

PERFORMANCE ANALYSIS OF PATIENT SPECIFIC ELMAN-CHAOTIC OPTIMIZATION MODEL FOR FUZZY BASED EPILEPSY RISK LEVEL CLASSIFICATION FROM EEG SIGNALS

Dr.R.HariKumar¹, T.Vijayakumar²

¹Professor, Department of ECE, Bannari Amman Institute of Technology Sathyamangalam

²Lecturer, Department of IT, Bannari Amman Institute of Technology Sathyamangalam

Abstract- This paper aims to analyze the optimization of Epilepsy risk levels from EEG signals using Fuzzy based Elman-Chaotic Optimization. The EEG (Electroencephalogram) signals of twenty patients are collected from Sri Ramakrishna Hospitals at Coimbatore. The raw EEG signals are sampled and various parameters like energy, energy, variance, peaks, sharp and spike waves, duration, events and covariance. The fuzzy techniques are applied as a first level classifier to classify the risk levels of epilepsy by converting the EEG signal parameters in to code patterns by fuzzy systems. Elman-Chaotic optimization is identified as post classifiers on the classified data to obtain the optimized risk level that characterizes the patient's epilepsy risk level. This classification provides a better way of treating the epileptic patients. This project aims to safeguard a patient's life when critical situation occurs. Future scope is to design an embedded system which collects the raw EEG signals from the brain and directly gives the level of epilepsy. It will make the neural surgeons to give appropriate remedial measures.

Keywords- EEG Signals, Epilepsy risk levels, Fuzzy Logic, Chaotic Optimization, Elman Neural Network

I INTRODUCTION

Epileptic seizures are result of the transient and unexpected electrical disturbance of the brain. Unfortunately, the occurrence of an epileptic seizure seems unpredictable and its process is very

little understood [1]. Twenty –five percent of the world’s 50 million people with epilepsy have seizures that cannot be controlled by any available treatment [2]. Electroencephalogram (EEG) as a representative signal containing information of the electrical activity generated by the cerebral cortex nerve cells, has been the most utilized signal to clinically assess brain activities, and the detection of epileptiform discharges in the EEG is an important component in the diagnosis of epilepsy. The detection of epilepsy, which includes visual scanning of EEG recordings for the spikes and seizures, is very time consuming, especially in the case of long recordings. In addition, bio-signals are highly subjective so disagreement on the same record is possible, so the EEG signal parameters extracted and analyzed using computers, are highly useful in diagnostics.

A. General Techniques

The early methods of automatic EEG processing were based on a Fourier transform[36]. This approach is based on earlier observations that the EEG spectrum contains some characteristic waveforms that fall primarily within four frequency bands. Such methods have proved beneficial for various EEG characterizations, but fast Fourier Transform (FFT), suffers from large noise sensitivity. Parametric methods for power spectrum estimation such as autoregressive (AR), reduce the spectral loss problems and give better frequency resolution. Since the EEG signals are non stationary, the parametric methods are not suitable for frequency decomposition of these signals [37]. Chaos" is a tricky thing to define. In fact, it is much easier to list properties that a system described as "chaotic" has rather than to give a precise definition of chaos. A dynamical system displaying sensitive dependence on initial conditions on a closed invariant set will be called chaotic [31].

By a chaotic solution to a deterministic equation we mean a solution whose outcome is very sensitive to initial conditions (i.e., small changes in initial conditions lead to great differences in outcome) and whose evolution through phase space appears to be quite random. Its combination of novel mathematics and high speed computing, has produced new insights into the behavior of complex systems and reveals surprising results even in the simplest non-linear models. Non-linear systems are characterized by having "bifurcation-points", regions where the system sits on a knife edge, as it were, and may suddenly change its qualitative behavior. Systems sometimes enter regions of highly erratic and chaotic behavior. In such cases it becomes impossible to predict the future behavior of the system even when based on its entire past history. From

moment to moment the system jumps violently in its behavior; moreover, it may be infinitely sensitive to any external change of fluctuation. One such non-linear signal is EEG (Electroencephalogram) non-linear dynamics theory opens new window for understanding the behavior of EEG. In the analysis of EEG data, different chaotic measures are used in recent literature [32]. Jing and Takigawa [33] applied correlation dimensions techniques to analyze EEG at different neurological states. Lehnertz and Elger [34] used correlation dimension technique to test whether a relationship exists between spatio-temporal alterations of neural complexity and spatial extent and temporal dynamics of the epileptogenic area. Casdagial et al [35] showed that the techniques developed for the study of non-linear systems could be used to characterize the epileptogenic regions of the brain during interictal period. Correlation integral, the measures sensitive to a wide variety of non-linearities, was used for detection. Between seizures, the EEG of a patient with epilepsy may be characterized by occasional epileptic form transients-spikes and sharp waves. EEG patterns have shown to be modified by a wide range of variables including biochemical, metabolic, circulatory, hormonal, neuro electric and behavioral factors [4].

Exploring various analytical approaches, both linear and non linear methods to process data from medical database is meaningful before deciding on the tool that will be most useful, accurate, and relevant for practitioners. For example, assigning a new patient to a particular outcome class is a classification problem commonly described as “pattern recognition”, “discriminant analysis”, and “supervised learning” [14]. In the past, the Encephalographer, by visual inspection was able to qualitatively distinguish normal EEG activity from localized or generalized abnormalities contained within relatively long EEG records. The different types of epileptic seizures are characterized by different EEG waveform patterns. With real-time monitoring to detect epileptic seizures gaining widespread recognition, the advent of computers has made it possible to effectively apply a host of methods to quantify the changes occurring based on the EEG signals [5]. One of them is a classification of risk level of epilepsy by using Fuzzy techniques [9]. The recognition of specific waveforms and features in the Electroencephalogram (EEG) for classification of epilepsy risk levels has been the subject of much research.

B. Overview of Fuzzy Based Epilepsy Risk Level Classifier

The early methods of automatic EEG processing were based on Fourier transform. We cannot predict the optimized epilepsy risk level based on the fuzzy outputs which is a first level classifier. Elman-Chaotic optimization is implemented as post classifiers in optimizing the epileptic risk level of the patient classified by the fuzzy system. We also present a comparison of these methods based on their Performance Indices and Quality Values.

The block diagram of epilepsy classifier is shown in figure1. This is accomplished as:

1. Fuzzy classification for epilepsy risk level at each channel from EEG signals and its parameters.
2. Each channel results are optimized, since they are at different risk levels.
3. Performance of fuzzy classification before and after the Elman-Chaotic optimization methods is analyzed

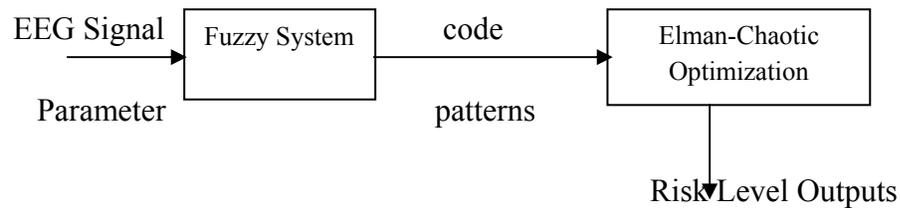


Figure 1 Elman-Chaotic- Fuzzy Classification System

The following tasks are carried out to classify the risk levels by Elman-Chaotic Optimization which are,

1. Converting the fuzzy encoded values in to numerical values for Elman-Chaotic Optimization (say a).
2. Introducing the numerical values in to the Poincare Equation for obtaining

$$b=A.*a.*(1-a).*(1-(2.*a)).^2;$$

$$c=A.*b.*(1-b).*(1-(2.*b)).^2;$$

$$d=A.*c.*(1-c).*(1-(2.*c)).^2;$$

3. With These values we can obtain difference values by using

$$A(i)=b(i)-b(i+1);$$

$$B(i)=c(i)-c(i+1);$$

$$C(i)=d(i)-d(i+1);$$

4. With the help of these difference results, a plot is drawn between A, B, C. Spherical CTM (discussed later) of the input parameters can be calculated from the three dimensional difference plot, see fig 5.
5. The CTM values are trained with Elman neural network to provide perfect classification.

II MATERIALS AND METHODS

The EEG data used in the study were acquired from twenty epileptic patients who had been under the evaluation and treatment in the Neurology department of Sri Ramakrishna Hospital, Coimbatore, India. A paper record of 16 channel EEG data is acquired from a clinical EEG monitoring system through 10-20 international electrode placing method. The EEG signal was band pass filtered between 0.5 Hz and 50Hz using five pole analog Butterworth filters to remove the artifacts. With an EEG signal free of artifacts, a reasonably accurate detection of epilepsy is possible; however, difficulties arise with artifacts. This problem increases the number of false detection that commonly plagues all classification systems. With the help of Neurologist we had selected artifact free EEG records with distinct features. These records were scanned by Umax 6696 scanner with a resolution of 600dpi.

A. EEG Data Acquisition and Preprocessing

Since the EEG records are over a continuous duration of about thirty seconds, they are divided into epochs of two second duration each by scanning into a bitmap image of size 400x100 pixels. A two second epoch is long enough to detect any significant changes in activity and presence of artifacts and also short enough to avoid any repetition or redundancy in the signal [6], [7],[10],[11].The EEG signal has a maximum frequency of 50Hz and so, each epoch is sampled at a frequency of 200Hz using graphics programming in C. Each sample corresponds to the instantaneous amplitude values of the signal, totaling 400 values for an epoch. The different parameters used for quantification of the EEG are computed using these amplitude values by suitable programming codes. The parameters are obtained for three different continuous epochs at

discrete times in order to locate variations and differences in the epileptic activity. We used ten EEG records for training and fifteen records for testing. These EEG records had an average length of six seconds and total length of 90 seconds. The patients had an average age of 31 years. A total of 720 epochs of 2 seconds duration are used.

B. Fuzzy System as Pre Classifier

The main objective of this research is to classify the epilepsy risk level of a patient from EEG signals

1. The energy in each two-second epoch is given by
$$E = \sum_{i=1}^n x_i^2 \quad (1)$$

Where x_i is signal sample value and n is number of samples. The normalized energy is taken by dividing the energy term by 1000.

2. The total number of positive and negative peaks exceeding a threshold is found.

3. Spikes are detected when the zero crossing duration of predominantly high amplitude peaks in the EEG waveform lies between 20 and 70 ms and sharp waves are detected when the duration lies between 70 and 200ms.

4. The total numbers of spike and sharp waves in an epoch are recorded as events.

5. The variance is computed as σ given by
$$\sigma^2 = \frac{\sum_{i=1}^n (x_i - \mu)^2}{n} \quad (2)$$

Where $\mu = \frac{\sum_{i=1}^n x_i}{n}$ is the average amplitude of the epoch.

6. The average duration is given by
$$D = \frac{\sum_{i=1}^p t_i}{p} \quad (3)$$

Where t_i is one peak to peak duration and p is the number of such durations.

7. Covariance of Duration: The variation of the average duration is defined by

$$CD = \frac{\sum_{i=1}^p (D - t_i)^2}{pD^2} \quad (4)$$

A sample value of extracted above seven features for the patient record 4 is shown in table 1.

Table.1 Average Values Of Extracted Parameters From Patient Record 4

Parameters	Epoch1	Epoch2	Epoch3
Energy	5.2869	8.581	10.10
Variance	1.1397	2.121	2.322
Peaks	9	38	35
Sharp & Spike	122	91	87
Events	185	154	145
Average duration	3.798	4.042	3.883
Covariance	0.5793	0.5123	0.5941

In the above abnormal case all the sixteen channels do not show high risk characteristics of EEG signals. There are certain regions (Channel IX & Channel XIII) which produce near normal features. Therefore it is indispensable to classify epilepsy risk level on channel basis using fuzzy techniques, since the parameter values are overlapping in between the normal and abnormal regions.

C. Fuzzy Membership functions

The energy is compared with the other six input features to give six outputs. Each input feature is classified into five fuzzy linguistic levels viz., *very low*, *low*, *medium*, *high* and *very high* [9]. The triangular membership functions are used for the linguistic levels of energy, peaks, variance events, spike and sharp waves, average duration and covariance of duration. The output risk level is classified into five linguistic levels namely *normal*, *low*, *medium*, *high* and *very high*.

D. Fuzzy Rule Set

Rules are framed in the format

IF Energy is low AND Variance is low THEN Output Risk Level is low

In this fuzzy system we have five linguistic levels of energy and five linguistic levels of other six features such as variance, peaks, events, spike and sharp waves, average duration and covariance of duration. Theoretically there may be 5^6 (that is 15625) rules are possible but we had considered the fuzzy pre -classifier as a combination of six two inputs and one output (2×1) system. With energy being a constant one input the other input is selected in sequential manner. This two inputs one output (2×1) fuzzy system works with 25 rules. We obtain a total rule base of 150 rules based on six sets of 25 rules each. This is a type of exhaustive fuzzy rule based system [11].

E. Risk Level Estimation in Fuzzy Outputs

The output of a fuzzy system represents a wide space of risk levels. This is due to sixteen different channels of input to the system in three epochs. This yields a total of forty-eight input output pairs. Since we deal with known cases of epileptic patients, it is indispensable to find the exact level of risk the patient. This will also aid in the development of automated systems that can precisely classify the risk level of the epileptic patient under observation. Hence an optimization of the outputs of the fuzzy system is initiated. This will improvise the classification of the patient’s state and can provide the EEGer with a clear picture. A specific coding method processes the output fuzzy values as individual code. Since working on definite alphabets is easier than processing numbers with large decimal accuracy, we encode the outputs as a string of alphabets. The alphabetical representation of the five classifications of the outputs is shown in table.2

Table.2 Representation Of Risk Level Classifications

Risk Level	Representation
Normal	U
Low	W
Medium	X
High	Y
Very High	Z

A sample output of the fuzzy system with actual patient readings is shown in fig. 2, for eight channels over three epochs. It can be seen that the Channel I shows low risk levels while channel VII shows high risk levels. Also, the risk level classification varies between adjacent epochs

Epoch 1	Epoch 2	Epoch 3
WYYWYY	WYYWYY	WZYYWW
YZZYXX	YYYYXX	YYYYXY
YYZXYY	YYYYYY	YYYYYY
YZZYXY	XZZXYY	YYYYYY
ZZZYYY	WYYYXX	YYYYXY
YYZXXX	WYZYYY	YZZYYY
ZZZYYY	YYYYYY	ZZZYYY
YYYYXX	YYYYXX	YYYXZY

Figure 2. Fuzzy logic Output

The fuzzy method's classification efficiency is evaluated from the following parameters. The Performance of Fuzzy method is defined as follows [3],

$$PI = \frac{PC - MC - FA}{PC} \times 100 \quad (5)$$

Where PC – Perfect Classification; MC – Missed Classification; FA – False Alarm

$$PI = [(0.5 - 0.2 - 0.1) / 0.5] * 100 = 40\%$$

The perfect classification represents when the physicians and fuzzy classifier agrees with the epilepsy risk level. Missed classification represents a true negative of fuzzy classifier in reference to the physician and shows High level as Low level. False alarm represents a false positive of fuzzy classifier in reference to the physician and shows Low level as High level. The performance for Fuzzy classifier is as low as 40%. The sensitivity is defined as [29]

$$S_e = [PC / (PC + FA)] * 100 \quad (6)$$

$$S_e = (0.5 / 0.6) * 100 = 83.33.5\%$$

The specificity is defined as [22]

$$S_p = [PC/PC+MC] * 100 \quad (7)$$

$$S_p = (0.5/0.7)*100=71.42\%$$

Due to the low value of performance index, sensitivity and specificity it is necessary to optimize the output of the fuzzy systems. Now we are about to identify the nonlinearities associated with fuzzy outputs in describing the epilepsy risk levels. The five risk levels are encoded as $Z>Y>X>W>U$ in binary strings of length five bits using weighted positional representation as shown in Table 3. Encoding each output risk level of the fuzzy output gives us a string of six codes (chromosomes), the value of which is calculated as the sum of probabilities of the individual codes. For example, if the output of an epoch is encoded as $ZZYXWZ$, its value would be 0.333331, [14]. Now the each input patterns are encoded in the numerical form of the range 0-1.

Table 3.Binary Representation Of Risk Levels

Risk Level	Code	Binary String	Weight	Probability
Very high	Z	10000	$16/31=0.51612$	0.086021
High	Y	01000	$8/31=0.25806$	0.043011
Medium	X	00100	$4/31=0.12903$	0.021505
Low	W	00010	$2/31=0.06451$	0.010752
Normal	U	00001	$1/31=0.03225$	0.005376
		11111=31	$\Sigma=1$	

Let the fuzzy outputs as shown in figure 2 is coded with appropriate numerical values. These numerical values are associated with the probability of each coded epilepsy risk level patterns. To illustrate the non linearity we have chosen the statistical measure of cross correlation between the

two adjacent epoch patterns. Thus the cross correlation function $r_{xy}(\mathbf{m})$ of the epochs $x(n)$ and $y(n)$ is defined by the equation (8) and assuming that both sequence have been measured from $n=0$ to $n=N-1$, in our case $n=1$ to 16 , [25]

$$r_{xy}(m) = \begin{cases} \frac{1}{N} \sum_{n=0}^{N-m-1} x(n+m)y(n), & \text{for } 0 \leq m \leq N-1 \\ \frac{1}{N} \sum_{n=0}^{N-|M|-1} x(n)y(n+M), & \text{for } -(N-1) \leq m \leq 0 \end{cases} \quad (8)$$

The cross correlation $r_{xy}(m)$ plot obtained through the equation (8) is shown in the “Fig.3”, which emulates the occurrence of highly non periodic patterns in the fuzzy outputs. Therefore any closed solution will be failed for this purpose of optimization. Hence, it is prudent to prefer non linear techniques instead of linear one, such a one is Chaotic optimization technique (post classifier) [15].

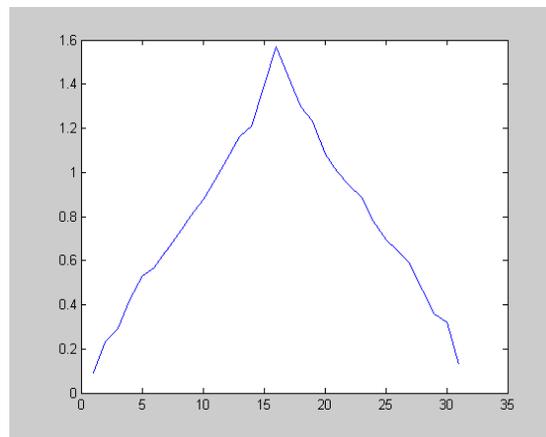


Figure.3 Cross Correlation Function plot for the Adjacent Epochs in fuzzy based Epilepsy Risk Level Outputs

III CHAOTIC OPTIMIZATION

One of the greatest attribute of a chaotic system is that they appear to have an unpredictable behavior called as deterministic disorder. The unstable behavior of chaotic system indicates that they tend not to resist output disturbance but instead it reacts in significant ways. Chaotic system exhibiting such a behavior is highly complex since it never repeats and continues to show the effect of disturbances.

Let we analyze the renowned Poincare equation which is otherwise called as Population equation [31]

$$a_{(n-1)}=A * a_n * (1-a_n) \tag{9}$$

A. Poincare Plot

This iterative function does not suddenly becomes chaotic but rather goes from the stage of convergence to a single value to bifurcation, additional bifurcation occurs. The system lacks periodicity and sensitivity to initial conditions, which is the most important feature of Chaos. Considering the above mentioned Poincare equation, A is a constant whose value decides the performance of the system. The recursion is dependent on selection of initial value a0, which is in the range of 0 and 1.

A slightly modified form of Poincare equation is given as

$$a_{(n+1)} = 12.25 a_n *(1-a_n)*(1-2a_n)^{(1/2)} \tag{10}$$

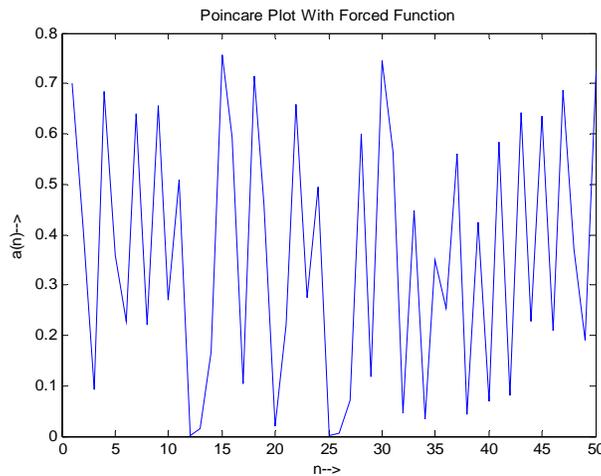


Figure 4: Poincare Plot with Forcing Function

Poincare Plot for Forced Function

As per equation (10) has forcing function as its last term. The constant can be changed but care should be taken such that the equation must fall within the chaotic area. Initial values for modified Poincare equation are nothing but coded numerical values from Fuzzy output. To analyze the variability and degree of theoretical chaos, we use third order differential plot which

is a graph plot between $(a(n+1)-a(n))$ vs $(b(n+1)-b(n))$ vs $(c(n+1)-c(n))$. It shows a better view of how for the points has been widely dispersed or clustered around the origin.

B. Center Tendancy Measure

Central Tendancy Measure (CTM) which is used to quantify the degree of variability in the third order differential plots(fig5). The CTM is computed by selecting a circular region of radius ‘r’ and dividing by the total number of points. Let t=total number of points and R is the radius of the central area,

$$\text{then CTM} = \frac{\sum_{i=1}^{t-2} \delta_i}{(t-2)}$$

$$\begin{aligned} \text{Where } \delta_i &= 1 \text{ if } \left[\left[(a_{(n+2)} - a_{(n+1)})^2 + (a_{(n+1)} - a_{(n)})^2 \right]^{1/2} < r \right. \\ &= 0 \text{ otherwise} \end{aligned} \tag{11}$$

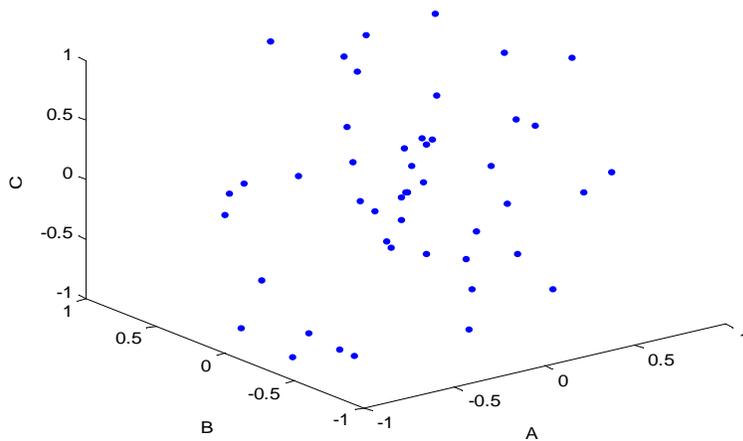


Figure 5: Spherical plot for finding CTM Using A,B,C

IV Elman Neural Networks For Optimization Of Fuzzy Outputs

As a result of Chaotic Optimization, 27 CTM values are obtained for each group. They are trained with the help of 27-4-1 neural network architecture. Artificial Neural Network (ANN’s) is a powerful tool in pattern recognition problems. Specifically, they are useful for automating diagnostic tasks carried out by experts (supervised classification tasks) [12]. The ANN’s

capability of learning from examples eases this knowledge acquisition problem [16]. On the other hand, the ANN gives opaque knowledge representation. Guoqiang (2000) and Jonathan Lee et al (1990) listed out the advantages of the neural networks in the following theoretical aspects [23],[24]. First, neural networks are data driven self-adaptive methods in that they can adjust themselves to the data without any explicit specification of functional or distributional form for the underlying model. Second, they are universal functional approximators in that neural networks can approximate any function with arbitrary accuracy. Third, neural networks are a nonlinear model, which makes them flexible in modeling real world complex relationships. Finally, neural networks are able to estimate the posterior probabilities, which provide the basis for establishing classification and performance. The Elman neural network is discussed in the following section of the report.

A. Elman Neural Network for Risk Level Optimization

The Elman neural network [22] is also known as partial recurrent network or simple recurrent network, the outputs of the hidden layer are allowed to feedback onto itself through a buffer layer, called context layer. This feedback allows Elman networks to learn, recognize and generate temporal patterns, as well as spatial patterns. Every hidden layer is connected to only one neuron of the context layer through a constant weight of value one. Hence, the context layer constitutes a kind of copy the state of the hidden layer, one instant before. The number of context neurons is consequently the same as the number of hidden neurons. Usually input, output and context neurons have linear activation functions, while hidden neurons have the sigmoidal activation function. The basic structure of the Elman neural network is illustrated in Fig 6.

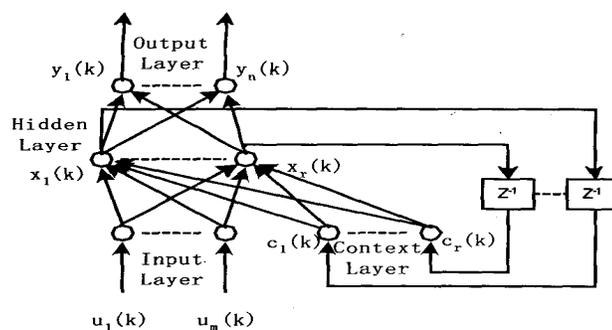


Figure.6 Structure of Elman Neural Network

It is easy to find that the Elman network mainly consists of four layers: input layer, hidden layer, context layer and output layer. There are adjustable weights connecting each two neighboring layers. Generally, it is considered as a special kind of feed forward neural network with additional memory neurons and local feedback. The self connections of the context nodes in the Elman network make it also sensitive to the history of input data which is very useful in dynamic signal modeling and analysis [28],[30].

The notation used in this section is given below:

w_{1ij} : The weight that connects node i in the input layer to the node j in the hidden layer.

w_{2ij} : The weight that connects node i in the input layer to the node j in the output layer.

w_{3ij} : The weight that connects context node i to the node j in the hidden layer.

m, n, r : The number of nodes in the input, output and hidden layers respectively.

$u_i(k), y_j(k)$: Input and outputs of the Elman neural Network, where $i=1,2,\dots,m$, and $j=1,2,3,\dots,n$.

$x_i(k)$: Output of the hidden node i , where $i=1,2,\dots,r$.

$c_i(k)$: The output of the context node i , i.e the output of the hidden node i of last time. z^{-1} : A unit time delay.

For each unit in the hidden layer an addition unit called context unit is added. The context unit is fully connected with all the hidden units in a forward manner. This means that there is a weight from every context unit to every hidden unit. Furthermore, there are recurrent connections from the hidden units back to the context units. But each hidden unit is connected to its associated context unit as shown in Fig.4. The weights of the recurrent connections are fixed and the forward weights get trained by using back propagation. In the forward phase the context units behave like input units. The values of the hidden units and of the output units get calculate in the same ways it is done in the feed forward networks. After calculating the outputs of the hidden units, the current values get copied into the corresponding units via the recurrent connections (through a unit delay). These values are used in the next time step. At the first time step they have to be set to some time step. During the backward phase of the training, target values for the

outputs are used and the forward weights are adjusted by back propagation. The inputs of network are: $u(k) \in R^m$, $y(k) \in R^n$, $x(k) \in R^r$, then the outputs in each layer can be given by

$$x_j(k) = f \left(\sum_{i=1}^m w_{2,i,j} u_i(k) + \sum_{i=1}^r w_{1,i,j} c_i(k) \right) \quad (12)$$

$$c_i(k) = x_i(k-1) \quad (13)$$

$$y_j(k) = g \left(\sum_{i=1}^r w_{3,i,j} x_i(k) \right) \quad (14)$$

Where, $f(\cdot)$ and $g(\cdot)$ are the linear or nonlinear output function of hidden layer and output layer respectively. Because the dynamic characteristics of Elman network are provided only by internal connection, so it needn't use the state as input or training signal. This is the advantage of the Elman network in contrast with static feed-forward network.

B. Learning and Testing Procedures for the Selection of Optimal Architecture in Elman networks

The primary aim of developing an ANN is to generalize the features (epilepsy risk level) of the processed fuzzy outputs. We have used different architecture of Elman networks for optimization. The network is trained using LM (Levenberg-Marquardt) algorithm to minimize the square output error. This error back propagation algorithm is used to calculate the weights updates in each layer of the network. The simulations were realized by employing Neural Simulator 4.0 of Matlab v.7.0 [21]. As the number of patterns in each database for training is limited, the technique of S-fold cross validation is employed to partition the data [19]. The available data is split up into Subsets each of equal size. The first subset is chosen to be test and the other S-1 subsets are combined to form the training and validation sets. After network is trained using these, the classification performance of test set is recorded. The process is then repeated so that each of the S-1 subsets acts as the test set in turn. The final classification performance is the average of the S test set results. In this paper, value of three was used for S. Since, we are using ten patients therefore ten models are selected. The use of cross validation removes any dependence of choice of pattern for the test set. The training process is controlled by monitoring the Mean Square Error (MSE) which is defined as [15], [17]

$$MSE = \frac{1}{N} \sum_{i=1}^N (O_i - T_j)^2 \quad (15)$$

Where O_i is the observed value at time i , T_j is the target value at model j ; $j=1-10$, and N is the total number of observations per epoch in our case it is 16. As the number of hidden units is gradually increased from its initial value, the minimum MSE on the testing set begins to decrease. The optimal number of hidden units is that number for which the lowest MSE is achieved. If the number of hidden units is increased beyond this performance does not improve and soon begins to deteriorate as the complexity of the neural network model is increased beyond that which is required for the problem. Based on the distribution of training patterns with MSE the learning rate is selected which is shown the fig.7. (Typically, a learning rate of 0.3 and a momentum term of 0.5 were used).

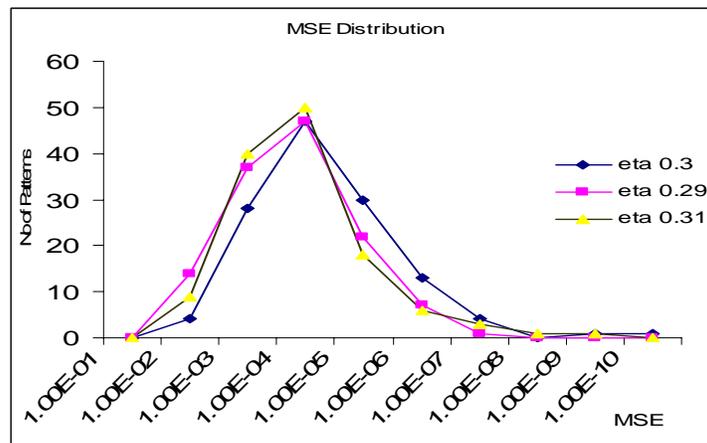


Figure 7 Selection of Learning Rate (eta) in Elman Neural Network

The squared error (e_i^2) from equation (12) between the input and the output of the ANN is converted into the confidence score using relation $C_i = \exp(-\lambda e_i^2)$ where refers to the neural network index [30]. In this paper we have chosen $\lambda=1$. The average confidence score for all Elman Network architecture is tabulated in the table.4. Table 4 shows the selection of Elman network architecture based on testing MSE. It is observed from table 4 the architecture 16-16-16 depicts the lowest number of training epochs and lesser MSE in testing. Once the optimal network architecture has been determine, the performance of the network models can be evaluated.

Table. 4 Estimation of MSE In Various Elman Network Architectures

Architecture	Mean Square Error (MSE)Index		Confidence score $C_i = \exp(-\lambda e_i^2)$
	Training	Testing	
9-9-9	0	3.874E-02	96.2
27-4-1	0	4.21E-03	99.65

In the Elman networks testing MSE index and number of epochs used for training are inversely proportional to each other. Therefore a compromise between them was achieved by taking into the consideration of larger training cost will ruin the system even though considerable accuracy is achieved in the targets (epilepsy risk levels) [18],[22]. Therefore we had selected 27-4-1 Elman network architecture which provides more accuracy in the classification which is depicted in fig.8.

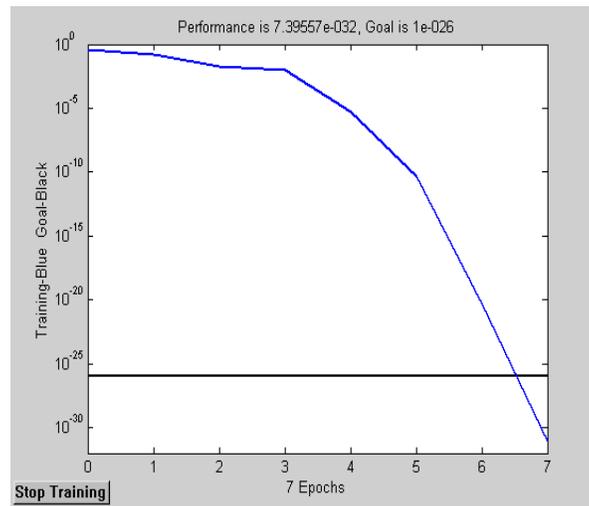


Figure 8. Training of Elman Neural Network (27-4-1)

V RESULTS AND DISCUSSION

The fuzzy outputs in three epochs for each patient are optimized by the neural network approach as a single epileptic risk level. The relative performance of the neural networks is studied through the Performance Index and the Quality Value parameters. These parameters

are calculated for each set of the patient and compared. The Performance Index (5) obtained by Fuzzy techniques, Chaotic optimization and Elman-Chaotic optimization are 40%, 79.2 and 100% respectively.

A. Quality Value

The goal of this research is to classify the epileptic risk level with as many perfect classifications and as few false alarms as possible. In Order to compare different classifier we need a measure that reflects the overall quality of the classifier. Their quality is determined by three factors.

- (i) Classification rate
- (ii) Classification delay
- (iii) False Alarm rate

The quality value Q_v is defined as,

$$Q_v = \frac{C}{(R_{fa} + 0.2) * (T_{dly} * P_{dct} + 6 * P_{msd})} \quad (16)$$

Where, C is the scaling constant

R_{fa} is the number of false alarm per set; T_{dly} is the average delay of the on set classification in seconds,

P_{dct} is the percentage of perfect classification, and P_{msd} is the percentage of perfect risk level missed.

A constant C is empirically set to 10 because this scale is the value of Q_v to an easy reading range. The higher value of Q_v , the better the classifier among the different classifier, the classifier with the highest Q_v should be the best. Table 5 shows the Comparison of the fuzzy and Chaotic optimization and Elman-Chaotic Optimization techniques. It is observed from table 5, that Chaotic-Elman network is performing well with the highest Performance Index and Quality Value. The Chaotic Elman network is a quick response method with least weighted delay of 2 seconds. In terms of false alarm Chaotic-Elman produces no false alarm when compared with Fuzzy and Chaotic optimization. Therefore for a given situation Chaotic-Elman is preferred than the Chaotic Optimization provided at the cost of loss of temporal information between the

adjacent channels of EEG signals. Hence, Chaotic Optimization is favored for long term analysis and Chaotic-Elman is adjudged for short term analysis.

Table 5 Results Of Classifiers Taken As Average Of Twenty Patients

Parameters	Fuzzy techniques before optimization	Chaotic optimization	Chaotic- Elman optimization
Risk level classification rate (%)	50	82.8	100
Weighted delay (s)	4	2.04	2
False-alarm rate/set	0.2	0.108	-
Performance Index %	40	79.2	100
Quality value	6.25	15.92	25

VI CONCLUSION

Chaotic Elman optimization provides a perfect classification when compared all other existing methods. The objective was to classify perfect risk levels with high rate of classification, a short delay from onset, and a low false alarm rate. Here we obtained PI of 100% and Zero false alarm rate. From this method we can infer the occurrence of High-risk level frequency and the possible medication to the patients. The future research is in the direction of an improved chaotic optimization models.

Acknowledgement

The authors thank the Management and Principle of Bannari Amman Institute of Technology, Sathyamangalam and for providing the research facility and encouragement. We thank the chief Medical officer Dr.Asokan, Sri Ramakrishna Hospital –Coimbatore and Dr.B.Rajalakshmi Diabetologist, Govt Hospital Dindigul for providing the patient EEG Signal and information.

REFERENCES

- [1] Leon D.Iasemidis et al., Adaptive Epileptic Seizure Prediction System, IEEE Transactions on Biomedical Engineering, May 2003,50(5): 616-627.

[2] K P Adlassnig, Fuzzy Set Theory in Medical diagnosis, IEEE Transactions on Systems Man Cybernetics, March 1986,16: 260-265.

[3] Alison A Dingle et al, A Multistage system to Detect epileptic form activity in the EEG,IEEE Transactions on Biomedical Engineering,1993, 40(12):1260-1268.

[4] Hauquo and Jean Gotman, A patient specific algorithm for detection onset in long-term EEG monitoring possible use as warning device, IEEE Transactions on Biomedical Engineering, February 1997, 44(2): 115-122.

[5] Arthur C Gayton, Text Book of Medical Physiology, Prism Books Pvt. Ltd., Bangalore, 9th Edition, 1996.

[6] J.Seunghan Park et al, TDAT Domain Analysis Tool for EEG Analysis, IEEE Transactions on Biomedical Engineering, August 1990,37(8): 803-811.

[7] Donna L Hudson, Fuzzy logic in Medical Expert Systems, IEEE EMB Magazine, November/December 1994,13(6): 693-698.

[8] C B Gupta and Vijay Gupta, An Introduction to Statistical Methods, 22nd Ed., Vikas Publishing House Lt., 2001.

[9] R.Harikumar and B.Sabarish Narayanan, Fuzzy Techniques for Classification of Epilepsy risk level from EEG Signals, Proceedings of IEEE Tencon – 2003, 14-17 October 2003,Bangalore, India, 209-213.

[10] Mark van Gils, Signal processing in prolonged EEG recordings during intensive care, IEEE EMB Magazine November/December 1997,16 (6): 56-63.

[11] R.Harikumar, Dr.(Mrs). R.Sukanesh, P.A. Bharthi, Genetic Algorithm Optimization of Fuzzy outputs for Classification of Epilepsy Risk Levels from EEG signals,I.E . India Journal of Interdisciplinary panels, May 2005, 86(1):9-17.

[12] Celement.C etal, A Comparison of Algorithms for Detection of Spikes in the Electroencephalogram, IEEE Transaction on Bio Medical Engineering, April 2003, 50 (4): 521-26.

- [13] Pamela McCauley-Bell and Adedeji B.Badiru, Fuzzy Modeling and Analytic Hierarchy Processing to Quantify Risk levels Associated with Occupational Injuries- Part I: The Development of Fuzzy- Linguistic Risk Levels, IEEE Transactions on Fuzzy Systems, 1996,4 (2): 124-31.
- [14] Joel.J etal, Detection of seizure precursors from depth EEG using a sign periodogram transform, IEEE Transactions on Bio Medical Engineering, April 2004,51 (4):449-458.
- [15] S.Vitabile etal.,Daily peak temperature forecasting with Elman neural networks, Proceedings of IEEE 2004, 2765-2769.
- [16] Nurettin Acir etal., Automatic detection of epileptiform events in EEG by a three-stage procedure based artificial neural networks, IEEE transactions on Bio Medical Engineering January 2005,52(1):30-40.
- [17] Drazen.S.etal., Estimation of difficult –to- Measure process variables using neural networks, Proceedings of IEEE MELECON 2004,May 12-15, 2004, Dubrovnik,Croatia, 387-390.
- [18] Moreno.L.etal., Brain maturation estimation using neural classifier, IEEE Transaction of Bio Medical Engineering , April 1995,42(2):428-432.
- [19] Tarassenko.L,Y.U.Khan,M.R.G.Holt,Identification of inter-ictal spikes in the EEG using neural network analysis, IEE Proceedings –Science Measurement Technology, November 1998,145(6):270-278.
- [20] Hwang et al., Recognition of Unconstrained Handwritten Numerals by A Radial Basis Function Network Classifier, Pattern Recognition Letters, 1997,18:657-664.
- [21] H.Demuth and M.Beale, Neural network tool box: User’s guide, Version 3.0 the math works, Inc., Natick, MA, 1998.
- [22] G.Fung etal, Fault Detection In Inkjet Printers Using Neural Networks, Proceedings of IEEE SMC, 2002.

- [23] Guoqiang Peter Zhang , Neural Networks for Classification: A Survey ,IEEE Transactions on Systems Man Cybernetics- Part C: Applications and Reviews, November 2000,30(4): 451-462.
- [24] Jonathan lee etal., A Neural Network Approach to Cloud Classification, IEEE Transactions on Geosciences and Remote Sensing, September 1990,28 (5): 846-855.
- [25] S.Haykin, Neural networks a Comprehensive Foundation, Prentice- Hall Inc. 2nd Ed. 1999.
- [26] Mu-chun Su, Chien –Hsing Chou, A modified version of the k-means clustering algorithm with a distance based on cluster symmetry, IEEE Transactions on Pattern Analysis and Machine Intelligence June 2001, 23 (6): 674-680.
- [27] Masaaki Tsujitani, Takashi Koshimizu, Neural Discriminant Analysis, IEEE Transactions on Neural Networks, November 2000,11 (6):1394-1401.
- [28] Rangaraj M. Rangayyan, Bio- Medical Signal Analysis A Case Study Approach, IEEE Press-John Wiley &sons Inc New York 2002.
- [29] Yuan-chu Cheng, Wei-Min Qi,WeiYou Cai, Dynamic Properties of Elman And Modified Elman Neural Network,Proceedings of IEEE, First International Conference on Machine Learning And Cybernetics, Beijing, 2002,637-640.
- [30] K.Sriramamurty and B.Yegnannarayana, Combining Evidence from Residual Phase and MFCC Features for Speaker Recognition, IEEE Signal Processing Letters, January 2006,13 (1): 52-55.
- [31] M.E. Cohen, Donna.L Hudson, and Prakash.C.D, applying continuous Chaotic modeling to cardiac signal analysis, IEEE EMBS Magazine, September 1996 vol,15, (5),pp 97-102.
- [32] Albano.A.M., Muench.J., Schwartz. C, Mees.A.I. and Rapp.P.E, (1988) ‘Singularvalue decomposition and the Grassberger-Procaccia algorithm’Physical Review A,pp.3017-3026.
- [33] Broomhead. D.S. and King .G.P, (1986) ‘Extracting qualitative dynamics from experimental data’ Physical Review D, pp.217-236.

- [34] Farmer. J.D., Ott.E., Yorke. J.A. (1983) , ‘The dimension of chaotic attractors’ Physical Review D, pp.153-180
- [35] Ivan Dvorak and Jarorair Siska, (1986), ‘On some problems encountered in calculating the correlation dimension of EEG’, International centre for Theoretical physics. International atomic energy agency, United Nations educational, scientific and cultural organization.
- [36] M.E. Cohen, Donna.L Hudson, Inclusion of ECG and EEG Analysis in Neural Network Models, Proceedings of International Conference of the 23rd Annual IEEE EMBS October 25-28, 2001, Istanbul Turkey, pp 1621-1624.
- [37] N.C. Bhavaraju, Mark.G.Frei and Ivan Osorio, Analog seizure Detection and Performance Evaluation, IEEE Transactions on BME Vol 53, (2), February 2006, pp 238-245.