

Optimization of BPNN parameters using PSO for EEG signals

P. Kshirsagar¹, S. Akojwar²

¹Department of Electronics Engineering, RG CoER&T, Chandrapur-442403, M.S, India.

²Department of Instrumentation Engineering, Government College of Engineering, Chandrapur-442403, M.S, India.
{pravindr88@gmail.com; sudhirakojwar@gmail.com}

Abstract: Epilepsy is a neurological disorder characterized by the existence of recurring seizures. Like many other neurological disorders, epilepsy can be appraisal by the electroencephalogram (EEG). The EEG signal is highly non-linear and non-stationary and consist of lot of data including significant data and artifact and hence, it is practically arduous to characterize and interpret it. However, it is a well-established clinical technique with low associated costs for detection of various neurological disorders. In this work, we propose a methodology for the automatic detection of normal, epilepsy and brain death from recorded EEG signals collected from clinic. Discrete wavelet transform is applied for feature extraction. Back Propagation neural network optimized by particle swarm optimization is used for classification of neurological disorders. Simple BPNN has several drawbacks which mainly include large time duration during EEG signal classification. This drawback is removed by PSO. In this paper, the proposed method used to detect the number of neurons in hidden layer of BPNN using optimization technique of PSO. Once the numbers of neurons in hidden layers are detected, optimum value for initial weights and biases for BPNN estimated which is further used for classification and sortilege of various neurological disorders. So that time duration decreases and accuracy increases. EEG signals are recorded for 30 minutes of three different patients that are epileptic, normal and brain death used to rehearse and test the proposed algorithm. A signal used to test is integrated signal by taking mean of 16 channels. By applying techniques to signals of epilepsy or normal or brain death patient which are known to us we find the more accurate results with less number of iteration and time .

Keywords: *EEG Signals, Discrete Wavelet transform (DWT), Back Propagation Neural Network (BPNN), Particle Swarm optimization (PSO), and Prediction.*

1 Introduction

Temporary electrical disturbance of the brain can cause epileptic seizures. Sometimes seizures may go unnoticed, depending on their strength, and sometimes may be confused with other events, such as strokes, which can also cause falls or migraines. Unfortunately, the occurrence of an epileptic seizure seems unpredictable and it may cause even brain death. So, more research is needed for a better understanding of the mechanisms which may cause epileptic disorders. Despite rapid advances of neuro imaging techniques, EEG recordings continue to play an important role in both the diagnosis of neurological diseases and the understanding of psycho physiological processes. In order to extract relevant information from recordings of brain electrical activity, variety of computerized-analysis methods has been developed. Most of them assume that the EEG signal is generated by a highly complex linear system, which results in characteristic properties like non- stationarity and difficulty of prediction.

In the last decades, the electroencephalogram (EEG) signal was studied, which endow information about only electrical activity of the human brain. As Electrical activity of brain is directly read on computers so it becomes effective tool in clinical research. But earlier EEG method was invasive method i.e. electrodes are implanted directly in to the grey matter of the brain during neuro-surgery thereby direct neural contact via pins. As research going on, EEG becomes Non invasive. So it becomes easier to researcher perform research related to brain activity. As brain is main part of nervous system, so study of brain is useful in many neurological disorders. So EEG can convey valuable clinical information about the current which is being widely used in the study of the nervous system properties, for monitoring sleep stages, and for the diagnosis of many disorders such as epilepsy, sleep disorders, dementia and brain death.

B. Iyer, S. Nalbalwar and R. Pawade (Eds.)

ICCASP/ICMMD-2016. Advances in Intelligent Systems Research.

Vol. 137, Pp. 385-394.

© 2017- The authors. Published by Atlantis Press

This is an open access article under the CC BY-NC license (<http://creativecommons.org/licenses/by-nc/4.0/>).



Artificial Neural Networks and Particle Swarm Optimization techniques designed in the form of a hybrid structure are used for diagnosis of epilepsy, normal and brain death patients via EEG signals. Idiosyncrasy of EEG signals are needed to be determined by employing EEG signals which are recorded from clinic. By using PSO first optimizing the parameter i.e. number of neuron in the hidden layer with the mean square error during the BPNN rehearsing process such that the mean square error of prediction was minimized. PSO is used for updating the HLNN during rehearsing BPNN and a program is constituted for classification of EEG signals. Education a recording processes were performed with different parameters by means of the constituted program. Testing is performed after the rehearsing process. This hybrid structure gives the fast result as compare to the simple BPNN model and gives more accuracy. The obtained findings show that the proposed method was effective for achieving accurate results as much as possible with the use of hybrid BPNN and PSO.

2 Literature Survey

Artificial neural networks have been showed their effectiveness in many real world problems such as signal processing, pattern recognition, and classification problems. Although they provide highly generalized solutions, there were several unanswered problems detects in using artificial neural networks. Determining the most appropriate architecture of artificial neural network is identified as one of those major problems. Generally, the performance of a neural network strongly depends on the size of the network. By increasing the number of layers generalization ability can be improved. However, this solution may not be computationally optimized. On the other hand, too many hidden neurons may over-train the data and which cause the poor generalization. Also, too few neurons under-fit the data and hence, network may not train the data properly. Thus, both too many and too few neurons show bad generalization. Therefore, determining the most suitable architecture is very important in artificial neural networks. As such, a large number of researchers have been carried out to model the hidden layer architecture by using various techniques. These techniques can be categorized as pruning techniques and constructive techniques. Pruning algorithms start with an oversized network and remove nodes until the optimal architecture occurs [28], [29], [30], [31] and [32].

In [33], author has been proposed a new pruning algorithm based on backpropagation training to design the optimal neural network. The optimal solution is obtained by two steps. First, the number of hidden layers in the most efficient network is determined. Then the network tends to the optimal solution by removing all unimportant nodes from each layer. The removable nodes are identified through the delta values of hidden neurons [34], [35]. The choosing of delta values was based on the fact that the delta values of the hidden layers are used to compute the error term of the next training cycle. Hence, delta value is a significant factor in error term. Thus, the delta values are used to identify the less saliency neurons and remove them from hidden neurons so that the error term tends to the desired limit faster than the backpropagation training.

In [36], author was proposed a new algorithm for multilayer hidden architecture. The algorithm is based on a pruning technique. Hidden neurons were pruned by using the delta values of hidden neurons. The correlation between the summation of delta values of each layer at the n th training cycle and the error of the $(n+1)^{st}$ training cycle was considered to identify the less saliency neurons. Moreover, neurons with zero delta values were recognized as unimportant neurons as they do not have much effect on updating the weights. Therefore hidden neurons with small positive or large negative values (depends on the correlation) can be successfully used to reduce the size of the multilayer artificial network. In his work, a benchmark comparison done with the back propagation and it demonstrates that new approach can be used to minimize the network by maintaining the same error rate as back propagation training with lesser number of training cycles. Further, the modified architecture can be obtained with very limited computations. Generally, 5% - 30% of neurons can be removed from hidden layers without degrading the performance of the output.

Artificial Neural Network is appropriate and most effective for pattern recognition, signal processing, and classification problems. For proper result ANN requires correct preprocessing of data, selection of architecture and training for the network. But the major problem is to select the number of hidden nodes, because the random selection of number of hidden nodes may cause the problem of Overfitting and Underfitting. Sometimes the network matches the data narrowly and loses its generalization ability over the test data which give rise to overtraining. In [37], author gives information regarding the competitive learning approach of finding hidden nodes for artificial neural network. The advantage of this proposed method is that it is not approximately calculating number of hidden nodes but based on similarity between input data. Here numbers of hidden nodes are not predefined but generated during the training time.

From the experiment author conclude that as increased similarity threshold, numbers of hidden node are decreased, improves MSE (min square error).

He concludes that all the previous methods to find numbers of hidden nodes are based on approximation; where have to decide the number of hidden nodes before training the neural network.

But proposed method is not based on approximation rather it gives exact solution. Also there is no need to decide number of hidden node before training of ANN. In proposed method number of hidden nodes is generated automatically during the training of ANN, based on the value of similarity threshold.

In [38], author proposed a novel and effective criterion based on the estimation of the signal-to-noise-ratio figure (SNRF) to optimize the number of hidden neurons in neural networks to avoid over fitting in the function approximation. SNRF can quantitatively measure the useful information left unlearned so that over fitting can be automatically detected from the training error only without use of a separate validation set. It is illustrated by optimizing the number of hidden neurons in a multi-layer perceptron (MLP) using benchmark datasets. The criterion can be further utilized in the optimization of other parameters of neural networks when Overfitting needs to be considered.

In [39], author proposed to optimize the number of hidden neurons in NN to avoid over fitting in function approximation. The method utilizes a quantitative criterion based on the SNRF to detect over fitting automatically using the training error only, and it does not require a separate validation or testing set. The criterion has been validated using benchmark datasets and compared with the common cross-validation method. The criterion is very easy to apply, consumes small amount of computations and is suitable for practical application. The same principle applies to the optimization of other parameters of neural networks, including the number of iterations in back propagation training to avoid overtraining or the number of hidden layers.

So this method is presented for reducing the number of units in the hidden layers of a feed-forward neural network. Starting with a net that is oversized, the redundant units in the hidden layer is eliminated by introducing an additional cost function on a set of auxiliary linear response units. The extra cost function enables the auxiliary units to fuse together the redundant units on the original network, and the auxiliary units serve only as an intermediate construct that vanishes when the method converges. Numerical tests on the Parity and Symmetry problems illustrate the usefulness of this method in practice.

In [40], author introduced number of auxiliary units equal to that of the hidden units, and each auxiliary unit has a full set of connections to all units in the hidden layer. Both the number of auxiliary units and their connectivity can be varied, which is a feature that may be of considerable use in problems involving many hidden units. In such a case it would be possible to limit the number of auxiliary units to which each hidden unit is attached, or to associate groups of auxiliary units with groups of hidden units. This would reduce considerably the computational requirements for large problems but still enable substantial reductions in network size. It is also possible to use other types of unit-attracting terms, or even to implement the method by successively eliminating the smallest components of each vector $\vec{\omega}_{(si)}$ until only one component remains on each auxiliary node.

3 Methodology

A BPNN based prediction scheme was proposed in this study. In the proposed research, effective estimation of EEG signal is accomplished by hybrid model of BPNN and PSO. Firstly EEG feature extraction is done using wavelet transform technique to extract important features. PSO is used to optimize the number of neurons in the hidden layer with the mean square error during the BPNN rehearsing process such that the mean square error of prediction was minimized. After rehearsing process to validate the BPNN rehearsing testing was performed. This hybrid structure give the fast result compare to the simple BPNN model and give more accuracy. Finally, the output stage converts the combined result back into a specific control output value.

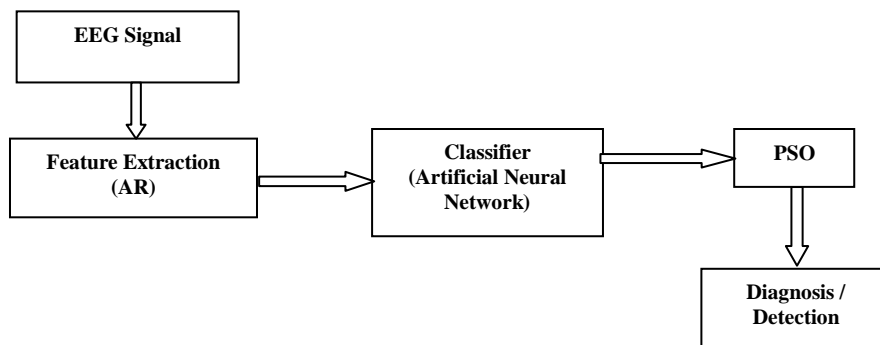


Fig.1. The design flow of the general co-design approach

3.1 Electroencephalography (EEG)

Electroencephalography (EEG) is electrophysiological monitoring method to record electrical activity of the brain which consists of billions of nerve cells called neurons. It is non – invasive method. Because of ionic current within the neurons of the brain there is voltage fluctuation and EEG read this voltage fluctuation. It is easily available tool and easy to use, that endows information about how the brain functions with respect to time. So EEG recording is one of the most useful tools for diagnosis of neurological diseases such as epilepsy, brain tumor, head injury, sleep disorder, monitoring depth of anesthesia during surgery. EEG was first introduced by ‘Hans Berger’ in 1929 that was neuropsychiatric from the University of Jena.

During EEG test, number of small discs called electrode are placed on different locations on the surface of scalp using 10-20 system of electrode placement with the help of temporary glue which is having good mechanical and electrical contact.

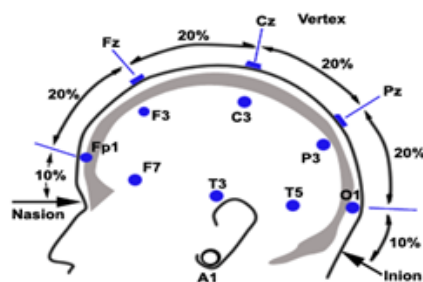


Fig.2. The international 10-20 electrode placement system

3.2 Data acquisition

In order to design PSO-BPNN hybrid algorithm, data records of the EEG signal are collected from clinic was obtained. The data were recorded at a sampling rate of 256 Hz. The original database contained records of 30 minutes. The dataset is a two-class problem either positive or negative for epileptic. Data set consist of signal with 16. Results of the hybrid model of BPNN and PSO give mean square error and optimize number of neurons in hidden Layer and final diagnosis give the result with features of unknown signal and more accuracy.

In the proposed system EEG signals are categorized into three possible cases epilepsy, normal, brain death. For testing, the input is given to the BPNN model. This BPNN model is used with the PSO which optimize the number of neurons in hidden layer and give mean square error.

3.3 Feature Extraction

EEG signals varying from 100 Hz to 1000 Hz, in order to obtain the best possible performances, it is necessary to work with a smaller number of values which describe some relevant properties of the signals. These values are known as “features”. Such features can have the power of the EEG signals in different frequency bands. Features are generally aggregated into a vector known as “feature vector”. Thus, feature extraction can be defined as an operation which transforms one or several signals into a feature vector. So there are different types for feature extraction. Neurologist found that in biological signals such as EEG signals, frequency is one of the most important parameter to analyze the brain activities. So we need Fourier transform.

As EEG signal are non stationary signals so Fourier Transform is not a suitable technique for these signals. As Fourier transform gives the spectral content of the signal, but it doesn't gives information regarding in which time these spectral components appears. Therefore, Fourier transform can be used for non-stationary signals.

Wavelet means a small wave. Wavelet analysis does not use a time-frequency region, but a time-scale region. The wavelet function $\Psi_{\tau, s}(x)$ for signal $x(t)$ is defined as

$$\Psi_{\tau, s}(x) = \frac{1}{\sqrt{s}} \Psi\left(\frac{x-\tau}{s}\right) \quad (1)$$

Where τ is called as translation coefficient and ‘s’ is called as scaling coefficient.

The discrete wavelet transform (DWT), provides sufficient information both for analysis and synthesis of the original signal, with a significant reduction in the computation time. In wavelet analysis there are approximations and details components. The approximations are the high-scale, low-frequency components of the signal. The details are the low-scale, high-frequency components. For many signals, the low-frequency content is the most important part, whereas high frequency content is not as much important as it contains noise. Here in this work upto 4th level decomposition is done. There are many different mother wavelet families. Among which Daubechies is best. In this presented work approximation components has been determined.

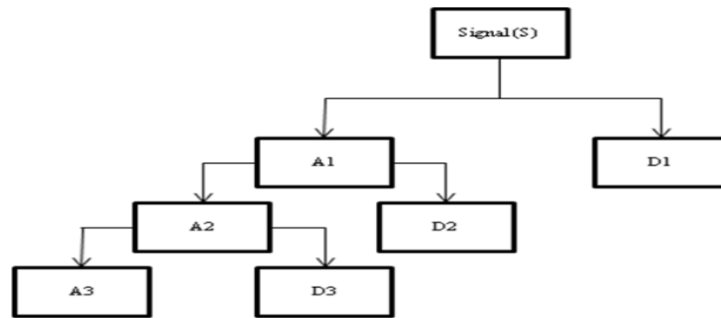


Fig .3. Multiple Level Wavelet Decomposition

3.4 Classification using Back propagation neural network

Back Propagation is the learning and the rehearsing algorithm. We have selected BPNN as Artificial neural network because BPNN is best for pattern recognition and mapping tasks.

The BP network is composed by one input layer and one or more hidden layers and one output layer. The learning process of network includes two courses, one is the input information transmitting in forward direction and another is the error transmitting in backward direction. In the forward action, the input information goes to the hidden layers from input layer and goes to the output layer. If the output of output layer is different with the wishful output result then the output error will be calculated, the error will be transmitted backward direction then the weights between the neurons of every layers will be modified in order to make the error as minimum as possible. Then the network is said to be rehearsed for the given data or application.

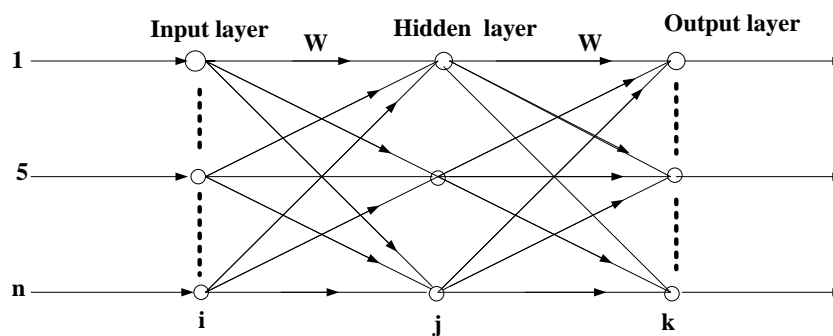


Fig.4. Structure of Back Propagation Neural Network

3.5 Optimization using Particle swarm optimization

Particle Swarm Optimization (PSO) is first described by James and Russell C. Eberhart in 1995. PSO [21] is a robust stochastic optimization technique based on the movement and intelligence of swarms. PSO applies the concept of social interaction to problem solving. It uses a number of agents (particles) that constitute a swarm moving around in the search space looking for the best solution

The PSO algorithm[22] works by simultaneously maintaining several candidate solutions in the search space. During each iteration of the algorithm, each candidate solution is evaluated by the objective function being

optimized, determining the fitness of that solution. Each candidate solution can be thought of as a particle “flying” through the fitness landscape finding the maximum or minimum of the objective function.

The velocity and position update step is responsible for the optimization ability of the PSO algorithm. The velocity of each particle in the swarm is updated using the following equation

$$v_i(t+1) = wv_i(t) + c_1 r_1 [\hat{x}_i(t) - x_i(t)] + c_2 r_2 [g(t) - x_i(t)] \quad (2)$$

Where

i is Particle index

w is Inertial coefficient

c_1, c_2 are Acceleration coefficients, $0 \leq c_1, c_2 \leq 2$

r_1, r_2 are Random values for each velocity update ($0 \leq r_1, r_2 \leq 1$) regenerated every velocity update

$v_i(t)$ is particle's velocity at time

$x_i(t)$ is particle's position at time t

$\hat{x}_i(t)$ is particle's individual best solution

$g(t)$ is swarm's best solution as of time t

4 Results

Performances of the implemented method is measured based on performance indices such as sensitivity, specificity, precision and accuracy parameters. The true positive, true negative, false positive and false negative values are calculated from the results obtained. The above four values are used to calculate performance indices as specified in the equations given below. Table 1 gives the performance evaluation for focal epilepsy, brain death, and slow wave activity and for normal waves.

The definitions of various parameters are:

Accuracy: Number of correctly classified patterns/total number of patterns.

$$\text{Accuracy} = \frac{TP+TN}{TP+FP+TN+FN} \quad (3)$$

Sensitivity: Number of correctly detected positive patterns/total number of actual positive patterns. A positive pattern indicates a detected seizure.

$$\text{Sensitivity} = \frac{TP}{TP+FN} \quad (4)$$

Specificity: Number of correctly detected negative patterns/total number of actual negative patterns. A negative pattern indicates a detected normal/non-seizure.

$$\text{Specifity} = \frac{TN}{FP+TN} \quad (5)$$

Table 1. Sensitivity and Accuracy for different disease signals

Parameter	Epilepsy	Normal	Brain Death
Sensitivity	96.33 %	90 %	100%
Accuracy	98.67 %	96.33 %	100%

Table 2. Optimized HLNN for best accuracy

Sr. No.	No of Hidden Layers	No. of Iterations	Time(sec)	Accuracy
1.	5	216	13	96.67 %
2.	7	138	5	100 %
3.	11	238	8	98.03 %
4.	15	450	26	93.33 %
5.	16	302	20	97.77 %
6.	21	298	18	95.54 %

4.1 Output of PSO

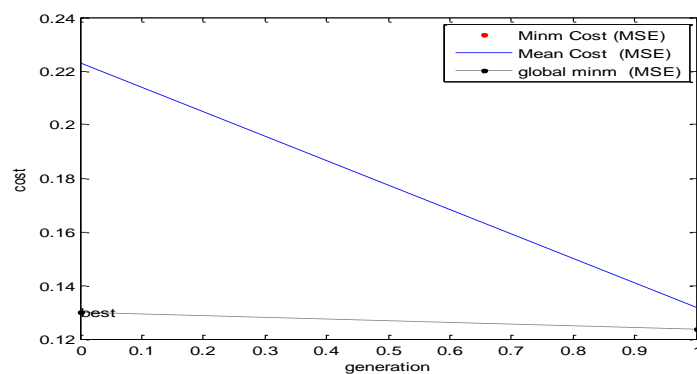


Fig.5. Graph of min. cost, mean cost and global minima

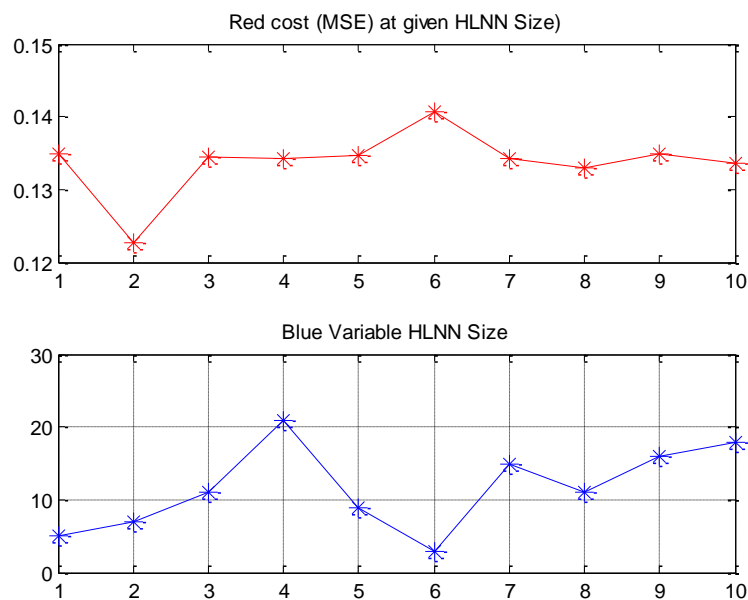


Fig.6. Graph of best HLNN Size and MSE for HLNN

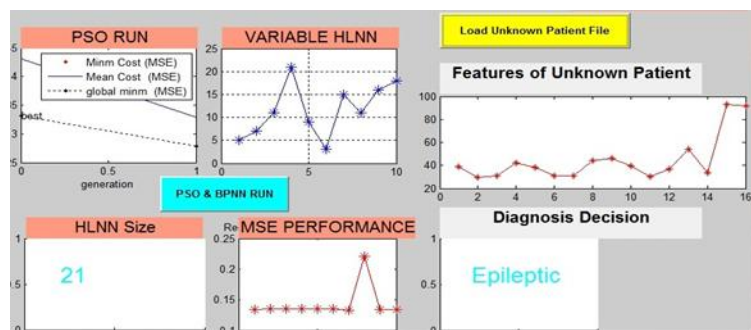


Fig.7. Diagnosis shows Epileptic patient

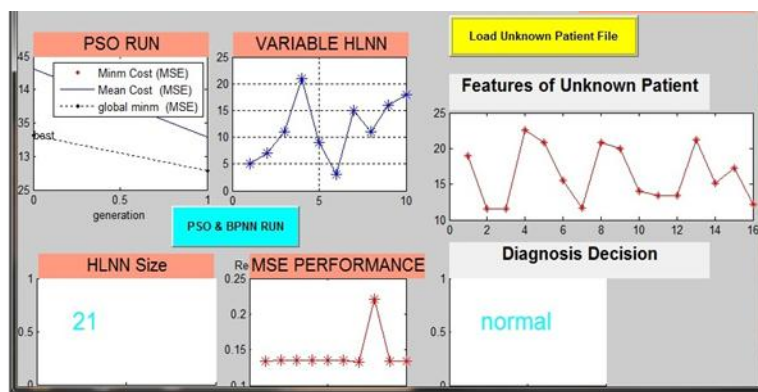


Fig.8. Diagnosis shows Normal patient

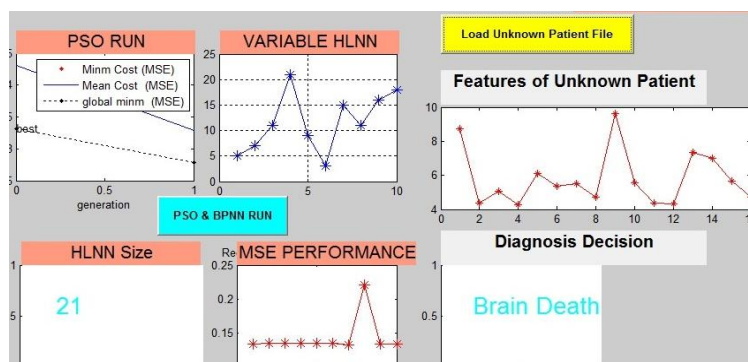


Fig.9. Diagnosis shows Brain Death patient

5 Conclusion

This paper study had been made to suggest and develop some tool which will eventually be useful to the Medical Science for easily recognize the patient by diagnosis the EEG signals with more accuracy and sensitivity in less time. This paper focused toward accuracy, sensitivity and time consumption which endow for the patients. The EEG signals of patients include normal patients, epileptic and brain death. For this we have constructed a signal consist of 16 channel EEG signal .This feature signals are given as a input to hybrid model of BPNN and PSO algorithm which give better rehearsing and testing with less time , number of neurons in hidden layer and mean square error. These factors are responsible for time consumption and accurate results. The relative importance and significance of these factors was made. This hybrid model gives rise to features of unknown patient and better diagnosis decision. The simple ANN model diagnoses the patient's signals but not optimize the parameter. This shows that ANN with PSO give perfect number of neurons in hidden layer and mean square error which give better result compare to the simple ANN model for EEG signals.

References

- [1] E. Niedermeyer, F. H. Lopes da Silva.: Electroencephalography - Basic principles, clinical applications and related fields, 3rd edition, Lippincott, Williams & Wilkins, Philadelphia, (1993).
- [2] R.D. Bickford: Electroencephalography. In: Adelman G. ed. Encyclopedia of Neuroscience, Birkhauser, Cambridge (USA), pp.371-373, (1987).
- [3] T.W. Picton et. al.: Guidelines for using event-related potentials to study cognition: Recording standards and publication criteria. Psychophysiology 37, 127-152. the Wavelet Transform" IEEE , (1997)
- [4] <http://ric.uthscsa.edu/facts/erp.html>, The Research Imaging Center, University of Texas.

- [5] A. S. Gevins, A. Rémond (eds.): *Electroencephalography and clinical neurophysiology, Methods of analysis of brain electrical and magnetic signals*, Elsevier, Amsterdam, (1987).
- [6] Alvaro Fuentes Cabrera, Dario Farina , Kim Dremstrup.: Comparison of feature selection and classification methods for a brain–computer interface driven by non-motor imagery, *International Federation for Medical and Biological Engineering*, Springer, (2009)
- [7] R.sobhani janbehsaraei , M.R.daliri , R. ebrahimpour: Feature Vectors Combination to Increase the Efficiency of a Brain-Computer Interface (BCI) System, *Journal of Basic and Applied Scientific Research*, Janbehsaraei et al , (2013).
- [8] Mai S. Mabrouk.: Non- Invasive EEG-based BCI system for Left or Right Hand Movement, *Majlesi Journal of Electrical Engineering*, Vol. 5, No. 3, (2011)
- [9] Mythra H V, Veenakumari H M, Sanjeev Kubakaddi: Multi-Class EEG Classification for Brain Computer Interface, *International Journal of Scientific & Engineering Research*, Volume 4, Issue 9, (2013)
- [10] Wu Ting a, Yan Guo-zheng a, Yang Bang-hua a, Sun Hong: EEG feature extraction based on wavelet packet, Elsevier, (2007).
- [11] Francesc Benimeli and Ken Sharman .,Electroencephalogram signal classification for brain computer interfaces using wavelets and support vector machines, *ESANN'2007 proceedings - European Symposium on Artificial Neural Networks Bruges (Belgium)*,pp 25-27 , (2007)
- [12] M . Rajya Lakshmi, Dr. T. V. Prasad , Dr. V. Chandra Prakash: Survey on EEG Signal Processing Methods, *International Journal of Advanced Research in Computer Science and Software Engineering* 4(1), pp. 84-91, (2014)
- [13] Sam Darvishi and Ahmed Al-Ani: Brain-Computer Interface Analysis using Continuous Wavelet Transform and Adaptive Neuro-Fuzzy Classifier, *Proceedings of the 29th Annual International Conference of the IEEE*, pp. 23-26, (2007).
- [14] Elif Derya Ubeyli : Combined neural network model employing wavelet coefficients for EEG signals classification, Elsevier, (2008)
- [15] Ericka Janet Rechy-Ramirez and Huosheng Hu : Stages for Developing Control Systems using EMG and EEG Signals: A survey, *Technical Report: CES-513* ,(2011)
- [16] Amir Ahangi, Mehdi Karamnejad, Nima Mohammad, Reza Ebrahimpour, Nasoor Bagheri: Multiple classifier system for EEG signal classification with application to brain–computer interfaces, *Neural Comput & Applic.*, Springer-Verlag London Limited , (2012)
- [17] Maan M. Shaker: EEG Waves Classifier using Wavelet Transform and Fourier Transform, *International Journal of Biological, Biomedical and Medical Sciences*, www.waset.org Spring ,(2006).
- [18] Vijay Khare, Jayashree Santhosh, Sneh Anand, Manvir Bhatia.: Performance comparison of three artificial neural network methods for classification of electroencephalograph signals of five mental tasks, *J. Biomedical Science and Engineering*, 3, 200-205 ,(2010).
- [19] Girisha Garg, Shruti Suri, Rachit Garg, Vijander Singh: Wavelet Energy based Neural Fuzzy Model for Automatic Motor Imagery Classification, *International Journal of Computer Applications*, Volume 28– No.7, (2011)
- [20] Jerry J. Shih, MD; Dean J. Krusienski, PhD; and Jonathan R. Wolpaw, MD: *Brain-Computer Interfaces in Medicine*, Mayo Foundation for Medical Education and Research, (2012)
- [21] Dennis J. McFarland, Jonathan R. Wolpaw: *Brain-Computer Interfaces for Communication and Control*, NIH Public Access, *Commun ACM*. Author manuscript; available in PMC ,(2011)
- [22] Pravin Kshirsagar and Dr. Sudhir Akojwar, Novel Approach for Classification and Prediction of Non Linear Chaotic Databases, In the proceeding of International Conference on Electrical, Electronics, and Optimization Techniques, March (2016).
- [23] Pravin Kshirsagar and Dr. Sudhir Akojwar, Prediction of Neurological Disorders using Optimized Neural Network, In the proceeding of International Conference on signal processing, Communication, Power and Embedded System ,October (2016).
- [24] Jessy Parokaran Varghese: *Analysis of EEG Signals For EEG-based Brain-Computer Interface*, School of Innovation, Design and Technology Mälardalen University Vasteras, Sweden, (2009)
- [25] Catherine Chesnutt, B.S.: *Feature Generation of EEG Data Using Wavelet Analysis*, Texas Tech University, (2012)
- [26] Fabien Lotte: *Study of Electroencephalographic Signal Processing and Classification Techniques towards the use of Brain-Computer Interfaces in Virtual Reality Applications*, (2009)
- [27] Lukas Vareka: *Methods for Signal Classification and their Application to the Design of Brain-Computer Interfaces*, University of West Bohemia in Pilsen, (2013).
- [28] Piotr Szachewicz : *Classification of motor imagenary for brain – computer interfaces*, Poznan University of Technology, (2013)

- [29] Le Cun Y., Denker, Solla S. A: Optimal Brain Damage, Advances in Neural Information Processing Systems D.S. Touretzky, Ed, San Mateo CA: Vol 2 pp 598-605 (1990)
- [30] Hassabi . B., Stork. D. G: Second Order Derivatives for Network Pruning: Optimal Brain Surgeon, Neural Information Processing Systems-vol 5, (1993).
- [31] Giovanna C., Anna M. F., Marcello P: An Iterative Pruning Algorithm for Feedforward Neural Networks, IEEE Transactions on Neural Networks, vol. 8, pp. 519-531, (1997).
- [32] Faisal Muhammad Shah, Md. Khairul Hasan, Mohammad Moinul Hoque, Suman Ahmmed: Architecture and Weight Optimization of ANN Using Sensitive Analysis and Adaptive Particle Swarm Optimization, IJCSNS International Journal of Computer Science and Network Security, vol. 10, no. 8, (2010).
- [33] A. N. Burkitt: Optimization of the Architecture of Feed-forward Neural networks with hidden layers by Unit Elimination, Complex Systems 5, pp. 371-380, (1991).
- [34] Clerk Maxwell, a Treatise on Electricity and Magnetism, 3rd ed., vol. 2. Oxford: Clarendon, pp.68-73 (1892).
- [35] B. G. H. Don R. Hush: Progress in Supervised neural networks, IEEE Signal Processing, vol. 10, pp. 8-39, (1993).
- [36] N. M. Wagarachchi, A. S. Karunananda: Mathematical Modeling of Hidden Layer Architecture in Artificial Neural Networks , International Proceedings of Computer Science and Information Technology, Power and Energy Systems II, Vol.56, pp154-159, (2012)
- [37] N. M. Wagarachchi, A. S. Karunananda: Optimization of Multi-layer Artificial Neural Networks Using Delta Values of Hidden Layers, IEEE Symposium on Computation Intelligence, Cognitive Mind and Brain , pp. 80-86, (2013)
- [38] N. M. Wagarachchi, A. S. Karunananda: A Novel Technique for Optimizing the Hidden Layer Architecture in Artificial Neural Networks, American International Journal of Research in Science, Technology, Engineering & Mathematics, 4(1), September- November, 2013, pp. 01-06
- [39] Foram Panchal, Mahesh Panchal: Optimizing Number of Hidden Nodes for Artificial Neural Network using Competitive Learning Approach, IJCSMC, Vol. 4, Issue. 5, May 2015, pg.358 – 364
- [40] Yinyin Liu, Janusz A. Starzyk, Zhen Zhu: Optimizing number of hidden neurons in neural networks
- [41] Anthony N . Burkitt,: Optimization of the Architecture of Feed-forward Neural Networks with Hidden Layers by Unit Elimination, Complex Systems 5 (1991) 371-380