

Prediction of the Epileptic Events 'Epileptic Seizures' by Neural Networks and Expert Systems

Kifah Tout, Nisrine Sinno, and Mohamad Mikati

Abstract—Many studies have focused on the nonlinear analysis of electroencephalography (EEG) mainly for the characterization of epileptic brain states. It is assumed that at least two states of the epileptic brain are possible: the interictal state characterized by a normal apparently random, steady-state EEG ongoing activity; and the ictal state that is characterized by paroxysmal occurrence of synchronous oscillations and is generally called in neurology, a seizure.

The spatial and temporal dynamics of the epileptogenic process is still not clear completely especially the most challenging aspects of epileptology which is the anticipation of the seizure. Despite all the efforts we still don't know how and when and why the seizure occurs. However actual studies bring strong evidence that the interictal-ictal state transition is not an abrupt phenomena. Findings also indicate that it is possible to detect a pre-seizure phase.

Our approach is to use the neural network tool to detect interictal states and to predict from those states the upcoming seizure (ictal state). Analysis of the EEG signal based on neural networks is used for the classification of EEG as either seizure or non-seizure. By applying prediction methods it will be possible to predict the upcoming seizure from non-seizure EEG.

We will study the patients admitted to the epilepsy monitoring unit for the purpose of recording their seizures. Preictal, ictal, and post ictal EEG recordings are available on such patients for analysis. The system will be induced by taking a body of samples then validate it using another. Distinct from the two first ones a third body of samples is taken to test the network for the achievement of optimum prediction. Several methods will be tried 'Backpropagation ANN' and 'RBF'.

Keywords—Artificial neural network (ANN), automatic prediction, epileptic seizures analysis, genetic algorithm.

I. INTRODUCTION

THE work accomplished in this study aimed to find among a set of EEG digital recordings [16], the set of features derived from the EEG which can be used to distinguish seizures from non-seizures sections of EEG [1]–[5] and [9]–[12].

The work is a continuation of a study made in 1999 at

K. Tout and N. Sinno are with the Department of Computer Science, Lebanese University, Hadath, Beirut, Lebanon (e-mail: ktout@ul.edu.lb; nisrine@ul.edu.lb).

M. Mikati is with the Department of Pediatrics and Adolescent Medicine, American University of Beirut Medical Center, Room C-662, Riad El-Solh, Beirut 11072020, Lebanon (phone: 961-1-353465 ext 5500; fax: 961-1-370781 e-mail: mamikati@aub.edu.lb, <http://www.aub.edu.lb/~webped/>).

Oxford University [6][13]. We have applied the classification methods on long term recordings of 5 patients having several seizures per patient, and then studied the application of neural networks for the classification of EEG, as either seizure or non-seizure. We have applied several methods to make the prediction of future seizures from a previous seizure section as well as from a non-seizure section.

The training of the neural network system was done using a set of data from a number of patients while the test of the system was done on a different set of data from other patients.

We have compared the results according to the sensitivity, and specificity of detection [6],[14] and we have concluded that with MLP we can predict a seizure 460ms before it happens, but if we can develop the study on time series networks, longer term prediction could be done. The following two figures; Fig. 1 and Fig. 2 represent an ictal section and a non-ictal section.

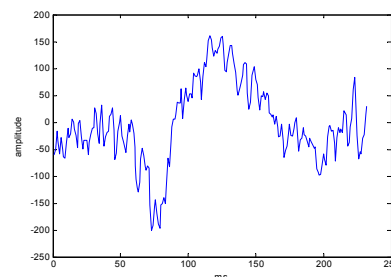


Fig. 1 Section of 250 ms of 'ictal' phase

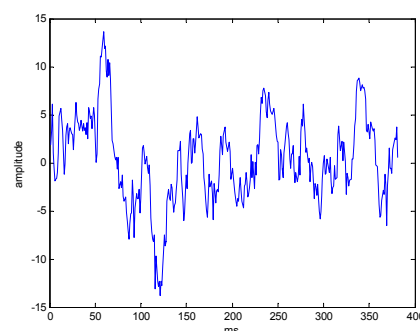


Fig. 2 Section of 400 ms of 'non-ictal' phase

The EEG signal has been partitioned in sections from 80ms to 1000ms, this choice of the length is made to assure that several seizures spikes can be included inside the sections away from the limits and the sections are small enough in order to have a stationary signal in the frequency domain.

II. METHODS AND RESULTS

Among the features, we have chosen the reflection coefficients k_1 and k_2 that provide frequency domain information.

We have calculated k_1 and k_2 using the linear prediction:

For a signal x the linear prediction (p order) of the element $y(n)$ is a linear combination of the p preceding values $y(n-1)$, $y(n-2)$,, $y(n-p)$ therefore:

$$\hat{y} = -\sum a_k y(n-k) \quad \text{with } k = 1, \dots, p \quad (1)$$

a_k are the prediction coefficients. Prediction error for p -order predictions is given by:

$$f_p(n) = y(n) - \hat{y}(n) \quad (2)$$

The construction of reflection coefficients (k_m) use the error generator ($g(n)$) as a filter with input y . The Filter is given by the following equation:

$$\begin{aligned} f_0(n) &= g_0(n) = y(n) \\ f_m(n) &= f_{m-1}(n) + k_m g_{m-1}(n-1) \\ g_m(n) &= k_m f_{m-1}(n) + g_{m-1}(n-1) \\ \text{with } m &= 1, 2, \dots, p \end{aligned} \quad (3)$$

The classification between ictal and non-ictal sections according to k_1 , k_2 is given in Fig. 3 and Fig. 4.

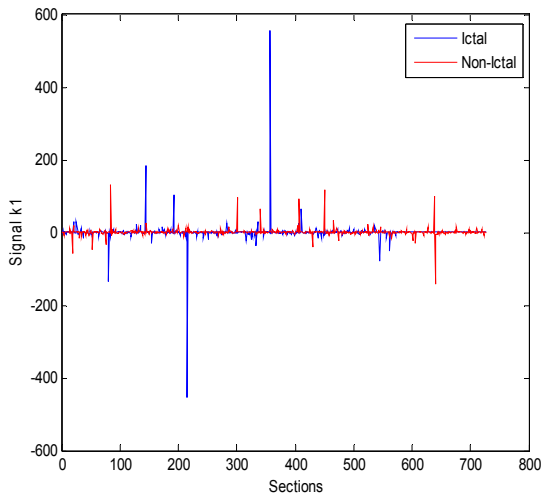


Fig. 3 Classification according to k_1

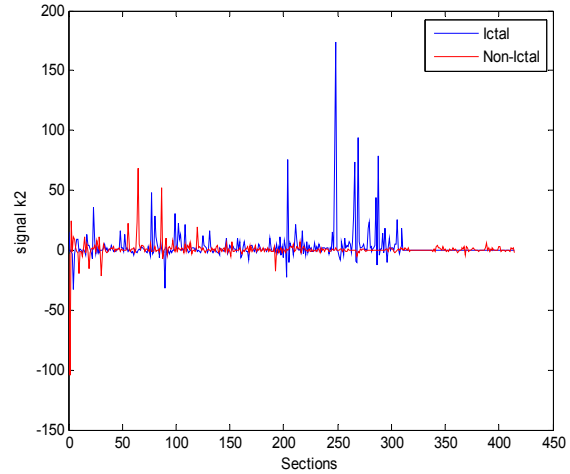


Fig. 4 Classification according to k_2

From Figs. 3 and 4 we can deduce that in ictal sections the reflection coefficient values are superior to their values in non-ictal sections.

Another feature has been used, the total power of the signal:

$$P_{\text{tot}} = \int |X(\omega)|^2 d\omega \quad (4)$$

With the Fourier transform coefficients calculated between 2 frequencies. The classification between ictal and non-ictal sections is represented as follows:

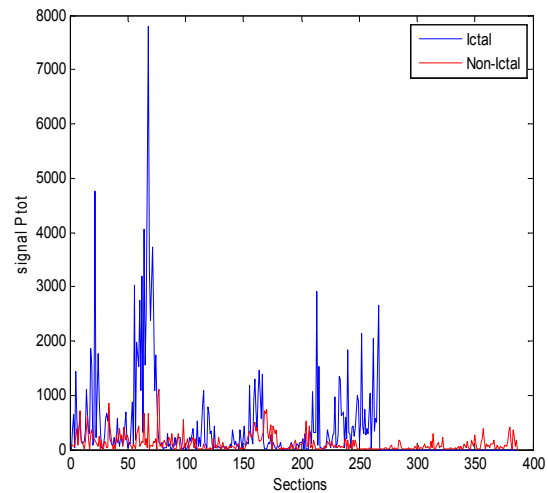


Fig. 5 Histogram of the total power P_{total}

We found that the total power parameter is a reliable classifier parameter for any kind of EEG waveform. The power in ictal sections is much higher than in Non-ictal sections.

Sometimes ictal sections can have the anomaly 'spike and wave' with spikes of EEG, accompanied with generalized wave represented by high amplitude over a frequency band of 20 Hz. Non-ictal sections also includes sometimes the

generalized wave. But what we experienced is that in all cases power in Ictal sections is always higher than in Non-Ictal sections.

We have also used the limited power after filtering the signal between 3-20Hz, in order to limit the effect of very low and very high frequency. The histogram is shown in Fig. 6:

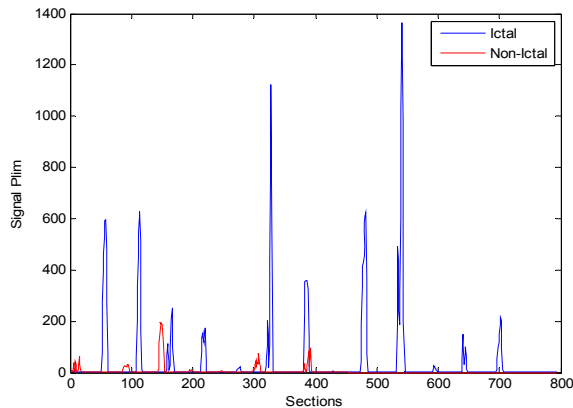


Fig. 6 Histogram of the limited power Plim

Among all those parameters k_1 , k_2 , P_{total} , P_{lim} , the total power is the most reliable parameter for classification even if the limited power is better than the reflection coefficients k_1, k_2 .

The rhythmicity feature of ictal sections in the EEG signal has been used for the classification between ictal and non-ictal sections:

It is possible that the power in certain frequency bands will reveal differences between ictal and non-ictal signals.

The rhythmicity is calculated by using the power in the dominant frequency

$$P_{\max} = \max\{(\Delta f / 2) (|X_{i-1}|^2 + 2 |X_i|^2 + |X_{i+1}|^2)\} \quad (5)$$

With

1. $0 \leq i \leq k$ k is the number of points in the frequency domain.
2. Δf is the frequency margin between successive points.
3. X_i is the i th point in the frequency domain.

The dominant power could be very close to the remaining signal power, in Ictal sections and much higher than the remaining signal power in Non-Ictal sections. Thus, we define

$$\text{rhythmicity} = \frac{\text{Power in the dominant frequency band}}{\text{Remaining signal power}} \quad (6)$$

The histogram of the rhythmicity, Fig. 7 shows that the rhythmicity is not a good factor of classification between Ictal and Non-Ictal sections because the values in both sections are close to each other.

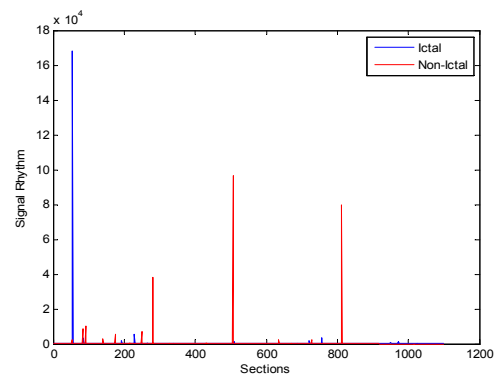


Fig. 7 Histogram of the rhythmicity

We represent the visualizing of pairs of parameters and their effect on the classification. Fig. 8, Fig. 9, and Fig. 10 show that with 2 parameters classification is better.

A) Prediction System Based on Neural Networks

The parameters P_{total} , P_{lim} , k_1 , k_2 , and the rhythmicity will be used as inputs to train several forms of neural networks.

The function of prediction is:

$$x(t+d) = f(x(t_1), \dots, x(t_n+1)) = f(y(t)) \quad (7)$$

with $(y(t))$ is the vector of n values of x at n instants and $d=1$, $f()$ predict the next value $x(t+1)$ of $x(t)$.

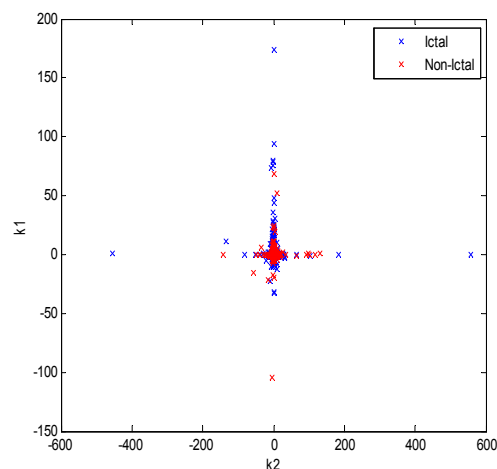


Fig. 8 k_1 versus k_2 for both ictal and non-ictal data

TABLE I

SAMPLE OF THE DATA-FILE REPRESENTING THE 5 PARAMETERS AND THE TYPE OF THE SECTION FROM WHERE THEY ARE USED IN EACH CASE Ictal AND NON-ICTAL CALCULATED

k1	k2	tot-p	b-lim-p	rhythm	Time (sec)	types
3.285997	0.693958	1106.464	243.111	0.000125	6.44	ep-sp
0.601432	-1.56869	638.8328	143.4019	0.074267	6.9	ep-sp
2.426349	5.637564	0	0	0.870225	368.46	non-sp
-3.27822	-15.8026	158.0583	4.484681	27.33341	198.72	ep-sp
-15.5915	1.780949	0	0	0.741701	369.38	non-sp
1.129749	-0.66082	0	0.267886	0.68781	369.84	non-sp
-4.08498	-1.2235	141.965	12.56221	0.710469	198.26	ep-sp
0.056722	0.168244	0	0	0.801182	368.92	non-sp
1.489458	-2.12027	3.206554	0.140827	353.8896	199.18	ep-sp
0.104254	1.217091	150.2014	0	2.355663	199.64	ep-sp

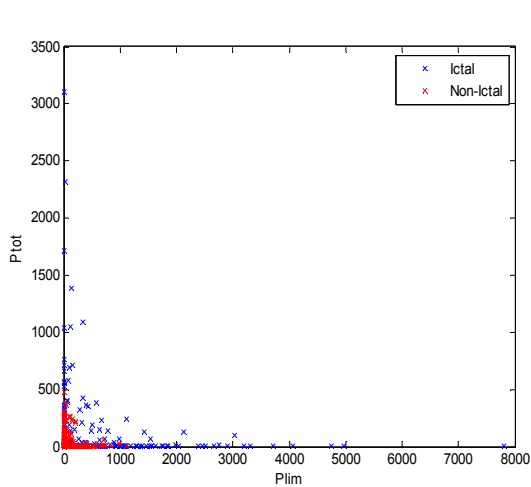


Fig. 9 Ptot versus Plim for both ictal and non-ictal data

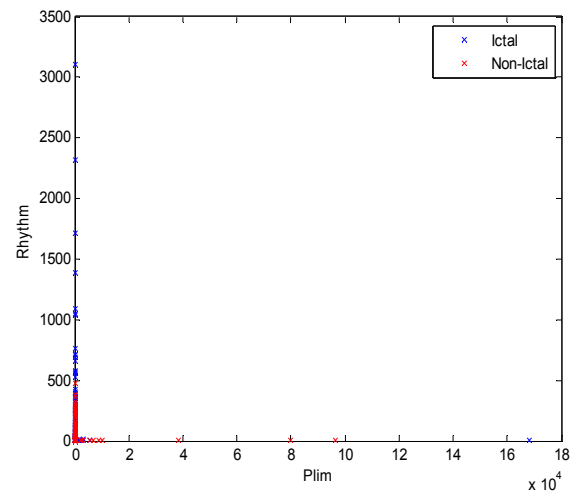


Fig. 10 Rhythmicity versus Plim for both ictal and non-ictal data

1) Training

At every prediction we have changed the dimension of the input vector, the number of hidden layers and the number of nodes. The optimal dimension of the input vector is given according to the efficiency of the prediction that is given by the two parameters sensitivity and specificity. We define sensitivity and specificity as follows:

$$\text{Sensitivity} = \frac{\text{correctly predicted positives}}{\text{total actual positives}} \quad (8)$$

$$\text{Specificity} = \frac{\text{correctly predicted negatives}}{\text{total actual negatives}} \quad (9)$$

(positives means ictal, and negatives means non-ictal).

The optimal numbers of hidden layers as well as the optimal number of nodes defined during the training are determined using the genetic algorithm.

2) Cross-Validation

Periodically the network is tested using the cross-validation, using other set of data than the one used for training and the performance must increase from test to test, if it doesn't the training is stopped. The cross-validation is a recommended criteria to stop training at the right moment.

3) Test

We have used another patient with new data file containing 500 sections as follows:

We have predicted ictal sections from previous ictal sections, and we have predicted ictal and preictal sections (sections that appears between two ictal sections) from both ictal and non-ictal sections.

TABLE II
NUMBERS OF EPOCHS USED FOR THE TEST

	Number of epochs (‘Training’)	Epochs for testing by (‘Cross validation’)
Ictal	500	250
Non ictal	500	250

In order to assure the convergence of the prediction we used random inputs (using the randomizing technique), and by varying the momentum m in the following formula where m is added at the last value of the weight:

$$\Delta w_{ij}(t) = \mu_i \delta_i y_i + m \Delta w_{ij}(t-1) \quad (10)$$

with $0 < m < 1$, as a new global parameter obtained during the test by calculating the error of prediction.

We have used the 5 input parameters k_1 , k_2 , P_{total} , $Plim$, Rhythmicity, but we predicted only one parameter that is k_1 , so the output of the network has one node.

a) Prediction of the Reflection Coefficient k_1 during the Training

The training was done with k_1 as input. The training needed a network MLP with 14 layers of 6 nodes each. Fig. 11 is the output detected and desired for k_1 .

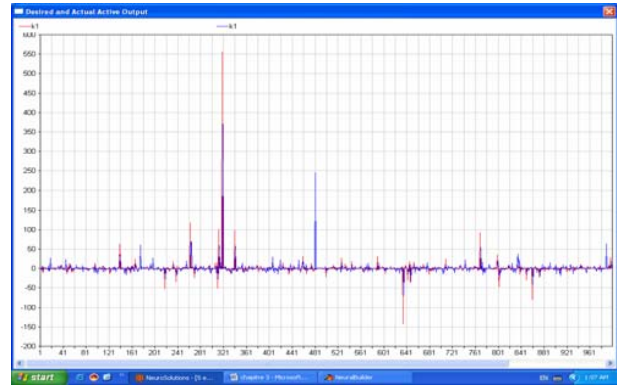


Fig. 11 Detection of k_1 with MLP during training (in blue k_1 detected and in red desired k_1)

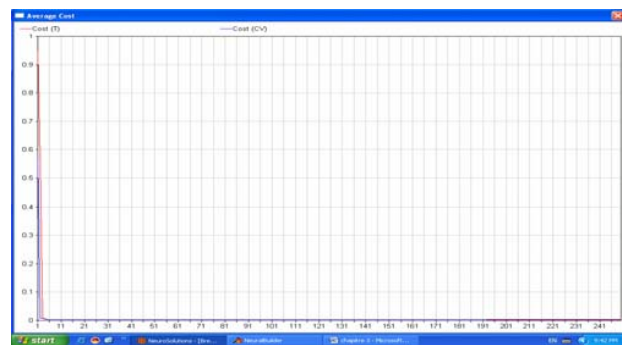


Fig. 12 The detection error of k_1 . Error is converging rapidly

We have tested the network trained on a third patient for prediction: Fig. 13 shows the prediction of next ictal section from previous ictal sections with one parameter k_1 as input of MLP.



Fig. 13 Test of 400 sections concerning the third patient by MLP network (in red, values desired and in blue, values predicted after 0.46s, predicting k_1 of the next ictal section).

The prediction is poor with only k_1 as input of MLP, the sensitivity concerning 400 epochs is 10 %.

Fig. 14 shows the prediction of next ictal section from 'ictal' and 'non ictal' sections with one parameter as input of MLP:

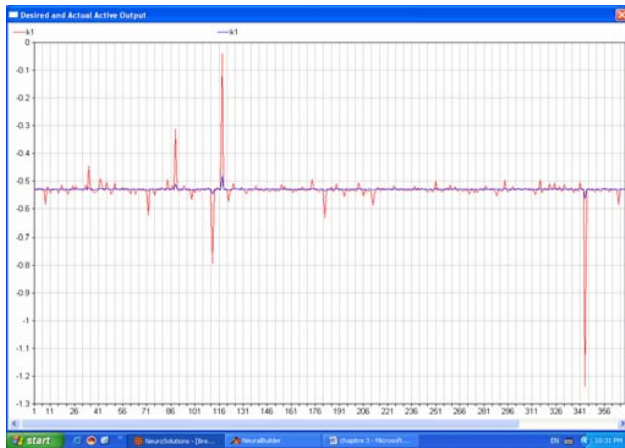


Fig. 14 Test of 400 sections onset of the third patient by MLP (in red, desired values and in blue predicted values after 0.46s predicting k1 of the next section)

Sensitivity = 10 % Specificity = 2 %

The prediction of k1 is poor even when 'ictal' and 'non ictal' sections are used, we explain those results from the fact that using only k1 as input parameter is not enough, it is better to use the 5 defined parameters as input.

b) Prediction of Coefficient k1 from k1, k2, P_{tot}, P_{lim}, rhythmicity as Input Parameters the Training Results in a Network of 20 Layers with 5 Nodes Each

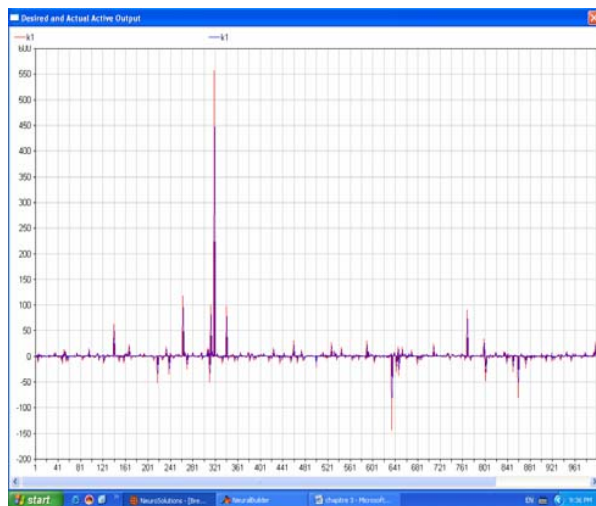


Fig. 15 Detection of k1 while training with MLP (in blue), and desired k1 (in red)

The results of the detection is much better, detected and desired k1 have close values.

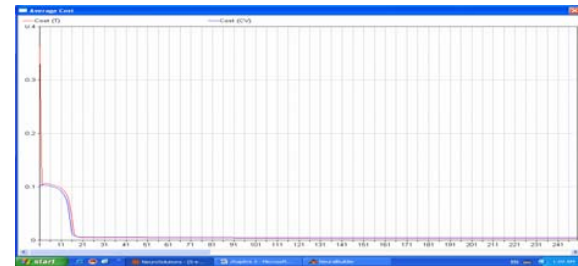


Fig. 16 Training error (blue) and cross validation error (in red)

The training with 5 input parameters and with one input parameter, results in the rapid convergence of the detection error to 0.

We have tested the prediction of ictal sections from previous 'ictal' sections with 5 parameters as inputs of the trained MLP, on a third patient and Fig. 17 shows the result:

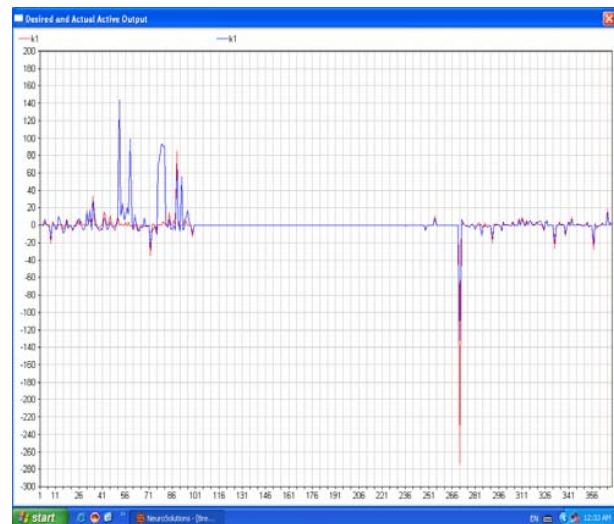


Fig. 17 Test of MLP prediction network for the third patient

In this case the prediction with 5 input parameters is better, values predicted (in blue) and desired in red. **The prediction is good because sensitivity is calculated and it is 88%.**

Now we have done the prediction of k1 in the next ictal section from both previous 'ictal' and 'non ictal' sections, Fig. 18 shows the result:

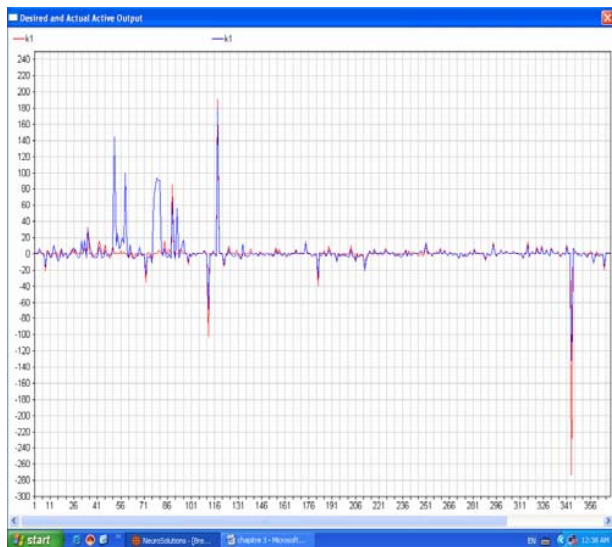


Fig. 18 The test of MLP on the third patient

Sensitivity = 88% Specificity = 92 %

The conclusion is that with an MLP of 100 nodes it is possible to predict the next section before 0.46s with a high efficiency from both previous ictal and non-ictal sections.

III. CONCLUSION

In this research we have tried to predict the epileptic seizure based on neural networks. We have applied the parameters that most likely can represent the long term EEG signal as inputs of the multilayer neural network. We have trained the network in order to detect the ictal and non-ictal sections, then we have tested the network for prediction and we have determined the sensitivity and specificity of the prediction. We can conclude that with 5 parameters used as inputs of the MLP network, the prediction has a high sensitivity and a high specificity (88%). The MLP can predict at short term (0.46s), only the next section. In order to predict long term sections we must use recurrent networks TLRN ('Time Lag Recurrent Network'), RN ('Recurrent Network'), and Elman & Jordan [14], [19].

But it has been demonstrated in [7], [18] that in some cases of chaotic time series MLP networks are superior to recurrent networks in term of rapid convergence and results stability.

We suggest that in further studies we apply the recurrent networks for the long term prediction of ictal sections in order to find the best way of seizure prediction either with MLP or with recurrent network.

But recurrent networks need much more layers and much more nodes, as well as more onsets on more patients; professional neural network software is needed to realize our aims.

REFERENCES

- [1] Josefina Gutierrez, Rogelio Alcantara, and Veronica Medina, "Analysis and Localization of Epileptic Events Using Wavelet Packets" *Medical Engineering & Physics*, Oct. 23, 2001.
- [2] M. A. Cody, "The fast wavelet transform", *Dr Dobb's Journal*, pp. 16-28., April 1992.
- [3] A. Bruce, D. Donoho, and H. Y. Gao, "Wavelet Analysis". *IEEE Spectrum*, pp. 26-35, October 1996.
- [4] Josefina Gutierrez, Rogelio Alcantara, "Spikes characterization on EEG signal by wavelet coefficients". *Fifth Conference of the European Society for Engineering and Medicine Abstracts*, 1999, pp.159-160.
- [5] R. Cerf, H. el Ouasdad, and M. el Amri, "EEG-detected episodes of low-dimensional self-organized cortical activity and the concept of a brain attractor", In Uhl C(ed) *Analysis of neurophysiological brain functioning*. Springer, Berlin Heidelberg New York, pp.126-144, 1999.
- [6] H. Haken, *Principles of brain functioning: a synergetic approach to brain activity, behavior and cognition*. Springer Verlag, Berlin Heidelberg New York, 1996.
- [7] A. Fuchs, J.A.S. Kelso, H. Haken "Phase transitions in the human brain spatial mode dynamics". *Int J Bifurc Chaos* 2, pp. 917-939.
- [8] Roger Cerf, El Hassan El Ouasdad, Philippe Kahane, "Criticality and synchrony of fluctuations in rhythmical brain activity: pretransitional effects in epileptic patients", 2004.
- [9] Generoso Gascon and Mohamed Mikati, "Seizures and Epilepsy", Department of Clinical Neurosciences, Brown University School of Medicine and Department of Pediatrics, American University of Beirut, School of Medicine.
- [10] Walter Van Emde Boas and Jaime Parra, "Long-Term Noninvasive Video Electroencephalographic Monitoring in Temporal Lobe Epilepsy", Department of Electroencephalography and Epilepsy Monitoring Unit, Epilepsy Clinic Meer & Bosch, Dutch Epilepsy Clinics Foundation, the Netherlands,.
- [11] Leon Iasemidis, Deng-shan shiau, Chris Sackellares, Panos Pardalos, and Awadhesh Prasad, "Dynamical Resetting of the Human Brain at Epileptic Seizures". *IEEE Transactions on Biomedical Engineering*, vol. 51, no. 3 march 2004.
- [12] Roger Cerf, El Hassan El Ouasdad, and Philippe kahane, "Criticality and Synchrony of Fluctuations in Rhythmical Brain Activity: Pretransitional Effects in Epileptic Patients", *Biological Cybernetics*, March 2004.
- [13] L. Tarassenko, "Neural Network Detection of Epileptic Seizures in the Electroencephalogram", Oxford University, Department of Engineering Science, February, 1999.
- [14] José Principe, Neil Euliano, and W.Curt Lefebvre. *Neural and Adaptive Systems*, John Wiley & Sons, Inc, 2000.
- [15] Russell Eberhart, Roy Dobbins. "Neural Network PC Tools", Academic Press, Inc, 1990.
- [16] M. Mikati, *Système d'enregistrement digital du signal EEG*, Internal Report, American University Hospital, Beirut, July, 2004.
- [17] Ahmad Fadi, "Analyse Des Evénements Epileptique a l'aide Des Transformées et Paquets d'Ondelettes", M.S. Thesis, Dept. Comp. Science, Lebanese University, Dec. 2004.
- [18] Jerome T. Connor, R. Douglas Martin, and L. E. Atlas, "Recurrent Neural Networks and Robust Time Series Prediction", *IEEE transactions on neural networks*, vol. 5, no. 2, March 1994.
- [19] R. Mikolajczak, "Comparative study of logistic map series prediction using feed-forward, partially recurrent and general regression networks", *In Proceedings of the 9th International Conference on Neural Information Processing ICONIP*, Volume 5, Issue 18-22, Nov. 2002, pp. 2364 - 2368.

Kifah Tout is professor of computer Science at the Lebanese University, Faculty of Sciences I, Beirut. He holds a PhD degree in computer science from Loughborough University of Technology (LUT), UK, 1992. He had more than 15 years Teaching and Research career at LUT, University College London, England, Institute of Cybernetics and Systems Theory, TU Berlin, Germany and NDU, Balamand, AUST, and the Lebanese University, Lebanon. Dr.

TOUT has a strong IT consulting experience with SAP, Brainware, Cap Gemini and Hermed. He was Technical Director for the European Projects PAPAGENA (Parallel Programming Environment for Genetic Algorithms) and HANSA (Heterogeneous Application Generator Standard Architecture for the use of Neural Networks) with collaboration from European partners such as Thorn EMI srl (England), Brainware (Germany), Olivetti Group (Italy), University College London, CAP Volmac (Holland), IMAG (France), German Computer Research Centre (Germany), and Institute for Cybernetic and Systems Theory (Germany). He is a member of IEEE, ACM, and EU Team of Experts. Dr. Tout currently supervises two PhD students in Knowledge Management and indexing and Searching multimedia content.

Nisrine Sinno, « Doctorat troisième cycle Electroniques » - Institut national polytechnique de Grenoble (I.N.P.G /E.N.S.E.R.G) 1983. Research work consists of circuits simulation and linear, nonlinear modeling, stochastic approach, performance of numerical analysis in simulating networks, Epileptic signals analysis using the wavelet packets approach.

Teaching since 1984,

- Lebanese University-Faculty of engineering 1984-1988, taught electronics, signal processing, telecommunications, Networks, instrumentation and supervised final projects.
- Lebanese University-Faculty of Sciences –1989 à 1995, 2000-2008, Taught advanced programming, networks microprocessors, computer systems and architecture.

Mohamad Mikati MD, Director, Adult and Pediatric Epilepsy Program, Professor and Chairman, Department of Pediatrics and Adolescent Medicine American University of Beirut Medical Center, Room C-662 Riad El-Solh, Beirut 1107 2020, Lebanon.