A Comparison of Elman and Radial Basis Function (RBF) Neural Networks in Optimization of Fuzzy outputs for Epilepsy Risk Levels Classification from EEG Signals

R. Harikumar, T. Vijayakumar

Abstract— In this paper; we investigate the optimization of fuzzy outputs in the classification of epilepsy risk levels from EEG (Electroencephalogram) signals using two categories (Recurrent &Non Recurrent) of neural networks. The fuzzy techniques are applied as a first level classifier to classify the risk levels of epilepsy based on extracted parameters like energy, variance, peaks, sharp and spike waves, duration, events and covariance from the EEG signals of the patient. Elman neural network (with error Back propagation training) & Radial Basis Function (RBF) neural network are identified as post classifiers on the classified data to obtain the optimized risk level that characterizes the patient's epilepsy risk level. The Performance Index (PI) and Quality Value (QV) are calculated for the above methods. A group of ten patients with known epilepsy findings are used in this study. High PI such as 97.87 %, and 98.92% was obtained at QV's of 23.31, and 23.98 in Elman and RBF neural network optimization when compared to the value of 40% and 6.25 through fuzzy techniques respectively. We find that the RBF (Non Recurrent) neural network out performs Elman Network in optimizing the epilepsy risk levels.

Index Terms— EEG Signals, Epilepsy Risk Levels,Fuzzy Logic, RBF, Elman Neural Networks, Back propagation.

I. INTRODUCTION

Epileptic seizures result from a temporary electrical disturbance of the brain. Some times seizures may go unnoticed, depending on their presentation, and some times may be confused with other events, such as a stroke, which can also cause falls or migraines. Approximately one in every 100 persons will experience a seizure at some time in their life. Unfortunately, the occurrence of an epileptic seizure seems unpredictable and its process is very little understood [1]. Since its discovery by R.Caton, theElectroencephalogram (EEG has been the most utilized signal to clinically assess brain activities. Twenty –five percent of the world's 50 million people with epilepsy have seizures that cannot be controlled by any available treatment [2]. The need for new therapies, and success of similar devices to treat cardiac arrhythmias, has spawned an explosion of research into algorithms for use in implantable therapeutic devices for epilepsy. Most of these algorithms focus on either detecting unequivocal EEG onset of seizures or on quantitative methods for predicting seizures in the state space, time, or frequency domains that may be difficult to relate to the Neuro physiology of epilepsy [3]. 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, neuroelectric 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.

In this paper, we investigate the efficacy of Elman &RBF neural networks 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 outline of the paper is as follows: Section II outlines the procedure used for collecting the EEG Recording and summarizes the analysis on EEG Signal by the fuzzy methods. Section III presents the analysis of Elman & RBF Neural networks in optimization of fuzzy outputs for risk level classification. Section IV interprets the results of the optimization methods and section V concludes the paper.

II. MATERIALS AND METHODS

The EEG data used in the study were acquired from ten 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 Butter worth filters to remove the artifacts.

Manuscript received on January, 2013.

Dr.R.Harikumar,Professor, ECE Department,.Bannari Amman Institute of Technology, Sathyamangalam India.

T.Vijayakumar Assistant Professor, (Sr.Gr) Department of IT, Bannari Amman Institute of Technology, Sathyamangalam.

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 (Golden standard with 100% sensitivity &100% specificity), 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 both training and testing. These EEG records had an average length of six seconds and total length of 60 seconds. The patients had an average age of 31 years. A total of 480 epochs of 2 seconds duration are used.

B. Fuzzy System as Pre Classifier

Fuzzy system is shown in figure 1. The main objective of this research is to classify the epilepsy risk level of a patient from EEG signals. 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 neural networks optimization methods is analyzed.
4. A comparison of supervised recurrent neural network (Elman) & hybrid non recurrent neural network (RBF) is analyzed.

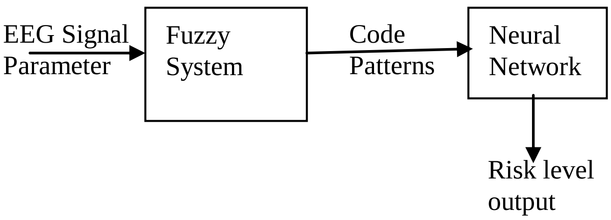

Figure 1 Neuro- Fuzzy Classification System

1. The energy in each two-second epoch is given by

$$E = \sum_{i=1}^{n} x_i^2 \qquad (1)$$

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

$$\sigma^2 = \frac{\sum_{i=1}^{n}(x_i - \mu)^2}{n} \qquad (2)$$

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

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

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

A sample value of extracted above seven features for the patient record 4 is shown in table I. In the table I abnormal case all the sixteen channels do not show high risk characteristics of EEG signal. 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.

Table I - 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	1	2	2
Total	9	38	35
Sharp &Spike	8	6	6
Total	122	91	87
Events	12	10	10
Total	185	154	145
Average duration	3.798	4.042	3.883
Covariance	0.5793	0.5123	0.5941

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 II

Table II - 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	YYYXYY
YYZXYY	YYYYYY	YYYYYY
YZZYXY	XZZXYY	YYYYYY
ZZZYYY	WYYYXX	YYYXYY
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],

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

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* (6)

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

The specificity is defined as [22]

S_p= *[PC/PC+MC] *100* (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 III. 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 III- 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	Σ=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.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

Thus the cross correlation function $\mathbf{r_{xy}(m)}$ of the epochs x(n) and y(n) is defined by the equation (6) and assuming that both sequence have been measured from n=0 to n=N-1, in our case n=1to 16,[25]

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

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 Neural network optimization technique (post classifier) [15]. A pertinent elucidation for the neural network optimization is given below.

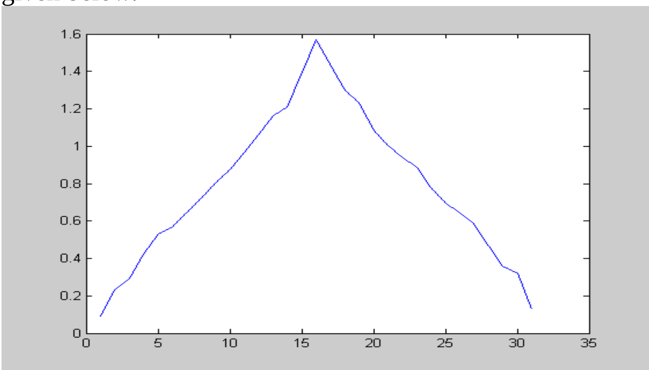

Figure.3 Cross Correlation Function plot for the Adjacent Epochs in fuzzy based Epilepsy Risk Level Outputs

III. NEURAL NETWORKS FOR OPTIMIZATION OF FUZZY OUTPUTS

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 etal(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 & RBF neural networks are discussed in the following section of the paper.

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.4.

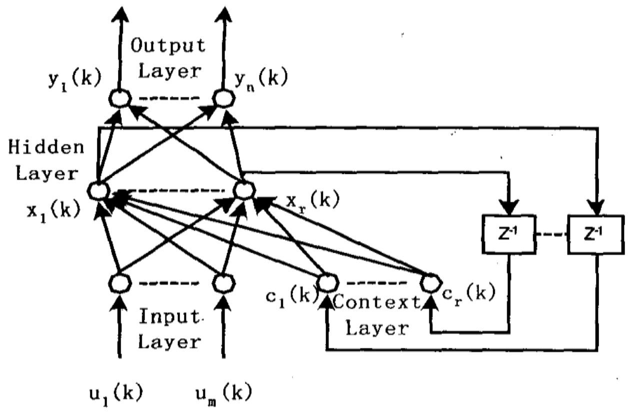

Figure.4 Structure of Elman Neural Network

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:

$w1_{ij}$: The weight that connects node i in the input layer to the node j in the hidden layer.

$w2_{ij}$: The weight that connects node i in the input layer to the node j in the output layer.

$w3_{ij}$: 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,…m, and j=1,2,3,…n.

x_i (k): Output of the hidden node i, where i=1,2,..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) € R^m , y(k) € R^n , x(k) € R^r , then the outputs in each layer can be given by

$$x_j(k) = f\left(\sum_{i=1}^{m} w2_{i,j} u_i(k) + \sum_{i=1}^{r} w1_{i,j} c_i(k)\right) \quad (9)$$

$$c_i(k) = x_i(k-1) \quad (10)$$

$$y_j(k) = g\left(\sum_{i=1}^{r} w3_{i,j} x_i(k)\right) \quad (11)$$

Where, f (·) and g (·) 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-Maruardt) 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-1subsets 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]

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(O_i - T_j)^2 \quad (12)$$

Where Oi 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.5. (Typically, a learning rate of 0.3 and a momentum term of 0.5 were used).

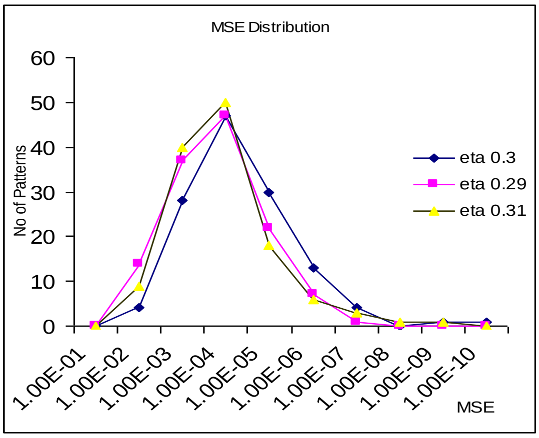

Figure 5 Selection of Learning Rate (eta) in Elman Neural Network

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 $\mathbf{C_i=exp(-\lambda e_i^2)}$ where refers to the neural network index [31]. In this paper we have chosen λ=1. The average confidence score for all Elman Network architecture is tabulated in the table IV. Table IV shows the selection of Elman network architecture based on testing MSE. It is observed from table IV 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 IV
ESTIMATION OF MSE IN VARIOUS ELMAN NETWORK ARCHITECTURES

Architecture	Training Epochs	Mean Square Error (MSE)Index		Confidence score $C_i=exp(-\lambda e_i^2)$
		Training	Testing	
16-16-16	7	0	2.067E-05	99.99
8-8-8	20	0	3.71E-03	99.62
4-4-4	73	0	1.669E-04	99.98
2-2-2	256	0	4.3E-07	99.99
1-1-1	337	1.3E-08	1.3E-08	99.99

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 16-16-16 Elman network architecture which requires lesser number of training epochs and the same is depicted in the fig. 6. Analysis of Radial Basis Function (RBF) Neural network is focused in the next section of the paper.

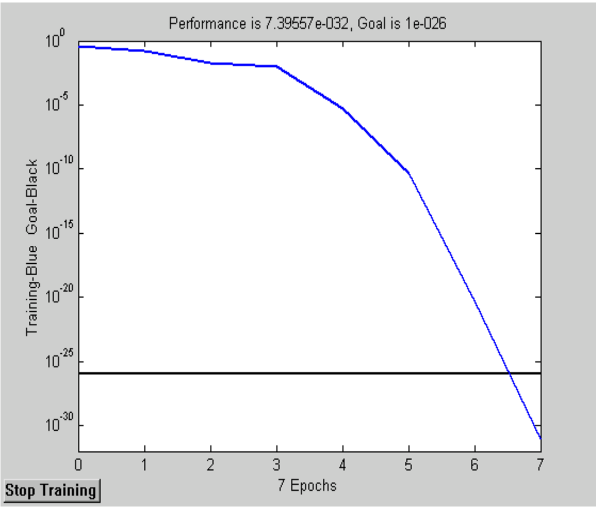

Figure 6.Training of Elman Neural Network (16-16-16)

C. Radial Basis Function (RBF) Neural Network as a post Classifier for Risk Level Optimization

The Radial Basis Function (RBF) neural network is widely used for function approximation, pattern classification and recognition due to its structural simplicity, universal approximators, and faster learning abilities due to locally tuned neurons [20]. In its basic form a RBF neural network involves three functionally distinct layers. The input layer is simply a set of sensory units. The second layer is a hidden layer of sufficient dimension which applies a non-linear transformation of the input space to a higher dimensional hidden. An RBF neural network is generally trained in two steps one after another. In the first step, the centers of hidden layer neurons are selected. Then the weights between the hidden and output layers are estimated. The centers of the hidden layer neurons of an RBF neural network are selected in different ways. Generally, these centers are selected by using some clustering algorithm like, k-means, fuzzy c-means, etc, [26].

D. Architecture of an RBF neural network

The architecture of an RBF neural network is shown in fig, 7. It consists of one input layer, one hidden layer and one output layer. Each input neuron is corresponds to an element of an input vector and is fully connected to the n hidden layer neurons and the bias neuron. Again, each of the hidden neuron and the bias neuron also fully connected to the output neurons. The output of a hidden layer neuron is usually generated by a Gaussian function as follows:

$$\varphi_i(X) = \begin{cases} \exp\left(-\dfrac{\|X - t_i\|}{2\sigma_i^{\,2}}\right) & ; i = 1,2,....,n \\ 1; i = 0(biasneuron) \end{cases} \qquad (13)$$

Where X is an input vector and t_i σ_i are the center and the width of the respective field of the i^{th} neuron of the hidden layer respectively. The number of neurons in the output layer is equal to the possible classes of the given problem. Each output layer neuron computes a linear weighted sum of the outputs of the hidden layer neurons as follows [27]:

$$z_j = \sum_{i=0}^{n} \varphi(X) w_{ij};\ j = 1,2,...c, \qquad (14)$$

Where w_{ij} is the weight between i^{th} hidden layer neuron and j^{th} output layer neuron.

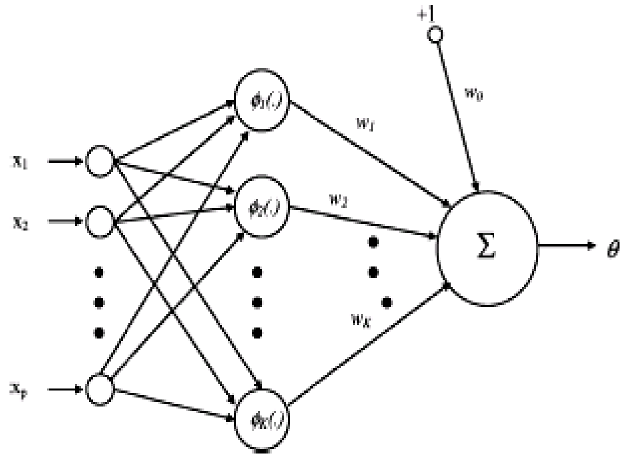

Figure.7. Radial Basis Function Network

E. Training and Testing Procedures for the Selection of Optimal Architecture in RBF Networks

The primary aim of developing an ANN is to generalize the features (epilepsy risk level) of the processed fuzzy outputs. We have applied different architectures of RBF networks for optimization. Even though RBF is an unsupervised network, the cluster centers of the hidden layers are identified as the target codes (**ZZYZZZ** -epilepsy risk level) for a particular model. The weights of the linear connections between the hidden layer and output layer network are trained with error back propagation algorithm to minimize the square output error to zero. The simulations were realized by employing Neural Simulator 4.0 of Matlab v.7.0 [21]. Since our neural network model is patient specific in nature, we are applying 48 (3x16) patterns for each RBF model. There are ten models for ten patients. As the number of patterns in each database for training is limited, each model is trained with one set of patterns (1x16) for zero mean square error condition and tested with other two sets of patterns (2x16). After network is trained using these, the classification performance of test set is recorded. The testing process is monitored by the MSE by the equation (12). 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. The average confidence score **C_i=exp (-λe_i^2)** for each RBF Network architecture is tabulated in the table V. Table V shows the selection of RBF network architecture based on their testing MSE. It is observed that the architecture 1-16-1 depicts lesser test MSE index and it is selected.

Even though 8-2-8 architecture exhibits zero test MSE index is not selected due to its unstable nature. Once the optimal network architecture has been determine, the performance of the network models can be evaluated.

Table V - Estimation Of Mse In Various Rbf Network Architectures

Elman Neural Network		RBF Neural Network	
Test Risk Level Code	Mean Square Error(MSE)	Test Risk Level Code	Mean Square Error(MSE
ZZYZZZ	1.95E-05	ZZYZZZ	1.85E-08
YYYXYY	7.45E-06	YYYXYY	2.45E-08
YYYYYY	4.52E-05	YYYYYY	1.5E-08
YYXXYY	3.56E-05	YYXXYY	2.4E-08
ZZYYXY	4.5E-05	ZZYYXY	4.5E-08
XXZYXY	9.12E-03	**XXZYXX**	5.12E-06
ZYYYZZ	1.98E-05	ZYYZZZ	1.4E-08
YYYXXX	3.04E-05	YYYXXX	2.14E-08
ZYZYXW	1.78 E-06	ZYZYXW	1.65E-08
XYYZWZ	1.62 E-06	XYYZWZ	1.7E-05

IV. RESULTS AND DISCUSSIONS

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, Elman neural network and RBF optimization are 40%, 97.875%, and 98.92% respectively. The following table VI, Depicts the comparison of the epilepsy risk level estimation for ten patient specific (16-16-16) Elman Neural Network and (1-16-1) RBF Neural network. Both the neural networks are performing well in the classification tasks expect in the patient model 6, and this is due the inherent mixed risk level patterns exhibited by the patient's EEG Signal parameters. While achieving the target codes, RBF networks demonstrates very low MSE values when compare to the MSE values of Elman networks. This illustrates the quantitative capabilities of RBF networks than the Elman networks.

Table VI - Comparison of Mse Estimation At Ten Patient Specific Elman &Rbf Neural Networks

Architecture	Mean Square Error (MSE)Index		Confidence score $C_i=\exp(-\lambda e_i^2)$
	Training	Testing	
1-16-1	3.3 E-08	3.3E-08	99.99
2-8-2	4.21E-07	4.21E-07	99.99
4-4-4	3.4 E-07	3.4 E-07	99.99
8-2-8	0	0	100
16-1-16	0	2.94E-04	99.97

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 [11],

$$Q_V = \frac{C}{\left(R_{fa}+0.2\right)*\left(T_{dly}*P_{dct}+6*P_{msd}\right)} \quad (15)$$

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 VII shows the Comparison of the fuzzy and neural networks optimization techniques. It is observed from table VII, that Elman &RBF neural network are performing well with the highest Performance Index and Quality Values. The Elman neural network is a quick response method with least weighted delay of 1.95 seconds. Even though both neural networks are performing well in terms of parameters but the training over head for Elman network is higher than the unsupervised RBF network. In terms of false alarm Elman produces more false alarm than the RBF networks. Therefore for a given situation RBF is preferred than the Elman provided at the cost of loss of temporal information between the adjacent channels of EEG signals. Hence, Elman is favored for long term analysis and RBF is adjudged for short term analysis

Table VII - Results of Classifiers Taken As Average Of Ten Patients

Parameters	Fuzzy Techniques before Optimization	Elman Neural Network	RBF Neural Network
Risk level classification rate (%)	50	97.92	98.92
Weighted delay (s)	4	1.95	1.978
False-alarm rate/set	0.2	0.0208	0.0108
Performance Index %	40	97.87	98.92
Quality value	6.25	23.31	**23.98**

Computational burden increases in the Elman network as the number of neurons are increased in the hidden layer.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

Since Elman consider optimization problem as a whole and it does not divide them into a multiple steps which leads to heavy computational overhead on the network. However, RBF's computational overhead is a reduced one when compared to Elman network. Since RBF network splits the problem into two steps as

(i). Recognition of center of clusters through k-means algorithms.

(ii). Estimation of linear weights between hidden layer and output layer through Back propagation algorithm.

When compared to Elman network though the increment of hidden layer neurons in RBF network does not imply any additional computational burden.

V. CONCLUSION

This paper investigates the performance of neural networks in optimizing the epilepsy risk level of epileptic patients from EEG signals. The parameters derived from the EEG signal are stored as data sets. Then the fuzzy technique is used to obtain the risk level from each epoch at every EEG channel. The objective was to classify perfect risk levels with high rate of classification, a short delay from onset, and a low false alarm rate. Though it is impossible to obtain a perfect performance in all these conditions, some compromises have been made. As a high false alarm rate ruins the effectiveness of the system, a low false-alarm rate is most important. Elman and RBF Neural network optimization techniques are used to optimize the risk level by incorporating the above goals. The classification rate of epilepsy risk level of above 98% is possible in our method. The missed classification is almost nil for a short delay of 2 seconds. The number of cases from the present ten patients has to be increased for better testing of the system. From this method we can infer the occurrence of High-risk level frequency and the possible medication to the patients. Also optimizing each region's data separately can solve the focal epilepsy problem. The future research is in the direction of a comparison between non heuristic optimization models with neural networks.

ACKNOWLEDGEMENTS

The authors express their sincere thanks to the Management and the Principal of Bannari Amman Institute of Technology, Sathyamangalam for providing the necessary facilities for the completion of this paper. This research is also funded by AICTE RPS.:F No 8023/BOR/RID/RPS-41/2009-10, dated 10th Dec 2010.

REFERENCES

1. Leon D.Iasemidis etal., Adaptive Epileptic SeizurePrediction 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. Haoqu 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 Transaction 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 IEEE2004, 2765-2769.
16. Nurettin Acir etal., Automatic detection of epileptiform events in EEG by a three-stage procedure based artificial neural networks, IEEE transaction 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. Arassenko.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. Meng Joo Er,Shiqian Wu,Juwei Lu, Face Recognition with Radial Basis Function (RBF) Neural Network, IEEE Transactions on Neural Networks, May 2002,13 (3): 697-710.
27. 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.
28. Masaaki Tsujitani, Takashi Koshimizu, Neural Discriminant Analysis, IEEE Transactions on Neural Networks, November 2000,11 (6):1394-1401.
29. Rangaraj M. Rangayyan, Bio- Medical Signal Analysis A Case Study Approach, IEEE Press-John Wiley &sons Inc New York 2002.
30. 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 Cybernatics, Beijing, 2002,637-640.

31. 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.

AUTHORS PROFILE

Dr. R. Harikumar received his B.E (ECE) Degree from REC (NIT) Trichy in 1988..He obtained his M.E (Applied Electronics) degree from College of Engineering, Guindy, Anna University Chennai in 1990. He was awarded Ph.D. in I &C Engg from Anna university Chennai in 2009. He has 22 years of teaching experience at college level. He worked as faculty at Department of ECE, PSNA College of Engineering & Technology, Dindigul. He was Assistant Professor in IT at PSG College of Technology, Coimbatore. He also worked as Assistant Professor in ECE at Amrita Institute of Technology, Coimbatore. Currently he is working as Professor ECE at Bannari Amman Institute of Technology, Sathyamangalam. He has published fifty six papers in International and National Journals and also published around ninty three papers in International and National Conferences conducted both in India and abroad. His area of interest is Bio signal Processing, Soft computing, VLSI Design and Communication Engineering. He is guiding nine PhD theses in these areas. He is life member of IETE, ISTE , IAENG,and member IEEE.

Mr.T.Vijayakumar received his B.E (ECE) degree from Bharathiyar University in 2003 .He obtained his M.E (Computer Communication)degree from Anna university Chennai in 2007. He has Six years of teaching experience at college level. He is pursuing his doctoral programme as a research scholar in Bio-Medical Engineering under guidance of

Dr. R.Harikumar at Department of ECE, Bannari Amman Institute of Technology Sathyamangalam. Currently he is working as Assistant Professor(Senior Grade) in Department of IT at Bannari Amman Institute of Technology, Sathyamangalam. He has published eleven papers in International and National Journals and also published around thirteen papers in International and National Conferences conducted both in India and abroad. His area of interest is Bio signal Processing, Soft computing, Embedded systems and Communication Engineering. He is Member of ISTE.