage, J. Timmer, J., "Seizure prediction: The impact of long prediction horizons" Epilepsy research 73 : 213-217, 2007 .

[4] Guler, I., and E. D .Ubeyli, Application of adaptive neuro-fuzzy inference system for detection of electrocardiographic changes in patients with partial epilepsy using feature extraction, Expert Systems with Applications 27 (2004) 323–330.

[5] Subasi A., Application of adaptive neuro-fuzzy inference system for epileptic seizure detection using wavelet feature extraction, Computers in Biology and Medicine 37 (2007) 227–244

[6] B. Litt, R. Esteller, J. Echauz, M. D'Alessandro, R. Shor, T. Henry, et al., "Epileptic seizures may begin hours in advance of clinical onset: a report of five patients", *Neuron*, 30, 2001, pp. 51–64.

[7] R. Esteller, J. Echauz, M. D'Alessandro, G. Worrell, et al., "Continuous energy variation during the seizure cycle: towards an on-line accumulated energy", *Clinical Neurophysiology*, 116, 2005, pp. 517-526.

[8] S. Gigola, F. Ortiz, C. D'Atellis, W. Silva, and S. Kochen, "Prediction of epileptic seizures using accumulated energy in a multiresolution framework", *Journal of Neuroscience Methods 138*, 2004, pp. 107-111.

[9] W. Chaovalitwongse, LD Iasemidis, PM Pardalos, PR Carney, DS Shiau, JC. Sackellares, "Performance of a seizure warning algorithm based on the dynamics of intracranial EEG", *Epilepsy 64*, 2005, pp. 93–113.

[10] Subasi, "Epileptic seizure detection using dynamic wavelet network", *Expert System with applications*, 29, 2005, pp. 343-355.

[11] M. Kemal Kiymik, Í. Guler, A.Dizibuyuk, M. Akin, "Comparison of STFT and Wavelet Transform Methods in Determining Epileptic Seizure Activity in EEG Signals for real-time application", *Computers in Biology and Medicine*, 35, 2005, pp.603-616

[12] G. Ouyang, X. Li, Y. Li, X. Guan, "Application of wavelet-based similarity analysis to epileptic seizures prediction", Computers in Biology and medicine, 37, 2007, pp.430-437

[13] Freiburger Zentrum fur Datenanalyse und mollbildung, Available: http://www.fdm.uni-freiburg.de/groups/timeseries/epi/EEGData/download/infos.txt.

[14] M. Le Van Quyen, F. Amor, and D. Rudrauf, "Exploring the dynamics of collective synchronizations in large ensembles of brain signals" J Physiol Paris. In press., 2007.

[15] The Mathworks, Inc

[16] C. Merkwirth,, U. Parlitz, I. Wedekind, W. Lauterborn , "TSTOOL User Manual", Version 1.11. http://www.dpi.physik.uni-goettingen.de/tstool/HTML/index.html.

Acknowledgements

This research is supported by European Union FP7 Framework Program, EPILEPSIAE Grant 211713 and by CISUC (FCT Unit 326). The authors express their gratitude to Freiburg Center for Data Analysis and Modeling (FDM) of Albert Ludwigs University of Freiburg, for the access to the epilepsy database.