Middle-East Journal of Scientific Research 22 (6): 816-821, 2014
ISSN 1990-9233

DOI: 10.5829/idosi.mejsr.2014.22.06.21952

A New Classification Method of Epileptic Eeg Signals Using Differential Evolution Optimally Pruned Extreme Learning Machine

[1]K.G. Parthiban and [2]Dr. S. Vijayachitra

[1]Research Scholar, Anna University, Chennai, Tamilnadu, India
[2]Department of Electronics and Instrumentation Engineering, India
Kongu Engineering College Perundurai, Erode, Tamilnadu, India

Abstract: An epileptic seizure is a transient event of symptoms due to abnormal neuronal action in the brain. Electroencephalography (EEG) is the neuro physiological measurement of electrical activity in the brain as recorded by electrodes placed in the cerebral cortex. An epilepsy EEG is based on three approaches. First, a scaling and wavelet function of the Multi Wavelet Transform (MWT) offers orthogonality and symmetry. Second, Feature Extraction reduces the dimensionality. Finally, Extreme Learning Machine (ELM) is utilized to train a single hidden layer of feed forward neural network (SLFN) features for classification of EEG signal. In this paper, a proposed method introduces a machine learning algorithm referred to as Differential Evolution Optimally Pruned Extreme Learning Machine (DE-OPELM). The DE-OPELM is utilized to detect and classify the epileptic EEG in the diagnosis of patients with epilepsy. The DE-OPELM is applicable to high-dimensional complex optimization problems. Hence, the DE-OPELM is accurate than ELM for automatic epileptic seizure detection. For inter-ictal EEG and ictal EEGs, the DE-OPELM produces 98.33% testing accuracy. As compared with other learning machines, the DE-OPELM detects only 0.0007sec of time.

Key words: Multi Wavelet Transform (MWT) · Root Mean Square (RMS) · Extreme Learning Machine (ELM) · Differential Evolution Optimally Pruned Extreme Learning Machine (DE-OPELM)

INTRODUCTION

Epilepsy [1] is a group of long-term neurological brain disorder characterized by epileptic seizures. In which clusters of neurons in the brain sometimes signal abnormally. Neurons typically generate electrochemical impulses that follow up on other neurons, organs and muscles to make human thoughts, feelings and actions. In seizure [2], the pattern of neuronal activity becomes abnormal behavior, disturbed, or sometimes convulsions and loss of consciousness. During a seizure, neurons may fire 500 times per second; it is much faster than normal. This happens infrequently in few people; but for others, it may happen up to several times a day. Mostly the cause is unknown. Even though compare with others, some people mostly get epilepsy as the effect of brain injury, brain cancer and drug misuse. The diagnosis generally involves ruling out different conditions that might cause such as syncope and also evaluating whether any immediate causes occur. Frequently, Electroencephalography (EEG) has been an important component in the confirmation of epilepsy. The EEG [3] gives essential information about EEG background and epileptic form discharges. The EEG [4-5] is required for the diagnosis of specific electro-clinical syndromes. Research on seizure detection exhibited in different techniques. To describe effectively one of the techniques is the Fourier transform (FT). The FT [6] is a tool utilized for more scientific functions and well suited only for stationary signals only. If all frequencies have an infinite coherence time, FT cannot be functional to non-stationary signals. To resolve some of the problems associated with non-stationary signals, the modified version of the Fourier transform introduced short-time Fourier transform (STFT). EEG signals are non-stationary signals. In [7-8], the STFT analysis of EEG signals and its extracted features depend upon the pseudo Wigner–file distribution. So, those features were utilized as inputs to an artificial neural network (ANN) for classifications. But the main drawback

Corresponding Author: K.G. Parthiban, Anna University, Chennai, Tamilnadu, India.

of the STFT that cannot be addressed all issues of concern. This drawback can be made processed effectively to utilize wavelet transforms in EEG. As opposed to fixing the time and the frequency resolutions, both time-frequency resolutions keep in order to acquire a multi resolution analysis.

In [8-9], Wavelet transforms has developed as one of the best techniques in analyzing non-stationary signals like EEG. Its potential in transforming a time domain signal into time and frequency domains helps to know the behavior of a signal. The EEGs were decomposed with wavelet transform into distinctive sub-groups and some statistical information extracted from the wavelet coefficients. To decompose an EEG signal, Multi Wavelet Transform (MWT)) is applied to the components of the EEG signal resolution levels (δ, θ, α, β and γ). The popular EEG features are Root Mean Square (RMS) and Mean Absolute Value (MAV) [10-11] for dimensionality reduction.

To classify in different medical applications, the MWT and neural network (NN) are utilized in the signal processing. Most of the classifiers are in both linear and nonlinear resolutions, such as, Artificial neural network (ANN), support vector machine (SVM) and Linear Discriminate Analysis (LDA). Neural networks have been trained to perform complex functions in various fields. Particularly, Artificial Neural Networks (ANNs) are utilized for complex pattern recognition and classification tasks. In ANNs, knowledge about the difficulty is allocated in neurons and its connection weights of links between the neurons. ANNs are formed of cells simulating the low-level functions of biological neurons. Back Propagation Neural Networks (BPNNs) and Linear Discriminate Analysis (LDAs) have been demonstrated to enhance the classification accuracy compared with linear techniques. Support vector machines (SVM) [12] are the most succeeded and utilized algorithms to model the EEG characteristics. However, execution time for SVMs training process is more. The main weakness of these techniques is that they do not work well for nonlinear problems. To accurate and fast training process without compromising in theoretically and experimentally, a new classification algorithm referred to as extreme learning machine (ELM) algorithm.

ELM [13-14] algorithm challenges all kinds of conventional learning methods and theories. The reason is that ELM [15-16] arbitrarily chooses the input weights and bias of the SLFNs as opposed to tuning. ELM is proposed by Huang, which is a single-hidden-layer feed-forward network (SLFN) and which could be utilized within regression and classification issues. But it can be achieved an execution performance related to generalization comparable to other accurate yet costly learning techniques. In order to improve the robustness of ELM we introduced a new technique that Differential Evolution Optimally Pruned Extreme Learning Machine (DE-OPELM). First of all, OPELM [17-18] is attained a three stage methodology which includes steps for hidden neurons fast ranking and model selection. Then, Differential Evolution (DE) [19] is based on the differences of randomly sampled pairs of solutions in the population. Hence, DE is used to solve the problem of model selection in OPELM to attain better performance as well as faster convergence with easier structure. The Epilepsy EEG techniques consist of three steps, such as, MWT, feature extraction and DE-OPELM, shown in Figure 1. MWT decomposed the EEG data to generate sub band signals to reduce noise, then for each sub band signal is extracted by using the feature extraction to reduce the dimensionality. And finally, the proposed technique DE-OPELM is used for classifying the EEG data into normal/non-epileptic seizure and epileptic seizure.

The organized sections as follows: Introduction about EEG and data collection is presented in section II. Section III describes MWT and its feature extraction. Section IV describes the proposed technique DE-OPELM algorithm. The discussion of the proposed technique and its results discussed in section V and VI respectively. Section VII summarizes the conclusion of the paper.

EEG Data Processing: Most number of brain disorders is diagnosed by visual check up of EEG signals. An EEG records electrical motion along the scalp and measures resulting voltage variations from ionic current flows within the neurons of the brain. In 1875, Richard Caton recorded EEG signal on animal brain. In 1929, Hans Berger recorded EEG signal on human brain. The EEG records different types of brain waves by their different frequency ranges, which is any type of brain activity. The brain waves are called alpha (α), beta (β), gamma (γ), delta (δ) and theta (θ). In 1929, Berger introduced alpha (α), beta (β). The frequency ranges alpha (α), beta (β) is 8-12Hz and 12-38Hz respectively. In 1936, Walter introduced delta (δ) in the range of 0.5-3Hz to designate all frequencies below the range of alpha (α). In 1938, Jasper and Andrews introduced gamma (γ) to refer to the waves in the range of 38-42 Hz. In 1944, Wolter and Dovey introduced theta (θ) within the range of 4-7.5Hz frequencies.

Fig. 1: Block Diagram of Proposed Method

Multi Wavelet Transform: A multi wavelet transform is characterized by multi-scaling functions and multi wavelet functions. The multi-scaling function $\phi(t)$ is defined as $\phi(t) = [\phi_1(t), \phi_2(t).....,\phi_r(t)]^T$. And for set of wavelet functions, the multi wavelet function $\psi(t)$ is defined as $\psi(t) = [\psi_1(t), \psi_2(t),....,\psi_r(t)$. While r is arbitrarily large, When r = 1, $\psi(t)$ is a scalar wavelet. The multi wavelets studied primarily for r = 2. The multi wavelet two-scale equations for scalar wavelets

$$\phi(t) = \sqrt{2}\ \sum_{m=-\infty}^{\infty} L_m \phi(2t-m), \tag{1}$$

$$\psi(t) = \sqrt{2}\ \sum_{m=-\infty}^{\infty} H_m \phi(2t-m) \tag{2}$$

Here $\{L_m\}$ and $\{H_m\}$ are the matrix of low pass filters and high pass filters, i.e. $\{L_m\}$ and $\{H_m\}$ are r x r matrices for every integer value m. Equations (1) and (2) can be illustrated the low-pass filter and high-pass filter coefficients. The MWT decomposition algorithm is written as:

$$c_{j-1,m} = \sum_n L_n c_{j,2m+n} \tag{3}$$

$$d_{j-1,m} = \sum_n H_n d_{j,2m+n} \tag{4}$$

The Equations (3) and (4) can be illustrated as the multi wavelet filter bank. $[c_{1,0,m}, c_{2,0,m}]^T$ is the multi wavelet system's initial expansion coefficients input signal vector. Four output streams are generated after the input signal vector is filtered. So, each of streams is down sampled by a factor of two. Then $[c_{1,-1,m}, c_{2,-1,m}]^T$ represent the multi scaling coefficient of the low-pass sub-band signal vector. And $[d_{1,-1,m}, d_{2,-1,m}]^T$ represent the multi wavelet coefficient of the high-pass sub-band signal vector. For high level decomposition, the same process repeats on the first level multi scaling coefficient vector to get the next level multi scaling and multi wavelet coefficients.

Feature Extraction

Root Mean Square (RMS): The RMS of EEG signal amplitude (alpha, beta and gamma) is measured to determine the power changes of the EEG signal. And it can be defined as;

$$RMS = \sqrt{\frac{1}{N}\sum_{n=1}^{N} x_n^2} \tag{5}$$

Here x_n represents the n^{th} sample of the EEG signal and N denotes the length of Ethe EEG signal window segment

Differential Evolution Optimally Pruned Extreme Learning Machine (De-Opelm): Principally, the Differential Evolution Optimally Pruned Extreme Learning Machine (DE-OPELM) methodology performs better than the ELM. For fast performance and lead to extremely accurate, first of all, construct an ELM, then exact positioning neurons are performed in the hidden layer and finally based on precise Leave-One-Out (LOO) error estimation technique decides how many neurons pruned. Differential Evolution is one of the powerful evolution computation algorithm and much simpler to implement for global numeric optimization. DE keeps a population of target vectors called members. Each target vector is a set of input weights and biases of hidden vectors, defined as;

$$\theta = \begin{bmatrix} w_{11}, w_{12}, \dots, w_{1d}, w_{21}, w_{22}, \dots, w_{2d}, \dots, w_{N1}, \\ w_{N2}, \dots, w_{Nd}, b_1, b_2, \dots, b_N \end{bmatrix} \tag{6}$$

Mutation, Crossover and Selection steps enhance the population generation. The main advantage of this algorithm is in dividing the computational time by hundreds and making the learning process of the neural network rather simplistic. It has been shown that the DE-OPELM method has advantages as compared with ELM. 1. A nonlinear model computational times close to the linear models. 2. Convergence is faster 3. The less no. of neurons needed to get the lowest LOO error. 3. Enhanced generalization performance with compact network 4. Less Complexity

Proposed Algorithm:

Step1: The input vector is $X = (x_1,x_2,x_3,...,x_d)^T \in R^d$, $W = (w_{i1}, w_{i2}, w_{i3},...,w_{id})$ is the input weight vector with the i^{th} hidden and the input nodes, b_i is the bias of i^{th} hidden node. The output weight vector connecting with output node is $\beta = (\beta_1, \beta_2, \beta_3,...\beta_N)^T$. The corresponding input observation value t_i. H is the hidden layer output matrix.

$$G_N(X) = \sum_{i=1}^{N} \beta_i g(w_i . X + b_i)$$

Set of distinct training data $\{(X_i^*, t_i^*)\}_{i=1}^{n} \subset \mathbb{R}^d \times \mathbb{R}$, such that;

$$\sum_{i=1}^{N} \beta_i g(w_i . X + b_i) = t_i, \quad j = 1,2,\dots,n.$$

$H\beta = T$

$$H = \begin{bmatrix} g(W_1.X_1 + b_1), & \dots, & g(W_N.X_1 + b_N) \\ \vdots & \ddots & \vdots \\ g(W_1.X_n + b_1), & \dots & g(W_N.X_n + b_N) \end{bmatrix}_{n\times N}$$

$$\beta = \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_N \end{bmatrix}_{N\times 1} \text{ and } T = \begin{bmatrix} t_1 \\ \vdots \\ t_n \end{bmatrix}_{n\times 1}$$

According to Huang's ELM,

$$\hat{\beta} = H^{\dagger} T$$

$H^{\dagger}$ is the Moore-Penrose generalized inverse of the hidden layer output matrix H.

Step 2: To position the hidden neurons according to their accuracy, the Multi Response Sparse Regression (MRSR) algorithm is useful, which is not only a simplification of the least angle regression (LARS) algorithm, but also able to locate an exact positioning for linear problems.

Step3: The prediction sum of squares (PRESS) provides subsequent closed expression for LOO error $P = (H^TH)^{-1}$.

$$\varepsilon_i^{PRESS} = \frac{y_i - h_i b_i}{1 - h_i P h_i^T}$$

The minimized LOO error $L = \operatorname{argmin}_{j\in\{1,\dots,N\}} \sum_{i=1}^{j} \varepsilon_i^{PRESS}$

Step 4: Based on Eq. (6), the initial generation target vector is $\theta_{i,0} \in R^D$, i = 1,...,NP; mutually exclusive integers r1, r2, r3 are randomly chosen in the range of [1, NP]. F is a scaling parameter.

The mutant vector V_i generates mutation step can be followed as;

$$V_{i,Gen+1} = \theta_{r1,Gen} + F \cdot (\theta_{r2,Gen} - \theta_{r3,Gen})$$

Step 5: The j^{th} evaluation of a uniform random number generator b(j), the crossover constant CR and randomly selected integer rnbr(i) ensures at least one parameter from V_i. The trial vector q_{ij} is formed.

$$q_{ij,Gen+1} = \begin{cases} V_{ij,Gen+1} & \text{if rand } b(j) \le CR \text{ or } j = rnbr(i) \\ \theta_{ij,Gen+1} & \text{if rand } b(j) > CR \text{ and } j \ne rnbr(i) \end{cases}$$

Step 6: The target vector represents the regularization and the objective function $g(V_i)$. The new generation $\theta_{i,Gen+1}$ is,

$$\theta_{i,Gen+1} = \begin{cases} q_{i,Gen} & \text{if } g(q_{i,Gen}) \le g(\theta_{i,Gen}) \\ \theta_{i,Gen} & \text{if } g(q_{i,Gen}) > g(\theta_{i,Gen}) \end{cases}$$

The correspond minimum objective function of the optimal vector is $g(\theta_{i,Gen}^*) = C \cdot L$

Step 7: Go to step4 until the maximum learning period is reached.

RESULT AND DISCUSSION

The accuracy, sensitivity and specificity of DE-OPELM, ELM, SVM and ANN techniques are shown in the Table 1. While using the sigmoid functions for ictal and inter-ictal EEG signals, the DE-OPELM classification results the highest classification accuracy (98.33%), sensitivity (96.80%), specificity (98.97%) and processing time (0.0007sec). In Table 2., the output of hidden layers: sigmoid and radial basis activation functions train and testing network results.

The accuracy, sensitivity and specificity of learning machine parameters are determined in (7) (8) (9) respectively.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{7}$$

$$Sensitivity = \frac{TP}{TP + FN} \tag{8}$$

$$Specificity = \frac{TN}{TN + FP} \tag{9}$$

where,

TP - Patients with the characteristic and tested positive.
FN - Patients with the characteristic and tested negative.
FP - Patients without the characteristic and tested positive.
TN - Patients without the characteristic and tested negative.

Table 1: EEG Classification Results With Different Machine Learning Methods

Learning Machine	Accuracy (%)	Sensitivity (%)	Specificity (%)	Processing Time (SEC)
DE-OPELM	98.33	96.80	98.97	0.0007
ELM	96.85	95.40	98.29	0.0011
SVM	87.85	85.24	90.46	1.8200
ANN	95.14	93.81	96.47	21.7200

Table 2: Performances of the DE-OPELM for Different Activation Functions

Activation Function	Training Accuracy %	Training Process Time (Sec)	Testing Accuracy %	Testing Process Time (Sec)	Average RMSE
Sigmoid	98.16	0.280	98.33	0.0007	0.125
Radial Basis	97.95	0.0608	95.12	0.0288	0.164

CONCLUSION

Epilepsy is detected more accurately by using DE-OPELM. ELM algorithm is utilized for classification. DE-OPELM is utilized to enhance the robustness of ELM. And it can be achieved through the initial SLFN, fast positioning of hidden neurons, PRESS, mutation, crossover and selection steps. These steps are performed fast accurate to detect the epilepsy from EEG signals. And DE-OPELM is achieved better accuracy (98.33%), better sensitivity (96.80%), good specificity (98.97%) and consumes less processing time (0.0007sec) as that of ELM, SVM and ANN methods.

REFERENCES

1. Robert Fisher, 1997. 2006. 2010. Overview of Epilepsy 1-56, copyright www.epilepsy.com.
2. Fisher, R., W. Van Emde Boas, W. Blume, C. Elger, P. Genton, P. Lee and J. Engel, 2005. Epileptic seizures and epilepsy Epilepsia, 46(4): 470-472.
3. Ince, N.F., F. Goksu, A.H. Tewfik and S. Arica, 2006. A Space-Time-Frequency Analysis Approach For the Classification Motor Imagery EEG Recordings in a Brain Computer Interface Task IEEE EMBS Annual International Conference, pp: 2581-2584.
4. Ana Filipa, Lopes and Jose Paulo Monteiro, 2014. Memory Functioning in Children with Epilepsy: Frontal Lobe Epilepsy, Childhood Absence Epilepsy and Benign Epilepsy with Centrotemporal Spikes. Hindawi Publishing Corporation Behavioural Neurology, pp: 1-8.
5. Catarina A. Guimarães and Li Li Min, 2006. Memory Impairment in Children with Temporal Lobe Epilepsy: a Review, Journal of Epilepsy and Clinical Neurophysiology, 12: 22-25.
6. Gao, R.X. and R. Yan, 2011. From Fourier transform to wavelet transform text book of wavelets Theory and Applications for Manufacturing, springer.
7. Mohseni, H., A. Maghsoudi, M. Kadbi, J. Hashemi and A. Ashourvan, 2006. Automatc detection of epileptic seizure using time–frequency distributions, IET 3rd international conference on advances in Medical, Signal and Information Processing, pp: 1-4.
8. Alexandros T. Tzallas, Markos G. Tsipouras and Dimitrios I. Fotiadis, 2009. Epileptic Seizure Detection in EEGs Using Time–Frequency Analysis IEEE Transactions on Information Technology in Biomedicine, 13(5): 703-710.
9. Jahankhani, P., V. Kodogiannis and K. Revett, 2006. EEG signal classification using wavelet feature extraction and neural networks, IEEE John Vincent Atanasoff 2006 International Symposium on Modern Computing, pp: 120-124.
10. Oskoei, M.A. and H. Hu, 2007. Myoelectric control systems - a survey Biomed. Signal Process. Control, 2(4): 275-294.
11. Naik, G.R., D.K. Kumar and S.P. Arjunan, 2010. Pattern classification of Myo-Electrical signal during different Maximum Voluntary Contractions: A study using BSS techniques - Meas. Sci. Rev., 10(1): 1-6.
12. Epileptic Seizure Detection for Multichannel EEG Signals with Support Vector Machines IEEE International Conference on Bioinformatics and Bioengineering (BIBE), pp: 39-43.
13. Enrique Romeroa and Rene Alquezar, 2012. Comparing error minimized extreme learning machines and support vector sequential feed-forward neural networks, ELSEVIER Neural Networks, 25: 122-129.
14. Erik Cambria and Guang-Bin Huang, 2013. Extreme Learning Machines, IEEE Intelligent Systems, 28(6): 30-59.
15. Li-Chen Shi and Bao-Liang Lu, 2013. EEG-based vigilance estimation using extreme learning machines, ELSEVIER Neuro Computing, pp: 135-143.

16. Huang, G.B., Q.Y. Zhu and C.K. Siew, 2004. Extreme learning machine: a new learning scheme of feed forward neural networks, in: Proceedings of the International Joint Conference on Neural Networks, pp: 985-990.
17. Tiago Matias and Rui Araujo, 2013. Genetically Optimized Extreme Learning Machine, IEEE Conference on Emerging Technologies & Factory Automation, pp: 1-8.
18. Yoan, Miche and Patrick Bas, 2008. A methodology for Building Regression Models using Extreme Learning Machine: OP-ELM, European Symposium on Artificial Neural Networks (ESANN), pp: 247-252.
19. Karaboga, Dervis, 2004. A Simple and Global Optimization Algorithm for Engineering Problems: Dierential Evolution Algorithm, Turkish Journal of Electrical Engineering & Computer Sci., 12(1): 53-60.