Interdiscip Sci Comput Life Sci (2012) 4: 209–214
DOI: 10.1007/s12539-012-0129-6

Epilepsy Diagnosis Based on Generalized Feed Forward Neural Network

P.A. KHARAT[1*], S.V. DUDUL[2]
[1](Department of Information Techanology, Anuradha Engineering College, Chikhli 443201, India)
[2](Department of Applied Electronics, S.G.B. Amaravati University, Amaravati 444602, India)

Received 22 November 2011 / Revised 11 March 2012 / Accepted 19 March 2012

Abstract: Epilepsy is a common neurological disorder that is characterized by recurrent unprovoked seizures. Epilepsy can develop in any person at any age. 0.5% to 2% of people will develop epilepsy during their lifetime. This paper aims to develop the clinical decision support system (DSS) for the diagnosis of epilepsy. In this paper a simple, reliable and economical Neural Network (NN) based DSS was proposed for the diagnosis of epilepsy. The generalized feed forward neural network (GFFNN) was designed for the diagnosis. Eleven statistical parameters along with the 64 FFT were extracted for the electroencephalogram (EEG) signal. Data used for the experimentation purpose was obtained from the University of Bonn. The classification rate of GFFNN was 100 % for the training data and 86.67% for the cross validation data.
Key words: generalized feed forward neural networks (GFFNN), decision support system (DSS), electroencephalogram (EEG), epilepsy, approximate entropy (ApEn).

1 Introduction

Epilepsy is a brain disorder in which clusters of nerve cells, or neurons in the brain sometimes have abnormal signal. In the epilepsy, the normal pattern of neurons activity becomes disturbed causing strange sensation, emotion, behavior and loss of consciousness (Engel, 1989; Robert, 2005). Epilepsy is a disorder due to many possible causes. Anything that disturbs the normal pattern of neuron activity may result in illness, brain damage, or abnormal brain development. EEG scan is a common diagnostic test for epilepsy and can detect abnormalities in the brain electrical activity. People with epilepsy frequently have changes in their normal pattern of brain wave, even though they are not experiencing a seizure. EEG plays an important role in the diagnosis of epilepsy.

Researchers proposed many automatic systems for the diagnosis of epilepsy. Akin *et al.* (2001) developed a classification method for the diagnosis. Another work (Szilagyi *et al.*, 2001) recommended the recognition of epileptic waveform by using the multi-resolution wavelet decomposition of EEG signal. Shrinivasan designed the approximate entropy based Elman neural network and probabilistic neural network for detection of epilepsy (Shrinivasan, 2007). The method proposed by Sriraam *et al.* (Shrinivasan, 2004) uses recurrent neural network classifier with wavelet entropy and spectral entropy features as the input for the automated detection of epilepsy.

This paper discusses the automated DSS for the epilepsy diagnosis using the GFFNN. Eleven statistical and 64 FFT were input to the GFFNN. Fig. 1 shows the block diagram of proposed DSS. As compare to the existing system the proposed DSS is simple and economical.

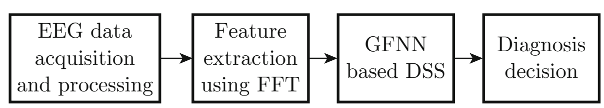

Fig. 1 Block diagram of proposed DSS.

2 Benchmark EEG data set

The EEG data considered for this work was extracted from EEG database of University of Bonn which is available in public domain (Ralph, 2001). The complete database was comprised of five sets of dataset referred as A-E. Each dataset contained 100 single channel EEG segment without any artifacts with 23.6 sec. Sets A and B contained recording obtained from surface EEG recording that was carried out on five healthy volunteers using a standardized 10-20 electrode place-

*Corresponding author.
E-mail: pravinakharat82@gmail.com

ment scheme. Sets C and D contained only activity measured during seizure free interval, recordings in set D were from the epileptognic zone and those in set C were from the hippocampal formation of apposite hemisphere of the brain. Set E only contained the seizure activity.

All signals were recorded with 128-channel amplifier system, using an average common reference. After 12 bit analog-to-digital conversion, the data were written continuously onto the disk of a data acquisition computer system at sampling rate of 173.61 Hz. Band pass filter setting was 0.53-40 Hz.

Two sets of EEG data were selected for the training and testing of neural network, set A for healthy subject and set D for epileptic subjects during a seizure free interval that indicates interictal activity.

3 Feature extraction

The feature vector was extracted by FFT and 11 statistical features namely standard deviation (STDV), min, max, mean, entropy, minima, maxima, power spectral density (PSD), approximate entropy (ApEn) and number of peaks. The data set was prepared for all 100 normal segments of the set A and also for all 100 abnormal segments of the set D. The ApEn was extracted by the method discussed by Shrinivasan *et al.* (2007). Other features were extracted by using MATLB and Microsoft Office Excel.

(1) The energy of EEG signal is given by

$$E = \sum_{i=1}^{n} x_i^2, \tag{1}$$

where x_i is the sample value and n is the number of samples.

(2) The standard deviation is calculated by using the formula

$$S = [\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2]^{\frac{1}{2}}, \tag{2}$$

where

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i. \tag{3}$$

(3) Min is the smallest element in EEG segment.

(4) Max is the largest element in EEG segment.

(5) Minima and Maxima are the average of all local minima and maxima in the EEG time series.

(6) Mean is the mean value of all the samples in EEG segment.

4 Generalized feed forward neural network

Generalized feed forward network is a generalization of the Multilayer Perceptron (MLP), such connections can jump over one or more layers. In theory, a MLP can solve any problem that a generalized feed forward network can solve. In practice, however, generalized feed forward networks often solve the problems much more efficiently (Arulampalam *et al.*, 2004; Derks *et al.*, 1995). A classic example of this is the two spiral problem. Without describing the problem, it suffices to say that a standard MLP requires hundreds of times more training epochs than the generalized feed forward network containing the same number of processing elements.

Learning from the data is the essence of neurocomputing. Every PE that has an adaptive parameter must change it according to some prespecified procedures. Back propagation is by far the most common form of learning. Here it is sufficient to say that the weights are changed based on their previous values and a correction term. The learning rule is the means by which the correction term is specified. Once the particular rule is selected, the user must still specify how much correction should be applied to the weights, referred to as the learning rate. If the learning rate is too small, then learning takes a long time. On the other hand, if it is set too high, then the adaptation diverges and the weights are unusable.

4.1 Step

The step uses gradient descent learning rules, e.g. back propagation and real time recurrent learning, and provides first order gradient information about the network's performance surface. In other words, they estimate which way is up. The most straightforward way of reaching the bottom (the minima) is to move in the opposite direction. With this scenario, the only variable is the step size, *i.e.* how far it should move before obtaining another directional estimate. If the steps are too small, then it will take too long to get there. If the steps are too large, then it may overshoot the bottom, causing rattle or even diverge.

4.2 Momentum

Momentum learning is a robust method to accelerate learning, and it is recommended as the default search rule for networks with nonlinearities. The momentum provides the gradient descent with some inertia so that it tends to move along a direction that is the average estimate for down. The amount of inertia (*i.e.*, how much of the past to average over) is dictated by the momentum parameter ρ. The higher the momentum is, the more it smoothes the gradient estimate and the less effect a single change in the gradient has on the weight change. The major benefit is the added ability to break out of local minima that a step component might otherwise get caught in. Note that oscillations may occur if the momentum is set too high. The momentum parameter is the same for all weights of the attached component. An access point has been provided for the step size and momentum allowing access for adaptive and scheduled learning rate procedures. The weight update equation for momentum is given in Equation (4).

$$\Delta w_i(n+1) = \eta_i \nabla w_i + \rho \Delta_i(n) \tag{4}$$

Where w is the weight.

4.3 Conjugate-Gradient

Standard gradient descent algorithms (like "step" and "momentum") use only the local approximation of the slope of the performance surface (error versus weights) to determine the best direction to move the weights in order to lower the error. Second order methods use or approximate second derivatives (the curvature instead of just the slope) of the performance surface to determine the weight update. If the performance surface is quadratic (which is only true in general for linear systems), then using a second order method can find the exact minimum in one step. The conjugate gradient method is an excellent tradeoff between speed of computation and performance.

4.4 Quick Propagation

The Quickprop implements Fahlman's quickprop algorithm. It is a gradient search procedure that has been shown to be very fast in a multitude of problems. It basically uses information about the second order derivative of the performance surface to accelerate the search.

4.5 Delta Bar Delta

Delta Bar Delta is an adaptive step-size procedure for searching a performance surface. The step size and momentum are adapted according to the previous values of the error at the PE. If the current and past weight updates are both of the same sign, it increases the learning rate linearly. The reasoning is that if the weight is being moved in the same direction to decrease the error, then it will get there faster with a larger step size. If the updates have different signs, this is an indication that the weight has been moved too far. When this happens, the learning rate decreases geometrically to avoid divergence (Jacobs, 1988).

5 Design of GFFNN classifier

For the GFFNN input processing elements (PEs) are equal to the number of input feature of EEG signal. We have used optimal 64 FFT and 11 statistical parameter of EEG signal. 75 PEs were used in the input layer and 2 PEs were used in output layer for normal and abnormal outputs. With the variation in number of hidden layers the randomized data was feed to the neural network and it was trained for N times. The weights were randomized to remove the biasing and ensure the true learning. It was observed that with single hidden layer GFFNN performed well. The number of processing elements (PEs) in hidden layer varied. The network was trained for N times and minimum MSE was obtained when 14 PEs were used in hidden layer as shown in Fig. 2. Rigorous experimentations were done by varying the number of inputs for training and cross validation (CV) dataset. The network was trained for three times with randomized weights, the minimum mean square error (MSE) and average classification accuracy were calculated for training as well as CV data. In Fig. 3 and Fig. 4 it is observed that the minimum

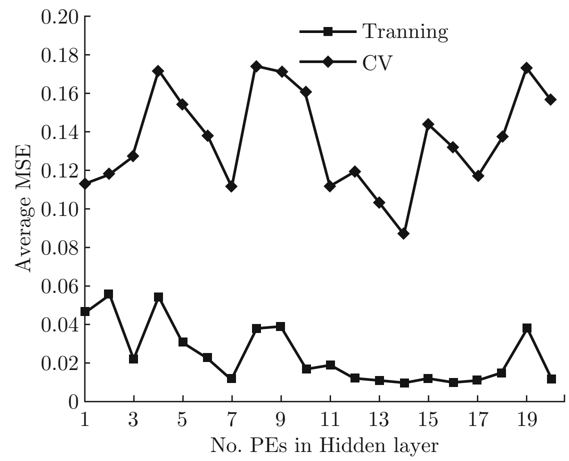

Fig. 2 Variation in average MSE with number of PEs in the hidden layer.

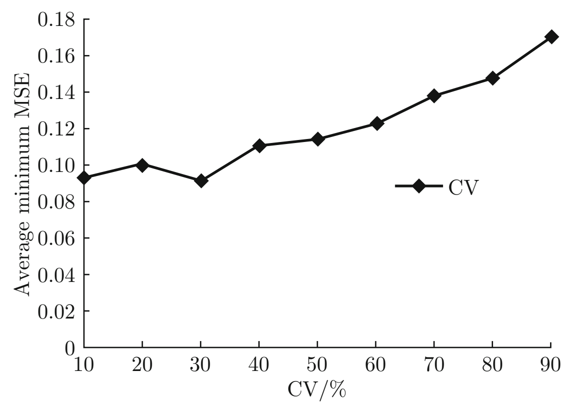

Fig. 3 Graph showing variation of average minium MSE on training and CV dataset with % CV data.

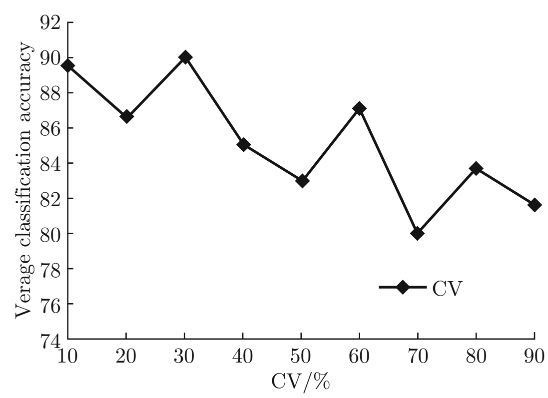

Fig. 4 Graph showing variation in % average classification accuracy with % CV data.

MSE and maximum classification accuracy is obtained when 30% data is used for CV and 70% data is used for training purpose.

Using 70% data for training and 30% for CV with transfer functions like Tanh, Sigmoid, Linertanh, Linear-sigmoid, Softmax, Bias axon, Linear axon and different learning rules viz. Momentum, Conjugate-Gradient, Quick Propagation, Delta Bar Delta, were used to train the GFFNN. Minimum mean square error (MSE) and average classification accuracy were measured and plotted in Figs. 5, 6, 7 and 8 respectively. It was found that Tanh transfer function and momentum

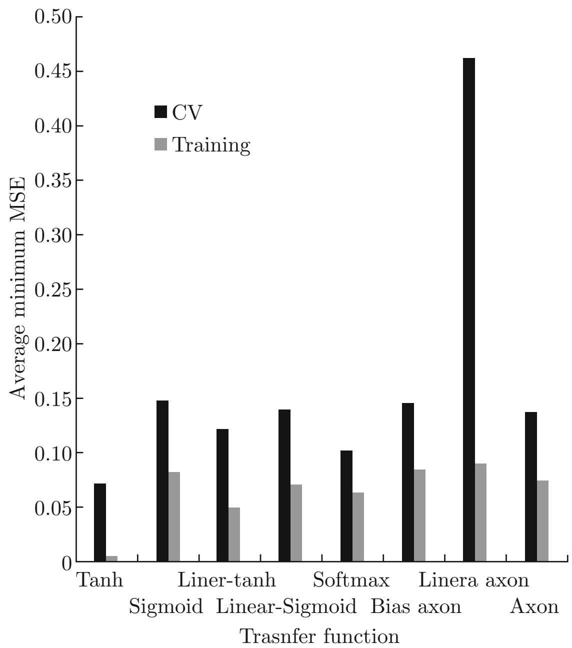

Fig. 5 Variation in average minimum MSE with transfer function.

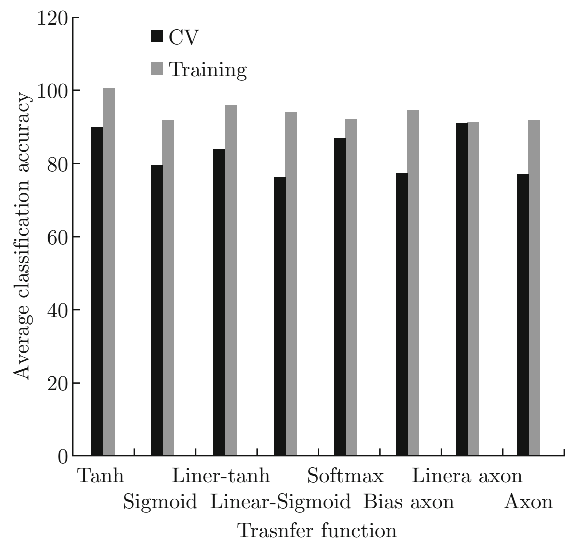

Fig. 6 Variation in average classification accuracy with transfer function.

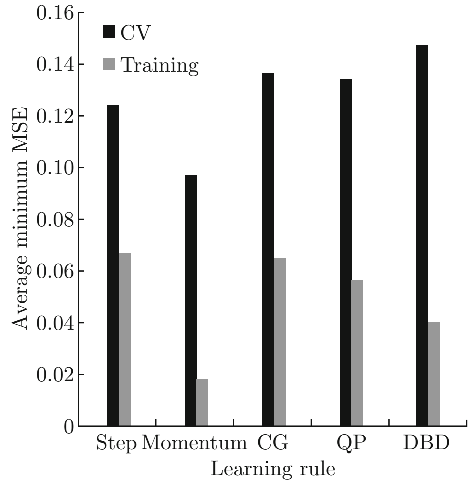

Fig. 7 Variation in average minimum MSE with learning rule.

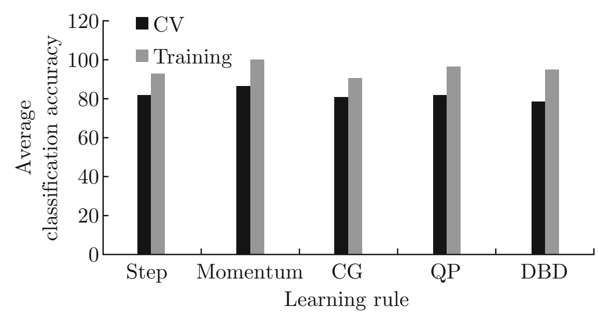

Fig. 8 Variation of average classification accuracy with learning rule.

learning rule gives the optimum result.

Time required to process the data and complexity of the neural network is an important performance parameter of any network in addition to the classification accuracy. The time elapsed per epoch per exemplar was found out to be 0.071 ms. With the above experimentation finally the GFFNN classifier is designed with the following parameter. For the experimentation the system with dual core, 1.73 GHz and 2 GB RAM configuration was used.

Number of inputs: 75
Number of hidden layers: 01
Number of PEs in hidden layers: 14
Number of PEs in output layers: 02
Hidden layer
Transfer function: Tanh
Learning rule: Momentum
Step size: 0.1
Momentum: 0.7
Output layer:
Transfer function: Tanh
Learning rule: Momentum
Step size: 0.1
Momentum: 0.7

Number of connections weights GFFNN (75-14-2): 1094

6 Results

With the rigorous experimentation and variation in parameters of neural network, the optimum DSS was designed and tested on CV and training dataset and results are portrayed in Table 1 and Table 4. The accuracy of classification was 100% for training dataset for normal as well as epileptic, and for CV dataset it was 87.5% for normal and 85.71% for epileptic. The performance of proposed DSS was measured by using the parameters like sensitivity, specificity and overall accuracy (Pradhan *et al.*, 1996). In medical diagnosis sensitivity gives the percentage of correctly classified disease individual and specificity gives the percentage of correctly classified individuals without the disease. The mathematical formulas for sensitivity, specificity and overall accuracy are given in Equations 5, 6 and 7 respectively. Table 5 shows the performance parameter for the proposed GFFNN.

$$Sensitivity = \frac{Positive\ correctly\ classified}{Total\ positive} \tag{5}$$

Table 1 Confusion matrix for training data set using GFFNN

Output / Desired	Normal	Epileptic
Normal	68	0
Epileptic	0	72

Table 2 Performance parameters for training data set using GFFNN

Output / Desired	Normal	Epileptic
MSE	0.017249	0.019038
MAE	0.089232	0.093095
% Accuracy	100	100

Table 3 Confusion matrix for CV data set using GFFNN

Output / Desired	Normal	Epileptic
Normal	28	4
Epileptic	4	24

Table 4 Performance parameters for CV data set using GFFNN

Output / Desired	Normal	Epileptic
MSE	0.105996	0.087438
MAE	0.183757	0.164396
% Accuracy	87.5	85.71

Table 5 Sensitivity, specificity and over all accuracy of GFFNN

Performance Measure	Total inputs	Correctly Classified	% result
Sensitivity	28	24	87.71
Specificity	32	28	87.5
Overall Accuracy	60	52	86.67

$$Specificity = \frac{Negative\ correctly\ classified}{Total\ negative} \tag{6}$$

$$\text{Overall accuracy} = \frac{Instances\ correctly\ classified}{Total\ instances} \tag{7}$$

7 Conclusion

This paper proposes the diagnosis of epilepsy based on GFFNN. For this purpose we have designed and tested the optimum GFFNN. The output was classified in two classes normal and epileptic. The train N time method was used to train the network. It was observed that the deviation in the percentage classification accuracy was 4% and assured accuracy of GFFNN was 86%.

References

[1] Akhil, M., Arserim, M.A., Kiymik, M.K., Turkoglu, I. 2001. A new approach for diagnosing epilepsy by using wavelet transform and neural network. EMBS International Conference, Istanbul.

[2] Andrzejak, R.G., Lehnertz, K., Mormann, F., Rieke, C., David, P., Elger, C.E. 2001. Indication of nonlinear deterministic and finite-dimensional structure in time series of brain electrical activity: Dependence on recording region and brain state. Phys Rev E 64, 061907.

[3] Arulampalam, G., Bouzerdoum, A. 2003. Generalized feedforward neura network classifier. Proceedings of the International Joint Conference on Neural Network, Portland, Oregon.

[4] Derks, E.P.P.A., Sanchez, M.S., Buydens, L.M.C. 1995. Robustness analysis of radial base function and multi-layer feed-forward neural network models. Chemo Metrics and Intelligent Laboratory System 28, 49–60.

[5] Engel, J. 1989. Seizure and Epilepsy. F. A. Davis Company, Philadelphia.

[6] Fisher, R.S., van Emde Boas, W., Blume, W., Elger, C., Genot, P., Lee, P., Engle, J. Jr. 2005. Epileptic seizure and epilepsy: Definition proposed by international league against epilepsy and the international bureau for epilepsy. Epileapsia 46, 470–472.

[7] Harikumar, R., Sabarish, N.B. 2003. Fuzzy techniques for classification of epilepsy risk level from EEG sig-

nal. Conference on Convergent Technique (TENCON), Banglor, India.

[8] Jacobs, R.A. 1988. Increased rates of convergence through learning rate adaptation. Neural Networks 1, 295–307.

[9] Pradhan, N., Sadasivan, P.K., Arunodaya, G.R. 1996. Detection of seizure activity in EEG by an artificial neural network: A preliminary study. Comput Biamed Rse 29, 303–313.

[10] Pravin, K.S., Sriram, N., Benakop, P.G. 2008. Automated detection of epileptic seizure using wavelet entropy feature with recurrent neural network classifier. Conference on Convergent Technique (TENCON), Hyderabad, India.

[11] Shrinivasan, V., Eswarann, C., Sriraam, N. 2007. Approximate entropy-based epileptic EEG detection using artificial neural network. IEEE Transaction on Information Technology in Biomedicine 11, 288–295.

[12] Shrinivasan, V., Eswaran, C., Sriraam, N. 2004. Epileptic detection using artificial neural network. International Conference on Signal Processing and Communication (SPCOM) IEEE, Bangalore, India.

[13] Szilagyi, L., Benyo, Z., Szilagyi, M.A. 2002. New method for epileptic waveform recognition using wavelet decomposition and artificial neural networks. Proceedings of Second Joint EMBS/BMES Conference, Houston, USA.