Eye state EEG signal classification using Semi-Supervised and Unsupervised Extreme Learning Machine

Sivachitra M*, Keerthika P1, Vadivel P2, Ponnibala M3

*.Department of EEE, Kongu Engineering College, Perundurai, Tamilnadu, India. Email.id: siva.chitra3@gmail.com

1. Department of EEE, Kongu Engineering College, Perundurai, Tamilnadu, India. Email.id: pkeerthika03@gmail.com

2. Department of Mathematics, Kongu Engineering College, Perundurai, Tamilnadu, India. Email.id: niviponvadivel@gmail.com

3. Department of EIE, Kongu Engineering College, Perundurai, Tamilnadu, India. Email.id: ponnibalakumar@gmail.com.

Abstract:

Brain Computer Interface (BCI) technique is the main source for the operation of EEG signals. BCI is a direct pathway to communicate between an enhanced or wired brain and an external device. The proposed work is the classification of EEG signals measured under eye opening and closing state using advanced learning classifiers. Eye Opening and closing dataset is a benchmark data obtained from UCI (University of California Irvine) database. The advance learning classifiers considered are SS-ELM (Semi-Supervised Extreme Learning Machine) and US-ELM (Unsupervised Extreme Learning machine). Semi supervised ELM (SS-ELM) and unsupervised ELM (US-ELM) are recently developed algorithms used for the EEG signal classification task. These algorithms can handle multi-class classification or multi-cluster clustering. SS-ELM algorithm performs better than US-ELM and other real valued algorithms in classifying eye opening and closing states. The improvement is due to the use of spectral techniques in embedding and clustering process.

***Keywords**: Clustering, Extreme learning machine, semi-supervised learning, un-supervised learning*

1 Introduction

In recent times, artificial neural networks are functional in various domains such as classification [1] [2] [3],[7],[13],[14],[15],[16] prediction[8], control systems, workstations, aeronautical engineering and human action recognition [11]. Nowadays BCI has a great progress in developing novel paradigms in the recording of EEG signal. The main goal of BCI system is to help disabled community to connect with the outside world. It is also used to restore the damaged hearing and sight movement. Hence, it requires accurate recognition of the stimuli to avoid false understanding. Therefore it is absolutely more important to determine which specific stimuli can be identified with suitable accuracy.

EEG based recording and monitoring of electric activity of brain is a safe and profitable technique. Feature extraction, pre-processing and classification are some of the efficient signal processing techniques which are essentially used for the practical implementation of the system. Mostly EEG based recording of electrical potentials probably could pose less harm and hence efficient for the observation of brain actions. This work emphasizes on using better classifiers for getting higher classification accuracy with less computation complexity. The basic pruning algorithm for a back propagation network has been studied in [1]. Moreover, the supervised learning algorithms in the complex domain have been represented in [10] and classification application in complex valued domain is dealt in [13].

Extreme Learning Machine (ELM) is used for training single-layer feed-forward networks (SLFNs). The main advantage of using ELM is that the networks input weights are assigned randomly and output weights are found analytically [5],[6] and [9]. ELMs are mainly used for supervised learning tasks such as classification and regression, which greatly limits their applicability. Occasionally, in a few applications attaining labels for a fully supervised learning are time consuming and expensive. Also, the foremost disadvantage of supervised learning is that they cannot use unlabelled data for training the network. It can be ensured that the unsupervised and semi-supervised ELMs can be positioned into a unified framework which possess two major steps i) random input feature mapping and ii) determining output weights. ELM can only be used for supervised learning and hence US-ELM and SS-ELM can be applied for unlabelled data.

In this paper, we have applied semi-supervised ELM (SS-ELM) and unsupervised ELM (US-ELM) [17] to classify eye opening and closing state data. The signal under two conditions namely, eye opening and closing state are recorded by the EEG recorder [12]. These signals are then trained using SS-ELM and US-ELM. The results show that the SS-ELM and US-ELM algorithms are competent with ELM algorithms in terms of efficiency. Also [18] has applied the above said algorithms for classifying motor imagery planning and relaxed states of EEG signals where again SS-ELM performed better than US-ELM in the classification task.

The organization of the paper is as follows, Section II and III describes learning algorithms namely, SS-ELM and US-ELM respectively. Section IV explains EEG eye data and the comparison results of the same. Conclusion and future scope is discussed in the section V.

II. SEMI SUPERVISED EXTREME LEARNING MACHINE

Gao Huang et al., recently developed the SS-ELM algorithm [17] and it has been explained briefly. In semi-supervised learning, a small number of labeled data and many unlabeled data are used. Moreover, when the labeled data are limited, SS-ELM uses the manifold regularization to control unlabeled data such that the classification accuracy can be improved. Skewed data involves poor generalization fit for testing dataset. Skewed data is one in which some classes have more data than the other. Consequently, the penalty factor is introduced to eliminate the effects and hence improves the generalization performance of the testing set. Suppose that x_i belongs to class t_i that has N_{t_i} training patterns, then we can associate e_i with a penalty of

$$a_i = \frac{a_0}{N_{t_i}}$$

Where a_0 could be a user defined parameter as given in earliest ELMs. Subsequently, the samples from the dominant classes will not remain over fitted and samples from the class with less samples will not get ignored.

The design equations of SS-ELM is given by

$$\min \quad \frac{1}{2}\left||\gamma|\right|^2 + \frac{1}{2}||C^{\frac{1}{2}}(\tilde{Y} - H\gamma)||^2$$

$$\gamma \varepsilon R^n \; h^{X n_0} + \frac{\lambda}{2} T_r(\gamma^T H^T L H \gamma) \qquad [1]$$

Where, $Y \in R^{(la+un)Xn_0}$) is that the training target with the initial la rows equal to Y_{la} and the remaining equal to zero. Further, the gradient of the objective function is computed with respect to γ

$$\nabla L_{SS-ELM} = \gamma + H^T C(\tilde{Y} - \mathrm{H}\gamma) + \lambda H^T L H \gamma \qquad [2]$$

SS-ELM Algorithm is given by,

<u>Algorithm 1- SS-ELM algorithm</u>

Inputs:

Labelled patterns, $\{X_{la}, Y_{la}\}=\{x_i, y_i\}_{i=1}^{la}$.

Unlabelled patterns, $X_{un} = \{x_i\}_{i=1}^{un}$.

Output:

The mapping function of SS-ELM: $f: R^{n_i} \rightarrow R^{n_0}$.

Learning Algorithm:

(i) Graph Laplacian L is constructed using each X_{la}, and X_{un}.

(ii) ELM network of n_h hidden neurons, random input weights and biases is constructed and the output matrix of the hidden neurons $H \in R^{(la+un)Xn_h}$ is calculated.

(iii) Trade-off parameter a_0 and λ. *a*re chosen

(iv) If n_h *is less than* N

the output weights are calculated using,

$$\gamma^* = (I_{n_h} + H^T \mathrm{C}\mathrm{H} + \lambda H^T L H)^{-1} H^T \mathrm{C}\tilde{Y}$$

Else the output weights are calculated using

$$\gamma^* = H^T (I_{i+u} + \mathrm{C}\mathrm{H}H^T + \lambda L H H^T)^{-1} \mathrm{C}\tilde{Y}$$

return the mapping function *f*(*x*) = *h*(*x*) γ.

III. UNSUPERVISED EXTREME LEARNING MACHINE

In this section, the US-ELM algorithm recently developed by Gao huang et al [17] has been briefly explained. This algorithm is introduced for learning data without targets. Additionally, it has also been stated that US-ELM gives satisfactory performance compared to recent clustering algorithms.

Here, the total training data $X\,\{x_i\}_{i=1}^{N}$ are unlabelled (N is the training patterns with different classes) and our intention is to search the enhanced structure for the initial data. The beginning of US-ELM follows from the formulation of SS-ELM. For the unlabelled data, Eq (1) is reduced to

$$\min_{\gamma \varepsilon R^{n}\ h^{Xn_0}} \quad \left|\left|\gamma\right|\right|^2 + \lambda T_r(\gamma^T H^T L H\gamma) \qquad (3)$$

This formulation reaches its minimum at $\gamma = 0$. After including constraints, the formulation is given by

$$\min_{\gamma \varepsilon R^{n}\ h^{Xn_0}} \quad \left|\left|\gamma\right|\right|^2 + \lambda T_r(\gamma^T H^T L H\gamma) \qquad (4)$$

$$\text{s.t.} \quad (H\gamma)^T\, H\gamma = I_{n_0}$$

US-ELM Algorithm is given by,

<u>Algorithm 2- US-ELM algorithm</u>

Inputs:

The training data: X ∈ $R^{N\times n_i}$;

Output:

For embedding task:N_+The embedding in a no- dimensional space*:*E ∈ $R^{N\times n_0}$

For clustering task*:*

The label vector of cluster index y ∈ ${N_+}^{N\times 1}$

(i) Graph Laplacian L is built from X**.**

(ii) An ELM structure of $n_h$hidden neurons with random input weights are constructed and the output matrix of the hidden neurons is calculated. H ∈ $R^{N\times n_h}$

(iii) ***If*** $n_h \le N$

the generalized Eigen vectors $v_2, v_3,\ldots, v_{n_0+1}$ of $(I_{n_h} + \lambda H^T LH)v = \Upsilon H^T Hv$ is found corresponding to the second through the n_{0+1} smallest Eigen values.

Let $\gamma = [\widetilde{v_2}, \widetilde{v_3}, \ldots \ldots \tilde{v} n_{0+1}]$, where

$\widetilde{v_i} = v_i / \| Hv_i \|$,i= 2,….., n_{0+1}

Else

The generalized eigenvectors of

$(I_u + \lambda LHH^T)u = \Upsilon HHH^T u$.

$u_2, u_3,\ldots, u_{n_0+1}$ are found corresponding to the second from the $n_0 + 1$smallest eigenvalues.

Let $\gamma = [\widetilde{u_2}, \widetilde{u_3}, \ldots \ldots \tilde{u} n_{0+1}]$,

Where $\widetilde{u_i} = u_i / ||HH^T u_i||$,*i= 2,….., n_{0+1}*

(iv)Embedding matrix: E = $H\gamma$ is computed

(v) (For clustering): Consider each row of *E as a* point, and N points are clustered into K clusters with the help of k-means clustering algorithm. For all the points, let y be the label vector of cluster index return *E* (for the embedding task) or *y* (for the clustering task).

The eigenvectors obtained from US-ELM are not directly used for data representations. However, the same are used as the parameters of the network, namely the output weights. The trained US-ELM model can be used for any presented data in the original input space. Hence, US-ELM provide a useful way for dealing with new patterns without again computing eigenvectors.

IV. DATASET DESCRIPTION AND PERFORMANCE COMPARISON OF US-ELM and SS-ELM FOR EEG DATA

The eye state dataset has been utilized from UCI repository [4]. The EEG signals are measured during two conditions, namely eye opening and closing conditions. Furthermore, a brief description about the experiment carried by [12] is presented. The experiment is performed in a noiseless room. The person is asked to sit in a relaxed condition by watching straight to the camera [12]. The nominee is made to blink without any restrictions and according to his will. The EEG signal is measured for 117 seconds. The frame rate of video camera is estimated to be four times that of the sampling rate of the EEG set. The eye states are manually recorded by video camera along with EEG headset. Here, both open and partly open were denoted as open condition of eye and only completely closed is represented as closed.

The total samples used in [12] paper is 14980. The training and testing data are divided as 75% and 25% individually. The dataset contains 14 input features along with one target vector. The target vector represents eye state opening and closing condition. One can refer [12] to know further particulars about the eye state conditions and its experimental details.

Table I Classification Performance of Eye Dataset

Classifier	Overall Testing Efficiency (%)	NHN/rules
ELM	60	300
USELM	**61.8758**	**800**
K Means	55.1135	300
SS-ELM	**77.0252**	**1000**

The results obtained using semi-supervised extreme learning machine (SS-ELM), unsupervised extreme learning machine (US-ELM), k means clustering and Extreme learning machine(ELM) are presented in table I. From Table I, it can be inferred that SSELM performs better than US-ELM, K means and ELM. Further, US-ELM performs better than K means and ELM network. Though the number of hidden neurons required by the classifiers is considerably larger, the above said networks can able to produce a better performance results. The performance efficiency of both classifiers is good because of the application of spectral techniques for embedding and clustering purpose.

V. CONCLUSION

The drawbacks of supervised learning algorithm is overcome in unsupervised and semi-supervised algorithms. These two algorithms are used for EEG signal classification task. A faster response has been achieved with both of these networks. In a few applications, obtaining labels for a totally supervised learning requires more time, although collecting a large number of unlabeled data is open and cost effective. The two versions of ELMs namely, semi-supervised and unsupervised learning practices spectral techniques for embedding and clustering. The overall testing efficiency of SS-ELM is 77.0252% and USELM is 61.8758% respectively. Furthermore, the results obviously indicate that the performance of SS-ELM and US-ELM classifier are better than other real valued classifier considered for evaluation. Henceforth, it shows that these two classifiers provide better classification efficiency and it can be applied for real world problems. For future work, hybrid procedures and complex valued operations can be incorporated in these extreme learning machines for better classification.

REFERENCES

1. E. D. Karnin, "A simple procedure for pruning back-propagation trained neural networks," *IEEE Transaction on Neural Networks*., vol. 1, no. 2, pp. 239–242,1990.
2. S. Suresh, R. V. Babu, and H. J. Kim, "No reference image quality assessment using modified extreme learning machine classifier," *Applied Soft Computing*, vol. 9, no. 2, pp. 541–552, 2009.

3. Shweta Sahu, and Rajen B. Bhatt, "Automatic classification of Electroencephalography signals Using packet analysis and fuzzy decision trees," *Proc. of 28th National Systems Conference (NSC-2004)*, Dec. 16-18, Vellore, INDIA, 2004.
4. C. Blake and C. Merz, "UCI repository of machine learning databases," Department of Information and Computer Sciences, University of California, Irvine, *[URL: http://archive.ics.uci.edu/ml/]*, 1998.
5. G.-B. Huang, Q.-Y. Zhu, and C.-K. Siew, "Extreme learning machine a new learning scheme of feed forward neural networks," *IEEE Int. Joint Conf. on Neural Networks*, vol. 2. IEEE, 2004, pp. 985–990.
6. G.-B. Huang, Q.-Y.Zhu,C.-K.Siew, Extreme learning machine :theory and applications, *Neurocomputing*,Vol. 70,no.1, pp.489–501,2006
7. R. Savitha, S. Suresh, N. Sundararajan, and H. J. Kim , "A fully complex-valued radial basis function classifier for real-valued classification problems," ,*Neurocomputing*, vol. 78, no.1, pp. 104-110, 2012.
8. S. Gokul, M. Sivachitra, and S. Vijayachitra, "Parkinson's disease prediction using machine learning approaches," *IEEE, Fifth International Conference on Advanced Computing (ICoAC)*, 2013.
9. G.-B. Huang, X.Ding ,H.Zhou , Optimization method based extreme learning machine for classification",*Neurocomputing*,Vol.74,No 1-3, pp. 155–163,2010.
10. N. Sundararajan, S. Suresh, and R. Savitha, Supervised Learning Algorithms in the Complex Domain, Springer-Verlag, Springer Berlin Heidelberg, vol. 421, pp. 1-189, ISBN: 978-3-642-29490-7, 2013.
11. R. Venkatesh Babu, R. Savitha, S. Suresh, and B. Agarwal, "Subject independent human action recognition using spatio-depth information and metacognitive RBF network*," Engineering Applications of Artificial Intelligence* , vol.26, no.9, pp.2010-2021, 2013.
12. Oliver Rosler and David suendermann, "A First Step towards Eye state prediction using EEG," *Hindawi Publishing Corporation*, 2014.
13. M.Sivachitra, and S.Vijayachitra, "Planning and relaxed state EEG signal classification using Complex Valued Neural Classifier for Brain Computer Interface", *IEEE International Conference on CCIP*,pp.1-4 ,2015.
14. R.Savitha , S.Suresh, N.Sundararajan, and H. J.Kim, "Fast learning fully complex-valued ELM classifiers for Real-valued classification problems", *8th International symposium of neural networks* ,pp 602-609,2011.
15. Nan-Ying Liang, Paramasivan Saratchandran, Guang-Bin Huang, and Narasimhan Sundararajan, "Classification of mental tasks from EEG signals Using extreme learning machine,"*International Journal of Neural Systems*, vol.16, no. 1 , pp. 29–38,2006.
16. D. Garrett, D. A. Peterson, C. W. Anderson, and M. H. Thaut, "Comparison of linear, nonlinear, and feature selection methods for EEG signal classification," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol.11 , pp. 141–144,2003
17. Gao Huang, Shiji Song, Jatinder N. D. Gupta, and Cheng Wu "Semi-supervised and unsupervised Extreme Learning Machines",*IEEE Transactions on cybernetics*, Vol. 44,no.12,pp.2405-2417,2014.
18. Dharani.M, M.Sivachitra,, 'Motor Imagery signal classification using semi supervised and unsupervised extreme learning machines', *IEEE Sponsored third international conference on innovations in Information Embedded and Communication Systems*, 17th and 18th March 2017, Coimbatore.