Performance Analysis of Extreme Learning Machine for Robust Classification of Epilepsy from EEG Signals

R. Harikumar,Ph.D
Professor,
Department of ECE,
Bannari Amman
Institute of Technology,
Sathyamangalam.

C.Ganeshbabu,Ph.D
Professor,
Department of ECE,
Bannari Amman
Institute of Technology,
Sathyamangalam.

M. Balasubramani,
Research Scholar,
Department of ECE,
Bannari Amman
Institute of Technology,
Sathyamangalam.

G.A. Nivhedhitha,
PG Scholar,
Department of ECE,
Bannari Amman
Institute of Technology,
Sathyamangalam.

ABSTRACT
Epilepsy is a common brain disorder that affects one out of hundred patients. EEG (electroencephalogram) is a signal that represents that effect of the superimposition of diverse processes in the brain. This paper investigates the possibility of Extreme Learning Machine (ELM) as a classifier for detecting and classifies the epilepsy of various risk levels from the EEG signals. The Singular Value Decomposition (SVD) is used for dimensionality reduction. Twenty patients are analysed in this study.

Keywords:
Extreme Learning Machine, Singular Value Decomposition, Epilepsy risk level, seizure.

1. INTRODUCTION
Epilepsy is a common and diverse set of chronic neurological disorders characterized by seizures. It is a paroxysmal behavioural spell generally caused by an excessive disorderly discharge of cortical nerve cells of brain and can range from clinically undetectable (electrographic seizures) to convulsions. Epileptic seizures result from abnormal, excessive or hyper synchronous neuronal activity in the brain. About 50 million people worldwide have epilepsy, and nearly 80% of epilepsy occurs in developing countries. Epilepsy is marked by the term "epileptic seizures".

During a seizure, neurons may fire as many as 500 times a second, much faster than normal. In some people, this happens only occasionally; for others, it may happen up to hundreds of times a day. About 25 to 30 percent of people with epilepsy still continue to experience seizures even with the best available treatment. The most common way to interfere the epilepsy is to analysis the EEG (electroencephalogram) signal which is non invasive, multi channel recording of the brain's electrical activity. It is also essential to classify the risk levels of the epilepsy so that the diagnosis can be made easy.

1.1 Methodology
In this paper, the SVD technique is used for dimensionality reduction and Extreme learning Machine as a classifier to classify the risk levels. The EEG data used in the study were acquired from twenty epileptic patients who had been under the evaluation and treatment in the Neurology department of Sri Ramakrishna Hospital, Coimbatore, India. A paper record of 16 channel EEG data is acquired from a clinical EEG monitoring system through 10-20 international electrode placing method. With an EEG signal free of artifacts, a reasonably accurate detection of epilepsy is possible. With the help of neurologist, the artifact free EEG data are selected for this study.

1.2 EEG Signal Acquisitions
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 artefacts and also short enough to avoid any repetition or redundancy in the signal. The EEG signal has a maximum frequency of 50Hz and so, each epoch is sampled at a frequency of 200Hz. Each sample corresponds to the instantaneous amplitude values of the signal, totalling 400 values for an epoch.

2. SVD TECHNIQUE FOR DIMENSIONALITY REDUCTION
The Singular Value Decomposition (SVD) technique is used to extract the features in this study. The SVD method has been a valuable tool in signal processing and statistical data analysis for dimensionality reduction [1]. A SVD of an $M \times N$ matrix X, representing the TFD of the signal x, is given by

$$X = USV^T \quad (1)$$

Where S is an M non square matrix with zero entries anywhere, except on the leading diagonal with elements S_i arranged in descending order of magnitude. Each S_i is equal to $\sqrt{\lambda_i}$ the square root of the eigen value of $C = X^TX$. A stem plot of these values against their index i is known as the singular spectrum. The smaller the eigen values are, the less energy along the corresponding eigenvector there is. Therefore, the smallest eigen values are often considered to be due to noise [2],[3]. The columns of V are an N x N matrix of column vectors which are the eigenvectors of C. The M x M matrix U is the matrix of projections of X onto the eigenvectors of C. If a truncated SVD of X is performed then the truncated SVD is given by $Y = US_PV^T$ and the columns of M x N matrix Y are the noise reduced signal. The columns of the orthonormal matrices U and V are called the left and right SVs, respectively. An important property of U and V is that they are mutually orthogonal. The singular values (σ_{ii}) represent the importance of individual SVs in the composition of the matrix. The vectors in one space are transformed to another space using SVD. SVD is advantageous since it combines two different uncertainty representations into a metric as total uncertainty and it also decomposes uncertainty measures (possibility, belief, probability etc.,) as a collection of vectors of different units, into a principle space. The risk

level code which contains different uncertainty units cannot be added directly this feature can be used in this study.

SVD is also used in various other techniques to reduce coupled non linear behaviour to uncoupled collections of linear behaviour and also enhances the signal to noise ratio[4]. The highest Eigen value obtained is considered as the pattern of the known patient's epilepsy risk level. Figure1 shows the data segmentation from time series into data matrix.

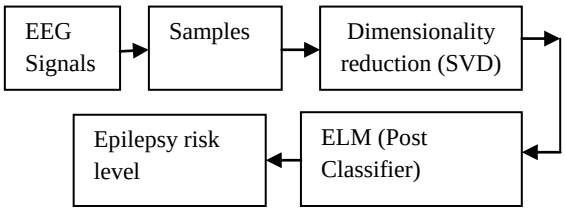

Fig 1: Block diagram of SVD-ELM based epilepsy risk level detector

3. EXTREME LEARNING MACHINE AS A CLASSIFIER

The traditional feed forward neural network parameters need to be tuned and due to this dependency between the different layers of parameters exists. It is known that the gradient descent-based methods have been used in various algorithms of feed forward neural networks. But due to the improper learning steps the learning methods of gradient descent based methods are very slow or may easily converge to local minima. To get better learning performance many iterative learning steps are also required.

In order to obtain high classification accuracy and short training time, Huang *et* [5],[6],[7] al proposed a new learning algorithm called the Extreme Learning Machine for single – hidden layer feed forward neural networks. The extreme learning machine is high popular due to its high generalization ability.ELM is used to classify the protein sequence classification[8] with ten class of super families obtained from a domain database. On comparing the result of the ELM with that of Back –Propagation Neural Networks, the ELM outperforms the BPNN.

In [9], R. Zhang et al developed an ELM for multi category classification in three Cancer Microarray Gene Expression datasets and the results prove that ELM can also avoid problems such as over-fitting, local minima, and improper learning rate. Apart from the field of bioinformatics, ELM has been applied to Biosignal Processing also. N. Y. Liang , et al [10] proposed ELM based classification scheme to classify five mental tasks from different subjects using EEG signals. The performance of the ELM with BPNN and Support Vector Machine(SVM) is compared and the results show that ELM needs an orders of magnitude less training time compared to SVMs and two orders of magnitude less training time compared to BPNN. In [11], a arrthythmia classification scheme proposed using Elm and Principle Component Analysis (PCA) , achieved 97.5% in average accuracy , 97.44% in average sensitivity , 98.46% in average specificity.

According to G.Geetha et al in [13], the ELM learning algorithm is much simple than the other learning techniques for feed forward neural networks. The existing learning techniques can be applied to only differentiable activation function, whereas the ELM learning algorithm can also be used to train SLFNs with many non differentiable activation functions.The ELM algorithm is explained as follows:

Guang-Bin Huang et.a1 in [14],[15]proposed that , suppose learning N arbitrary different instances (x_i,t_i), where $X_i =[x_{i1},x_{i2,.....},x_{in}]^T \in R^n$ and $t_i=[t_{i1},t_{i2,.....},t_{im}]^T \in R^m$ standard Single-layer Feedforward Networks with *N* hidden neurons and activation function g(x), are mathematically modelled as a linear system as

$$\sum_{i=1}^{\tilde{N}} \beta_i g(w_i . x_j + b_i) = T_j \qquad (2)$$

Where $w_i =[w_{i1},w_{i2,.....},w_{in}]^T$ denotes the weight vector connecting the *i*th hidden neuron and the input neuron, $\beta_i =[\beta_{i1},\beta_{i2,.....},\beta_{in}]^T$ denotes the weight vector connecting the *i*-th hidden neuron and output neurons, and b_i represents the threshold of the *i*-th hidden neuron. $w_i * w_j$ represents the inner product of w_i and x_j. If the Single-layer Feedforward Network with N hidden neurons with activation function *g(x)* is able to approximate N distinct instances (x_i,t_i) with zero error means that

$$H\beta=T \qquad (3)$$

Where

$$H(w_1 ,.....,w_{Nh} ,b_1,....,b_{Nh} , x_1 ,...,x_N) = \begin{bmatrix} g(w_1 . x_1 + b_1) & ... & g(w_{Nh} . x_1 + b_{Nh}) \\ . & . & . \\ g(w_1 . x_N + b_1) & ... & g(w_{Nh} . x_N + b_{Nh}) \end{bmatrix} \qquad (4)$$

$$W = \begin{bmatrix} w_1^T \\ . \\ . \\ w_{Nh}^T \end{bmatrix}_{Nh \times m} \quad T=\begin{bmatrix} t_1^T \\ . \\ . \\ t_N^T \end{bmatrix}_{N \times m} \qquad (5)$$

H is the hidden layer output matrix of the SLFN. Hence for fixed arbitrary input weights w_i and the hidden layer bias s, training a Single-layer Feed-forward Network equals to discovering a least-squares solution $\hat{\beta}$ of the linear system $H\beta=T$, $\hat{\beta}=H^{\dagger}T$ is the best weights , where $H^{\dagger}$ is the Moore-Penrose generalized inverse.

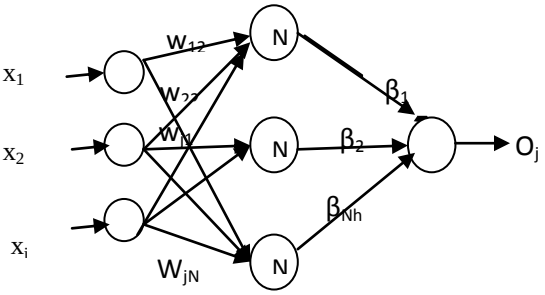

Fig 2: The structure of ELM

The sigmoid function with a gain parameter λ is used in training instead of threshold function directly [15] and it is given by

$$g(x) = 1/(1 + \exp(-\lambda x)) \qquad (6)$$

The sine function is given by

$$g(x)=0.5 \sin(x) \qquad (7)$$

The triangular basis function is given by

$$a = \text{tribas}(n) = 1 - \text{abs}(n), \text{ if } -1 <= n <= 1$$
$$= 0, \text{ otherwise} \qquad (8)$$

The radial basis function is given by

$$\phi(x) = \exp(-\frac{\|x-c\|^2}{\sigma^2}) \quad (9)$$

The procedure of ELM for single-layer feedforward networks is expressed as follows:

1) Choose arbitrary value for input weights w_i and biases of hidden neurons b_i.

2) Calculate hidden layer output matrix *H*.

3) Obtain the optimal $\hat{\beta}$ using, $\hat{\beta}=H^{\dagger}T$.

3.1 Training and Testing: 10 Fold Cross Validation

Among the variety of methods to divide the EEG signals, to reduce the bias of training and testing data, a 10 fold cross validation method is used [12]. This method is implemented during the training period to estimate the classification model that learns from the training data. Usually, the data set is divided into 10 subsets, and holdout approach is reiterated 10 times. Each time, one of the 10 subsets is utilized as the testing dataset and 9 other subsets are combined to form training data set. At last, the average error for 10 trials is calculated.

4. RESULTS AND DISCUSSION

The SVD output for the three epochs is optimized by the Extreme Learning Machine as a single epileptic risk level. The target values of each patient are obtained with different activation functions like sigmoid, sine, triangular basis function and radial basis function. The learning time of ELM is found to be only 0.0156 seconds. The Mean Square Error (MSE) of the actual target and the observed target are calculated as

$$MSE= \quad (T_j - O_j)^2 (i, j = 1, \dots, k) \quad (10)$$

The mean square error for the observed target of twenty patients are given as

Table 1: MSE Values of 10 hidden neurons

patient	sigmoid	sine	tribas	radbas
1	1.2E-07	2.0E-06	1.0E-04	2.0E-06
2	3.0E-06	2.0E-04	7.0E-04	2.0E-04
3	5.1E-06	2.0E-04	1.2E-03	1.0E-04
4	1.1E-06	4.0E-05	4.0E-04	5.0E-04
5	2.9E-06	2.0E-04	5.0E-04	2.0E-04
6	2.6E-08	4.0E-06	6.0E-05	1.0E-05
7	1.2E-07	2.0E-05	1.0E-04	4.0E-06
8	3.1E-08	4.0E-07	9.0E-05	1.0E-06
9	4.7E-08	2.0E-07	5.0E-06	1.0E-06
10	1.9E-08	2.0E-06	2.0E-05	2.0E-06
11	3.1E-06	8.0E-06	5.0E-05	4.0E-06
12	1.3E-07	4.0E-06	2.0E-04	3.0E-06
13	8.6E-08	4.0E-05	2.0E-05	2.0E-06
14	2.3E-06	5.0E-06	8.0E-05	3.0E-06
15	2.0E-08	7.0E-07	2.0E-05	2.0E-07
16	1.9E-08	3.0E-07	2.0E-06	1.0E-07
17	2.8E-08	9.0E-07	2.0E-05	7.0E-07
18	6.0E-07	2.0E-06	3.0E-06	9.0E-06
19	3.5E-08	1.0E-06	3.0E-06	6.0E-07
20	3.0E-06	2.0E-04	7.0E-04	2.0E-04

Table 2: MSE Values of 15 hidden neurons

patient	sigmoid	sine	tribas	radbas
1	1.8E-08	1.0E-05	9.7E-05	7.2E-06
2	6.4E-05	2.0E-04	4.1E-04	9.0E-05
3	1.4E-05	9.0E-05	5.2E-04	4.2E-04
4	7.0E-05	7.0E-05	6.2E-04	1.1E-04
5	3.2E-06	6.0E-05	4.0E-04	2.4E-04
6	1.9E-08	2.0E-06	3.1E-05	3.1E-05
7	4.0E-07	2.0E-06	4.2E-05	2.0E-05
8	4.9E-08	9.0E-07	4.2E-05	8.2E-07
9	1.2E-07	5.0E-08	8.9E-07	1.0E-06
10	4.9E-08	4.0E-07	1.7E-05	4.5E-06
11	2.7E-08	2.0E-06	4.3E-05	7.7E-06
12	2.8E-08	7.0E-07	8.2E-06	4.0E-06
13	4.6E-08	9.0E-06	5.8E-05	1.0E-06
14	8.5E-08	1.0E-06	4.4E-05	2.3E-06
15	1.2E-08	4.0E-07	3.7E-05	1.3E-07
16	1.3E-08	1.0E-07	5.9E-07	6.7E-08
17	1.5E-08	2.0E-07	2.4E-06	6.0E-07
18	4.5E-08	2.0E-06	1.4E-05	2.4E-06
19	2.4E-08	2.0E-07	3.1E-06	8.3E-07
20	6.4E-05	2.0E-04	4.2E-04	9.0E-05

Table 3: MSE Values of 20 hidden neurons

patient	sigmoid	sine	tribas	radbas
1	6.0E-09	2.0E-05	8.0E-05	4.6E-06
2	3.0E-05	2.6E-06	1.0E-04	1.4E-05
3	2.0E-05	1.0E-04	7.0E-04	2.0E-04
4	3.0E-07	2.0E-05	4.0E-04	6.0E-05
5	1.0E-06	2.0E-05	3.0E-04	6.0E-05
6	3.0E-08	4.0E-05	3.0E-05	1.0E-05
7	2.0E-07	4.0E-07	1.0E-04	6.0E-05
8	4.0E-08	9.0E-06	9.0E-06	1.0E-06
9	2.0E-07	3.0E-06	7.0E-06	6.0E-07
10	4.0E-08	1.0E-06	4.0E-06	8.0E-06
11	4.0E-05	5.9E-06	5.0E-05	1.0E-05
12	4.0E-08	4.9E-06	2.0E-05	4.0E-06
13	2E-07	1E-06	2E-05	5E-06
14	6E-08	1E-06	2E-05	3E-06
15	1E-08	3E-07	1E-06	1E-07
16	2E-08	2E-07	1E-06	1E-07
17	2E-08	1.0E-06	2E-06	4.045E-07
18	3E-08	5E-07	1E-05	2E-06
19	2E-08	1.6E-07	1E-06	9E-06
20	3E-05	2.6E-06	0.0001	1E-05

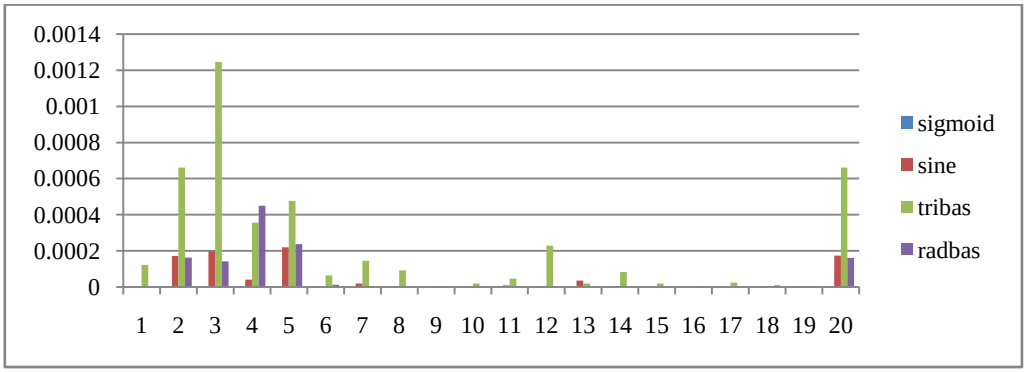

Fig 3: MSE for 10 hidden neurons

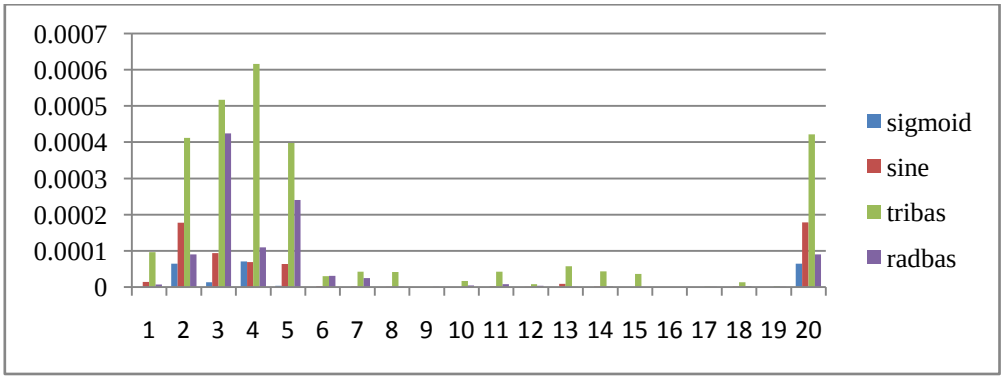

Fig 4: MSE for 15 hidden neurons

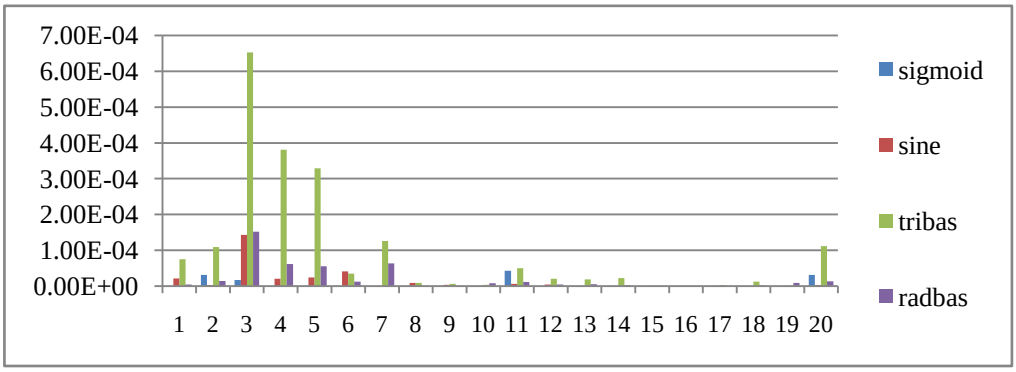

Fig 5: MSE for 20 hidden neurons

5. CONCLUSION

This paper analyses the performance of the Extreme Learning Machine in optimizing the epilepsy risk level of epileptic patients from EEG signals. The classification rate of 100% is achieved and the misclassification rate is nil for an epoch of 2seconds. From this method, the risk levels of the patients are identified and proper medication can be given to them. Also optimizing each region's data separately can solve the focal epilepsy problem.

6. REFERENCES

[1] R.Harikumar, C.Ganesh Babu, P.Sinthiya, M.Balasubramani,"Performance analysis of SVD neural networks for classification of epilepsy risk level from EEG signal", Lecture notes in Electrical Engineering 222,DOI:10.1007/978-81-322-1000-9_3,pp 27-34,Springer 2013.

[2] F. Jiang, R. Kannan, M. Littman, S. Vempala, Efficient Singular Value Decomposition Via Improved Document Sampling. Technical Report. CS-99-5, *Department of Computer Science*, Duke University, 1999.

[3] Abdi H, Singular Value Decomposition (SVD) and Generalized Singular Value Decomposition(GSVD). *Encyclopedia of Measurement and Statistics. Thousand Oaks*: Sage Publications; 2007,907–912.

[4] Stewart GW, "On the early history of the singular value decomposition", *SIAM Rev*1993, 35:551–566.

[5] G.B.Huang ,Q.Y. Zhu and C.K.Siew, "Extreme learning machine : A new learning scheme of feedforward neural networks" in proceedings of international joint conference on neural networks.(IJCNN 20004)(budapest, Hungary),25-29 July, 2004.

[6] G.B. Huang and C.K.Siew, "Extreme learning machine: RBF network case", in proceedings of eighth international conference on control automation, robotics and vision(ICARCV 2004), (Kunming, China) , 6-9 December 2004.

[7] G.B. Huang and C.K.Siew, "Extreme learning machine with randomly assigned RBK kernels", international journal of information technology, Vol.11, no.1,2005.

[8] Wang D. and Huang, G.B. (2005), "Protein sequence classification using extreme learning machine". Proceedings of the 2005 IEEE International Joint Conference on Neural Networks, Montreal, **3**, 31 July-4 August 2005, 53-59.

[9] R. Zhang, R., Huang, G.B., Sundararajan, N. and Saratchandran, P. , " Multicategory classifcation using an extreme learning machine for microarray gene expression cancer diagnosis" IEEE Transactions on Computational Biology and Bioinformatics, 2007 .4(3), 485-495.

[10] Liang, N.Y., Saratchandran, P., Huang, G.B. and Sundararajan, N. "Classification of mental tasks from EEG signals using extreme learning machine". International Journal of Neural Systems 2006, 16(1), 29-38.

[11] Kim, J., Shin, H., Lee, Y. and Lee, M. "Algorithms for classifying arrhythmia using extreme learning machine and principle component analysis", Proceedings of the 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Lyon, 22-26 August 2007, 3257-3260.

[12] Y.Song ,P.Lio, "A new approach for epileptic seizure detection: sample entropy based feature extraction and extreme learning machine", J. Biomedical Science and Engineering , 2010, 3, 556-567.

[13] G. Geetha, S .N. Geethalakshmi, "Detecting epileptic seizure using electroencephalogram: A new and optimized method for seizure classification using hybrid extreme learning machine". International conference on Process automation, control and computing (PACC), July 2011. Page no 1-6.

[14] Guang-Bin Huang and Haroon A. Babri, "Upper Bounds on the Number of Hidden Neurons in Feedforward Networks with Arbitrary Bounded Nonlinear Activation Functions", IEEE Transactions On Neural Networks, Vol. 9, No. 1, January 1998.

[15] Guang-Bin Huang, Qin-Yu Zhu, K. Z. Mao, K. Z. Mao, P. Saratchandran, N. Sundararajan, "Can Threshold Networks be Trained Directly?", IEEE Transactions On Circuits And Systems—II: Express Briefs, Vol. 53, No. 3, March 2006.