

NEURAL NETWORK CLASSIFICATION OF EEG SIGNALS BY USING AR WITH MLE PREPROCESSING FOR EPILEPTIC SEIZURE DETECTION

Abdulhamit Subasi^a, M. Kemal Kiymik^{a*}, Ahmet Alkan^a, Etem Koklukaya^b
e-mail : {asubasi, mkemal, aalkan}@ksu.edu.tr; ekaya@sau.edu.tr

^aDepartment of Electrical and Electronics Engineering, Kahramanmaraş Sütçü İmam University, 46100 Kahramanmaraş, Turkey.

^b Department of Electrical and Electronics Engineering, Sakarya University
54187 Sakarya, Turkey.

Abstract-The purpose of the work described in this paper is to investigate the use of autoregressive (AR) model by using maximum likelihood estimation (MLE) also interpretation and performance of this method to extract classifiable features from human electroencephalogram (EEG) by using Artificial Neural Networks (ANNs). ANNs are evaluated for accuracy, specificity, and sensitivity on classification of each patient into the correct two-group categorization: epileptic seizure or non-epileptic seizure. It is observed that, ANN classification of EEG signals with AR gives better results and these results can also be used for detecting epileptic seizure.

Keywords- EEG, Autoregressive method (AR), Maximum likelihood estimation (MLE), Artificial Neural Networks (ANN).

1. INTRODUCTION

EEG signals involve a great deal of information about the function of the brain. But classification and evaluation of these signals are limited. Since there is no definite criterion evaluated by the experts, visual analysis of EEG signals is insufficient. Since routine clinical diagnosis needs to analysis of EEG signals, some automation and computer techniques have been used for this aim. Since the early days of automatic EEG processing, representations based on a Fourier transform have been most commonly applied. This approach is based on earlier observations that the EEG spectrum contains some characteristic waveforms that fall primarily within four frequency bands— delta (< 4 Hz), theta (4–8 Hz), alpha (8–14 Hz), and beta (14–30 Hz). Such methods have proved beneficial for various EEG characterizations, but fast Fourier transform (FFT), suffer from large noise sensitivity. Parametric power spectrum estimation methods such as AR, reduces the spectral loss problems and gives better frequency resolution. Also AR method has an advantage over FFT that, it needs shorter duration data records than FFT [1,2]. Also it is faster than Continuous Wavelet transform techniques, especially in real time applications [18].

Numerous other techniques from the theory of signal analysis have been used to obtain representations and extract the features of interest for classification purposes. Neural networks and statistical pattern recognition methods have been applied to EEG analysis. Over the past two decades much research has been done with the use of conventional

temporal and frequency analyses measures in the detection of epileptic form activity in EEGs and comparatively good results have been obtained [1-6].

Neural network detection systems have been proposed by a number of researchers [7-23]. Pradhan[9] uses the raw EEG as an input to a neural network while Weng [8] uses the features proposed by Gotman [6] with an adaptive structure neural network, but his results show a poor false detection rate. Petrosian, et al., [12] showed that the ability of specifically designed and trained recurrent neural networks (RNN), combined with wavelet preprocessing, to predict the onset of epileptic seizures both on scalp and intracranial recordings. Esteller [23] uses linear predictor to find AR coefficients as an input to ANN. In this study we used maximum likelihood estimation (MLE) to find AR parameters and after that we used these AR coefficients as an input to ANN.

The purpose of the work described in this paper is to investigate the practicality of using an AR model by using MLE to extract classifiable features from human EEG. The success of this study depends on finding a signal representation contains the information needed to accurately classify epileptic seizure. Here, AR model with MLE was used to define representations. Various feature based on this model was classified with a multilayer, feedforward, neural network using the error back-propagation training algorithm. Discrimination was performed between a single pair of tasks. An AR with MLE representation resulted in the better classification percentages than FFT representation.

2. MATERIALS AND METHOD

2.1. EEG Data Acquisition and Representation

Epileptic seizure is an abnormality in EEG recordings and is characterized by brief and episodic neuronal synchronous discharges with dramatically increased amplitude. This anomalous synchrony may occur in the brain locally (partial seizures) which is seen only in a few channels of the EEG signal, or involving the whole brain (generalized seizures) which is seen in every channel of the EEG signal.

Subjects in different age group were recruited for this study. They were known epileptics with uncontrolled seizures and were admitted to the neurology department of the Medical Faculty Hospital of Dicle University. These signals belong to several healthy and unhealthy (epileptic patients) persons. The signals are collected by a data acquisition system which contains data acquisition card (PCI MIO-16-E+ type), signal processors and a personnel computer. Data can be taken in to computer memory quickly by using this card which is connected to PCI data bus of the computer. For this system LabVIEW programming language was used. The system provides real time data processing.

EEG signals are analyzed by using spectral analysis methods to diagnose some cerebral diseases. The power spectral density of the signal $P(f)$ found by applying

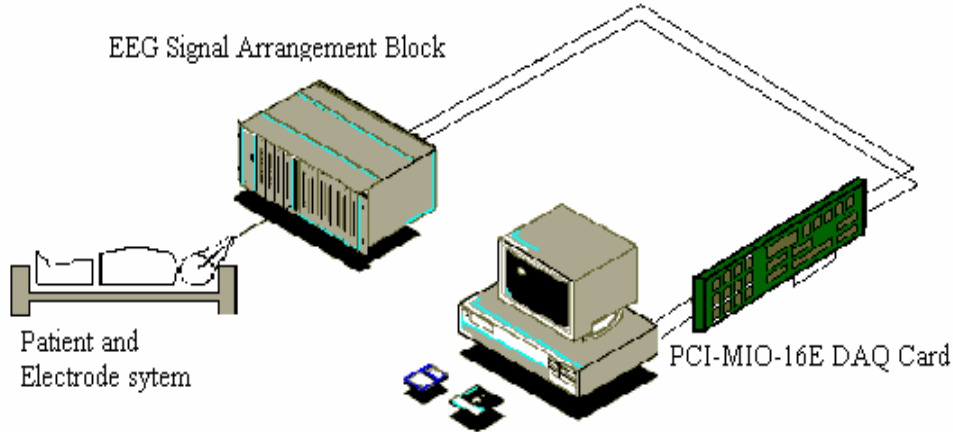


Figure. 1. The scheme of the EEG data acquisition system.

conventional and modern spectral analysis methods such as FFT and AR. The data acquisition system for the processing of EEG signals is shown in Fig. 1.

2.2. Autoregressive parameter estimation and MLE

In the AR model, to find out model parameters Levinson-Durbin algorithm which makes use of the solution of the Yule-Walker equations is used. Autocorrelation estimation is used for the solution of these equations. After those autocorrelation, AR model parameters are estimated. To do that biased form of the autocorrelation estimation is used which is given as

$$r_{xx} = \frac{1}{N} \sum_{n=0}^{N-m-1} x^*(n)x(n+m), \dots, m \geq 0 \quad (1)$$

The aim now is to estimate the AR model parameters by using MLE in the solution of the Yule-Walker equations from a record of EEG data. If the maximum likelihood estimate of a parameter exists under regular conditions, it is consistent, asymptotically unbiased, efficient, and normally distributed. Unfortunately, the maximum likelihood (ML) estimator is often too cumbersome to obtain. As this is the case for the EEG model, it is proposed to estimate the model parameters by maximizing an approximation of the log-likelihood function, known as Whittle's approximation, the derived estimator is expected to retain the properties associated with the ML estimator in an asymptotic sense, but with much less complexity. In fact, Whittle's estimate asymptotically retains the properties of the ML estimate for Gaussian random processes, but this is not generally true for the non-Gaussian case [2].

In many cases it is difficult to evaluate the MLE of the parameter whose power spectrum density function (PSDF) is Gaussian due to the need to invert a large dimension

covariance matrix. For example, if $x \sim N(0, c(\theta))$, the MLE of θ is obtained by maximizing

$$P(x; \theta) = \frac{1}{(2\pi)^{n/2} \cdot \det(c(\theta))} e^{[-1/2 X^t \cdot c^{-1}(\theta) \cdot X]} \quad (2)$$

If the covariance matrix cannot be inverted in closed form, then a search technique will require inversion of the $N \times N$ matrix for each value of θ to be searched. An alternative approximate method can be applied when x is data from a zero mean random process, so that covariance matrix is Toeplitz. In such a case, the asymptotic log-likelihood function is given by

$$\ln P(x; \theta) = \frac{N}{2} \ln 2\pi - \frac{N}{2} \int_{-1/2}^{1/2} \left[\ln P_{xx}(f) + \frac{I(f)}{P_{xx}(f)} \right] df \quad (3)$$

where

$$I(f) = \frac{1}{N} \left| \sum_{n=0}^{N-1} x(n) \cdot e^{-j2\pi f n} \right|^2$$

is the periodogram of the data and $P_{xx}(f)$ is the power spectral density (PSD). The dependence of the log-likelihood function on θ is through the PSD. Differentiation of (3) produces the necessary conditions for MLE

$$\frac{\partial \ln P(x; \theta)}{\partial \theta_i} = -\frac{N}{2} \int_{-1/2}^{1/2} \left[\frac{1}{P_{xx}(f)} - \frac{I(f)}{P_{xx}^2(f)} \right] \frac{\partial P_{xx}(f)}{\partial \theta_i} df \quad (4)$$

or

$$\int_{-1/2}^{1/2} \left[\frac{1}{P_{xx}(f)} - \frac{I(f)}{P_{xx}^2(f)} \right] \frac{\partial P_{xx}(f)}{\partial \theta_i} df = 0 \quad (5)$$

The second derivative allows the Newton-Raphson or scoring method may be implemented using the asymptotic likelihood function. This leads to simpler iterative procedures and is commonly used in practice.

In this study, to find MLE asymptotic form of the log-likelihood given by (3) is used. Since the PSD is

$$P_{xx}(f) = \frac{\delta_u^2}{|A(f)|^2} \quad (6)$$

After some calculations and derivations, the estimated auto correlation function is,

$$\hat{R}_{xx}(k) = \begin{cases} \frac{1}{N} \sum_{n=0}^{N-1-|k|} x(n)x(n+|k|) \dots |k| \leq N-1 \\ 0 \dots \dots \dots |k| \geq N \end{cases} \quad (7)$$

and the set of equations to be solved for the approximate MLE of the AR filter parameters becomes

$$\sum_{l=1}^p \hat{a}(l) \hat{R}_{xx}(k-l) = -\hat{R}_{xx}(k) \dots k=1,2,\dots,p \quad (8)$$

or in matrix form

$$\begin{bmatrix} \hat{R}_{xx}(0) & \hat{R}_{xx}(1) & \dots & \hat{R}_{xx}(p-1) \\ \hat{R}_{xx}(1) & \dots & \dots & \hat{R}_{xx}(1) \\ \dots & \dots & \dots & \dots \\ \hat{R}_{xx}(p-1) & \hat{R}_{xx}(p-2) & \dots & \hat{R}_{xx}(0) \end{bmatrix} \begin{bmatrix} \hat{a}(1) \\ \hat{a}(2) \\ \dots \\ \hat{a}(p) \end{bmatrix} = - \begin{bmatrix} \hat{R}_{xx}(1) \\ \hat{R}_{xx}(2) \\ \dots \\ \hat{R}_{xx}(p) \end{bmatrix}. \quad (9)$$

These are so-called estimated Yule-Walker equations and this is the autocorrelation method of linear prediction. Note that the special form of the matrix and the right-hand vector, which thereby allow a recursive solution known as the Levinson recursion [1]. To complete the discussion explicit form for the MLE of δ_u^2 must be determined. From (6)

$$\hat{\delta}_u^2 = \sum_{k=0}^p \hat{a}[k] \hat{R}_{xx}[-k] = \sum_{k=0}^p \hat{a}[k] \hat{R}_{xx}[k] \quad (10)$$

These parameters are used to compute AR spectral power spectral density as

$$P_{xx}^{YW}(f) = \frac{\hat{\delta}_{wp}^2}{\left| 1 + \sum_{k=1}^p \hat{a}(k) e^{-j2\pi f k} \right|^2} \quad (11)$$

In the AR modeling method, the order of the model, namely, the filter, is depend on the number of AR coefficients. In the AR method, the model order is identified according to different criteria. The selection of the model order in AR spectral estimation is a critical subject. Too low order results in a smoothed estimate, while too large order causes spurious peaks and general statistical instability. AR spectral estimator offers the promise of higher resolution. Their principal shortcomings are that in the case of AR spectral estimation, if the assumed model or the dimension of the autocorrelation matrix is inappropriate and if the model orders chosen incorrect, then poor spectral estimates will result. Heavy biases and/or large variability may be exhibited. In this study, Akaike information criteria (AIC) is taken as the base for choosing the model order. According to AIC, model order $p=6$ was taken because the determined model order was lower. In our AR model, MLE is used for the solution of the Yule-Walker equations to get AR model parameters[1]. We used 5 AR coefficient over time for a 6th order model. AR Coefficients over time for seizure-records from epileptic patient are shown in Fig. 2.

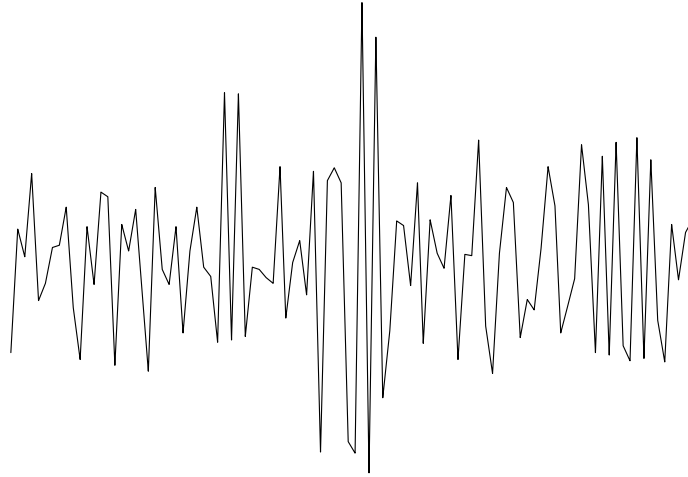


Figure 2: AR Coefficients over time for seizure-records from epileptic patient for a 6th order model.

2.3. Spectral analysis of EEG signals

In the FFT spectral analysis of the EEG signals some spurious frequencies are seen comparing with AR with MLE spectrum. The AR with MLE offers a good quality spectrum output in terms of frequency resolution. In Fig. 3 an epileptic EEG signal and FFT of this signal are given. If frequency spectrum of FFT is examined, it is seen that there are peaks at 1 Hz and 3 Hz. AR spectrum of the same signal is presented in Fig. 4. There are peaks at 3 Hz with higher amplitude, 6 Hz, 9.5 Hz and 13.5 Hz. When we compare these two spectrums it is seen that AR spectrum has got sharper peaks and less misleading peaks than FFT. Due to this better frequency solution, explanation and determination of the activities in the signal is easier by using AR method. Since the signal is taken from an epileptic patient, the results fit with the typical characteristics of epilepsy that is delta activity (low frequency range) [1].

Fig. 5 shows a normal EEG signal and FFT spectrum of this signal. Spectrum of the AR with MLE is given in Fig. 6. If these two spectrums are examined although FFT spectrum has got wide and misleading peaks AR spectrum has got sharp and clear peaks. If these spectrums are examined, delta activity, alpha activity, and beta activity can be seen easily. These results are true because it is a normal EEG signal. Higher variations and misleading peaks in FFT spectrum avoid the dominant alpha and delta activities.

As seen in these figures, the power spectrum obtained by using the FFT method does not have good frequency resolution. In addition, some misleading frequencies are seen on the FFT method's spectrums comparing with AR methods' spectrums. In the case of

nonepileptic EEG, lower frequencies are not clear due to misleading frequencies. EEG spectrums obtained by AR method are clearer and have higher spectral resolution compared with FFT. FFT is an inconsistent spectral estimator which continues to fluctuate around the true PSD. Often, the spectral leakage masks weak signals that are present in the data. Smearing and spectral leakage are particularly critical for spectra with large amplitude ranges, such as peaky spectra. Since, power spectrum obtained by using FFT is not clear, AR with MLE gives better performance for spectral resolution than FFT.

As a result, when AR with MLE approach is compared for their resolution and interpretation performance, it is determined that the AR approach is better for the use in ANN as preprocessing method in clinical and research areas, because of the clear spectra, which are obtained by it.

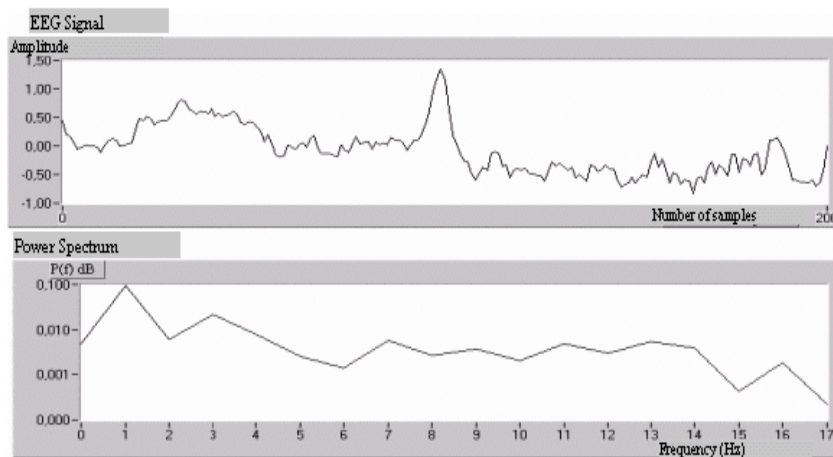


Figure 3. Epileptic EEG signal and its FFT.

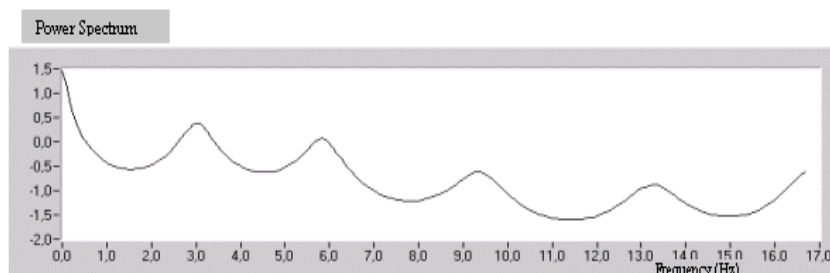


Figure 4. AR Spectrum of epileptic EEG signal.

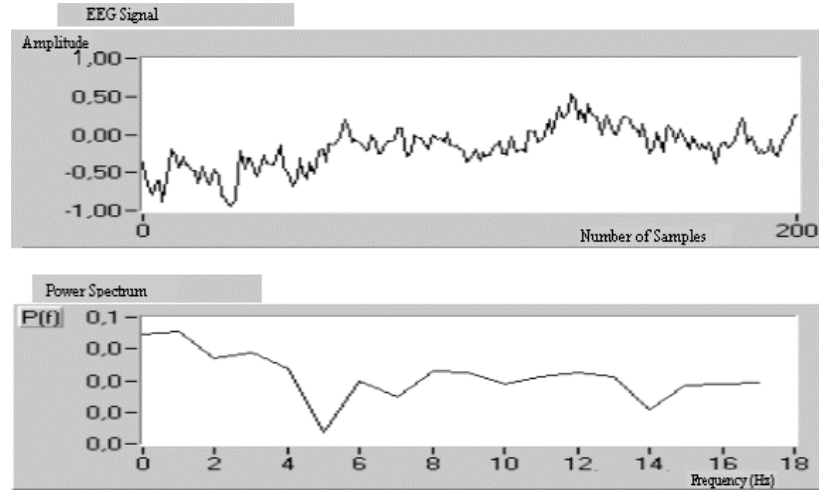


Figure 5. Normal EEG signal and its FFT.

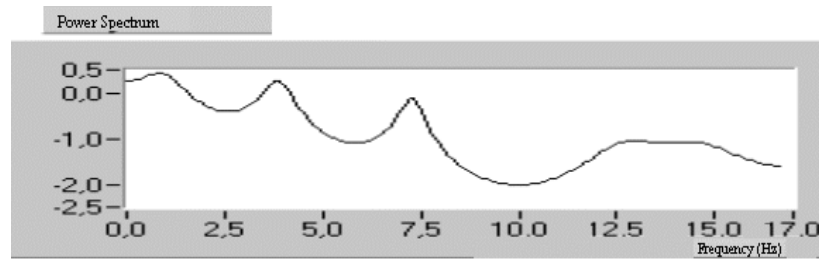


Figure 6. AR Spectrum of normal EEG signal

2.4. Artificial Neural Network Classifier

ANNs consist of a great number of processing elements (neurons), which are connected with each other; the strengths of the connections are called weights. For the modeling of physical systems, a feedforward multilayered neural network is commonly used. It consists of a layer of input neurons, a layer of output neurons and one or more hidden layers. In order to cope with nonlinearly separable problems, additional layer(s) of neurons placed between the input layer (containing input nodes) and the output neuron are needed leading to the MLP architecture, as shown in Fig. 7. We used five input neurons for AR and 18 input neuron for FFT. The network topology is the standard feedforward network with a single hidden layer. The number of output neuron is two, and the number of hidden unit neuron is chosen as 20. Since the intermediate layers do not interact with the external environment, they are called hidden layers and their nodes called hidden nodes. The addition of intermediate layers revived the perceptron by extending its ability to solve nonlinear classification problems. In ANNs, the knowledge lies in the interconnection weights between neurons. Therefore, training process is an important characteristic of the ANN methodology, whereby representative examples of the knowledge are iteratively

presented to the network, so that it can integrate this knowledge within its structure. No assumption is needed about the underlying data probability distributions when ANN is used for pattern classification. Once trained, it can be configured to perform adaptively to improve its performance over time [19-23].

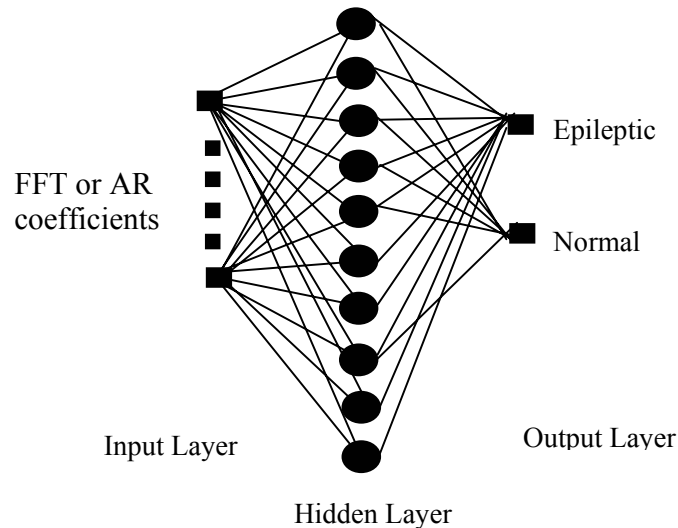


Figure 7. Artificial neural network architecture.

Although rules for neural network optimization are under development, neural network architectures are derived by trial and error. The determination of appropriate number of hidden layers is one of the most critical tasks in neural network design. Unlike the input and output layers, one starts with no prior knowledge as to the number of hidden layers. A network with too few hidden nodes would be incapable of differentiating between complex patterns leading to only a linear estimate of the actual trend. In contrast, if the network has too many hidden nodes it will follow the noise in the data due to overparameterization leading to poor generalization for untrained data. With increasing number of hidden layers, training becomes excessively time-consuming. The most popular approach to finding the optimal number of hidden layers is by trial and error [22]. In the present study, MLP network consisted of one input layer, one hidden layer, and one output layer and the decision about the number of hidden layers in use was given as empirically.

MLP ANN was used because it is appropriate for solving pattern classification problems where supervised learning is implemented with backpropagation algorithm. The advantage of using this type of ANN is the rapid execution of the trained network, which is particularly advantageous in signal processing applications [20-22].

In most applications of MLP, the weights are determined by means of the backpropagation algorithm, which is based on searching an error surface (error as a function of ANN weights) using gradient descent for points with minimum error [20-22]. During the training phase, the weights are successively adjusted based on a set of inputs and the corresponding set of desired output targets. Each iteration in backpropagation

constitutes two sweeps: forward activation to produce a solution, and a backward propagation of the computed error to modify the weights. The forward and backward sweeps are performed repeatedly until the ANN solution agrees with the desired value within a prespecified tolerance. The backpropagation algorithm provides the needed weight adjustments in the backward sweep [21, 22]. The backpropagation algorithm is a nonlinear procedure because of the nonlinear threshold element contained in each node, and its behavior is very complex because of the layered structure. However, this nonlinear behavior allows a perceptron to generate highly complex decision regions, which is a desirable property for pattern classification [21]. MLP neural networks employing backpropagation training algorithm are so versatile and can be used for signal processing, image compression, pattern recognition, medical diagnosis, prediction, classification, nonlinear system modeling, and control [19, 20]. Since backpropagation training algorithm has rapid execution and has widely used in pattern classification problems, MLP neural network employing backpropagation training algorithm was used to predict the presence or absence of epileptic seizure.

A fully-connected network is employed and the standard backpropagation algorithm with momentum and adaptive learning rate employed, with parameters as in Table 1. Learning coefficients and momentum values for different number of iterations are given in Table 2. With regard to the adaptive learning rate, the rate is increased (by a factor of LRI) following an improvement in the SSE, but if the error ratio exceeds MER, the learning rate is decreased by a factor of LRD as shown in Fig. 8. Such an adaptation speeds up convergence considerably. We performed the following cross-validation procedure for training the network as a way to control the over-fitting of training data. We randomly select 60% of the data set for training the network and 20% of the data for validation after each training epoch. The error of the network on the validation data is calculated after every pass, or epoch, through the training data. This best network is then applied to the remaining 20% of the data, referred to as the test set. All representations were classified using different random selections of train, validation, and test sets and initial weight values [13].

Table 1. Neural network training parameters

Sum squared error (SSE) goal	10^{-5}
Initial learning rate coefficient	0.01
Learning rate increase (LRI)	1.05
Learning rate decrease (LRD)	0.7
Momentum coefficient	0.95
Maximum error ratio (MER)	1.04

Table 2. Learning Coefficients and Momentum Values for Different Number of Iterations

Iteration	Learning coefficient	Momentum
2,000	0.001	0.02
4,000	0.0035	0.0042
8,000	0.00001	0.00001

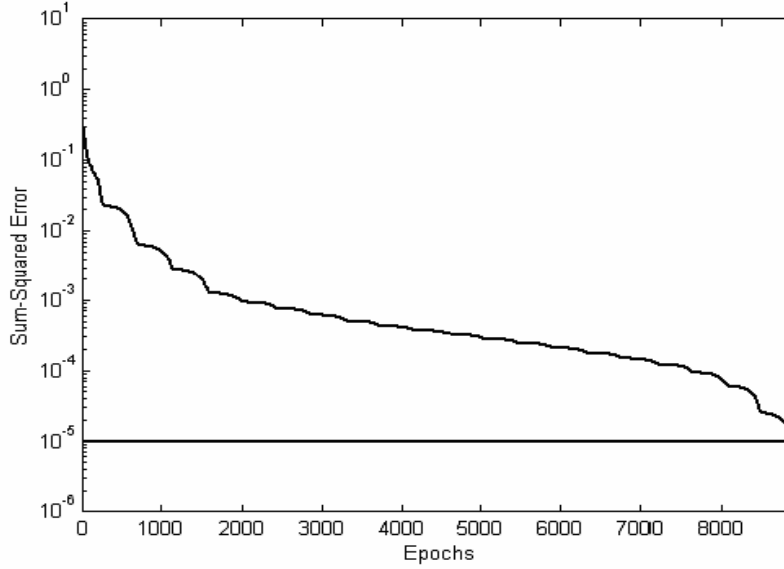


Figure 8. Graphics of Sum-Squared Error vs. epochs

3. RESULTS AND DISCUSSION

In this study we used FFT and AR with MLE parameters used as an input to ANN. ANN is trained with data preprocessed by FFT and AR with MLE. Potential confusions that the classifier can be identified by the relatively high responses of an output unit for test segments that do not correspond to the task represented by that output unit. For this trial, averaging over 20 segments results in 92% correct, but performance is not improved this much on all trials. The best classification performance for the network, averaged over all 30 repetitions, is achieved by averaging over all segments.

Results for the baseline performance of the best performing backpropagation neural network using AR with MLE and FFT are shown in Table 3. Accuracy is the total percentage of correct predictions. Specificity is the percentage of correct predictions for patients that had epileptic seizures during the EEG and sensitivity is the percentage of correct predictions for patients that did not have epileptic seizures.

We achieved a classification rate of 92.3% by using a neural network with a single hidden unit as a classifier. The results are averages over 30 runs, each run with different combinations of train, test, and validate sets and different initial weight values. The validation and test sets each contained 20% of the total number of feature vectors, with the rest in the training set. The classification rates are the 90% confidence intervals. Although the error on the training data decreases as the number of hidden unit increases, the error on the test data does not change significantly with the number of hidden units. This suggests that we can not do much better than a linear classifier with this representation. The classification percentages of AR with MLE on test data are above 92%. An average of 91% classification is achieved by using FFT as preprocessing in the neural net.

Table 3. Seizure classification performance of ANN with AR and FFT.

Preprocessing method	Accuracy	Specifity	Sensitivity
AR with MLE	92.3	96.2	90.3
FFT	91.6	93.6	89.8

From Table 3, the backpropagation artificial neural network outperforms on both overall accuracy and sensitivity. One can see when looking at the averages across subjects that the AR with MLE gives the best classification accuracies at 92.3%, but not by much. The FFT preprocessing performs slightly worse at 91%. The AR with MLE model is better at specificity, because it classifies almost 96.2 percent of the patients as epileptic seizure patients.

Our results show that the AR with MLE is the most consistent feature vector. However, if in the future many subjects are to be tested and computation time is an issue, FFT preprocessing appears to be the best choice.

4. CONCLUSION

In this study the FFT and AR spectrums which have Maximum Likelihood Estimation (MLE) optimization of epileptic and normal EEG signals are used as an input to an artificial neural network that could be used to discriminate between the two tasks with greater than 92% accuracy. To get AR method model parameters, MLE which has wide applications in statistics is used. This type of result might be marginally acceptable for a real-time system based on two commands for some subjects. As compared to methods such as FFT, it is seen that the classification accuracy of AR with MLE is better; but when the processing speed or time is considered FFT may be more suitable. If we compare to other methods (such as wavelet etc.) speed of the AR with MLE is better and classification is acceptable [18]. Cluster analysis was applied to learned weight vectors, revealing some of the acquired relationships between representation components and mental tasks. The results of clustering can be used both for the construction of lower-dimensional representations and for investigating hypotheses regarding differences in brain activity related to different cognitive behaviour. One of the strengths of this study is its rigorous training procedure involving cross-validation, early stopping, and a large number of training repetitions. Early stopping is one of the simplest methods for limiting the complexity of a network. The most likely route to better performance is to test other EEG signal representations. We have presented a method for the automatic classification of seizures. The system performance would be improved by replacing the linear AR predictive model with a nonlinear model and using the coefficients of the nonlinear model as the signal representation.

REFERENCES

1. I.Guler , M.K.Kiymik, M.Akin and A. Alkan, AR spectral analysis of EEG signals by using maximum likelihood estimation, *Comp in Biology and Medicine*; **31**, 441-450, 2001.
2. M. Akin, and M.K. Kiymik, Application of Periodogram and AR Spectral Analysis to EEG Signals, *Journal of Medical Systems* **24(4)**, 247-256, 2000.
3. J.Gotman and L.Wang, State-dependent spike detection: concepts and preliminary results, *Elec and Clinical Neurophysiology* **79**, 11-19, 1991.
4. Y.S. Tseng, R.C.Chen, F. C. Chong and T.S. Kuo, Evaluation of parametric methods in EEG signal analysis, *Med. Eng. & Phy.* **17**, 71-78, 1995.
5. S. Goto, M. Nakamura and K. Uosaki, On-line spectral estimation of nonstationary time series based on ar model parameter estimation and order selection with a forgetting factor, *IEEE Trans on Signal Processing* **43**, 1519-1522, 1995.
6. J. Gotman, Automatic recognition of epileptic seizures in the EEG, *Electroencephalography and Clinical Neurophysiology* **54**, 530-540, 1982.
7. B. O. Peters, G. Pfurtscheller, and H. Flyvbjerg, Mining multi-channel EEG for its information content: an ANN-based method for a brain-computer interface, *Neural Networks* **11**, 1429-1433, 1998.
8. W. Weng and K. Khorasani, An adaptive structure neural network with application to EEG automatic seizure detection, *Neural Networks* **9**, 1223-1240, 1996.
9. N. Pradhan, P.K. Sadasivan and G. R. Arunodaya, Detection of seizure activity in EEG by an artificial neural network: A preliminary study, *Computers and Biomedical Research* **29**, 303-313, 1996.
10. I. Osario, M. G. Frei and S. B. Wilkinson, Real-time automated detection and quantitative analysis of seizures and short-term prediction of clinical onset, *Epilepsia* **39**, 615-627, 1998.
11. H. Qu and J. Gotman, A Patient-Specific Algorithm for the Detection of Seizure Onset in Long-Term EEG Monitoring: Possible Use as a Warning Device, *IEEE Trans on Biomedical Eng.* **44**, 115-122, 1997.
12. A. Petrosian, D. Prokhorov, R. Homan, R. Dashei and D.Wunsch, Recurrent neural network based prediction of epileptic seizures in intra- and extracranial EEG, *Neurocomputing* **30**, 201-218, 2000.
13. C.W. Anderson, S.V. Devulapalli and, A.E. Stolz, *Neural Networks for Signal Processing*, IEEE Service Center, Piscataway, NJ, 1995.
14. S. Walczak and W. J. Nowack, An Artificial Neural Network Approach to Diagnosing Epilepsy Using Lateralized Bursts of Theta EEGs, *Journal of Medical Systems* **25**, 9-20, 2001.
15. T. Shimada, T. Shiina, and Y. Saito, Detection of characteristic waves of sleep EEG by neural network analysis, *IEEE Trans. on Biomedical Eng.* **47**, 369-379, 2000.
16. P.Wahlberg and L.Lantz, Methods for robust clustering of epileptic EEG spikes, *IEEE Trans on Biomedical Eng.* **47**, 857-868, 2000.

17. M. Sun and R.J.Sclabassi, The forward EEG solutions can be computed using Artificial neural networks, *IEEE Trans on Biomedical Eng.* **47**, 1044-1050, 2000.
18. M.K. Kiymik, *Realization of real time spectral analysis of EEG signals with parametric methods*, Project No:197E014 (EEEAG-249), The Scientific and Technical Research Council of Turkey, 2001.
19. W.G. Baxt, Application of artificial neural networks to clinical medicine, *Lancet* **346**, 1135–1138, 1995.
20. L. Fausett, *Fundamentals of neural networks architectures, algorithms, and applications*, Prentice Hall, Englewood Cliffs, NJ, 1994.
21. S. Haykin, *Neural networks: A comprehensive foundation*, Macmillan, New York, 1994.
22. A.S. Miller, B.H. Blott and T.K. Hames, Review of neural network applications in medical imaging and signal processing. *Medical and Biological Engineering and Computing* **30**, 449–464, 1992.
23. R. Esteller, *Detection of seizure onset in epileptic patients from intracranial EEG signals*, Ph.D dissertation, Georgia Institute of Technology, April 2000.