

Study of Variants of Extreme Learning Machine (ELM) Brands and its Performance Measure on Classification Algorithm

Dr. J. Samuel Manoharan

Professor/Dept. of ECE,
Sir Isaac Newton College of Engineering and Technology,
Papakoil, Nagapattinam – 611102,
TamilNadu, India.
drjasm1530@ieee.org

Abstract- Recently, the feed-forward neural network is functioning with slow computation time and increased gain. The weight vector and biases in the neural network can be tuned based on performing intelligent assignment for simple generalized operation. This drawback of FFNN is solved by using various ELM algorithms based on the applications issues. ELM algorithms have redesigned the existing neural networks with network components such as hidden nodes, weights, and biases. The selection of hidden nodes is randomly determined and leverages good accuracy than conservative methods. The main aim of this research article is to explain variants of ELM advances for different applications. This procedure can be improved and optimized by using the neural network with novel feed-forward algorithm. The nodes will mainly perform due to the above factors, which are tuning for inverse operation. The ELM essence should be incorporated to reach a faster learning speed and less computation time with minimum human intervention. This research article consists of the real essence of ELM and a briefly explained algorithm for classification purpose. This research article provides clear information on the variants of ELM for different classification tasks. Finally, this research article has discussed the future extension of ELM for several applications based on the function approximation.

Keywords: *Extreme Learning Machine*

1. INTRODUCTION

Nowadays, deep learning methods are widely used in many fields due to their significant advantages. Deep learning integrates feed-forward neural networks to train and test the model with higher accuracy [1]. Many complex non-linear mapping problems are handled by deep learning methods. The model is developed based on the natural artificial phenomena, which can be easily predicted by the parameter. The tuning of feed-forward network in the hidden layers can consume less computation time [2]. Therefore, many researchers are optimizing their neural network scheme with many validation tuning techniques. Figure 1 shows the simple construction of ELM architecture.

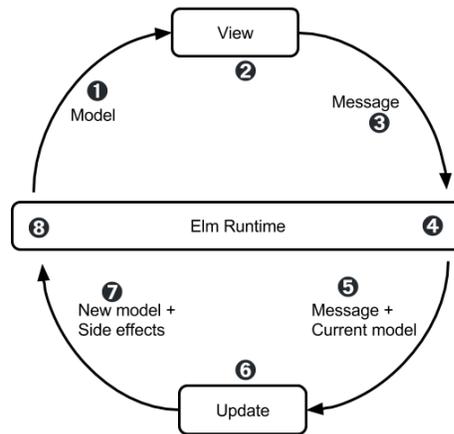


Figure 1 Model Construction for ELM Architecture

Single hidden layer feed-forward network (SLFNs) is used as a network structure for any performing classification and prediction with learning capabilities and fault-tolerant skills. The parameters are tuned by many iterative procedures to obtain a local minimum, which will be relatively slow [3]. There are many new fast learning algorithms to improve the accuracy and minimize the computation time of the overall framework. The learning rate's loss and epochs are used to tune the control parameters. The ELM can be implemented for the prediction process without performing an iterative tuning procedure [4] [5]. The learning efficiency is also faster than other traditional methods [6]. The learning parameters are tuned by concerning the input

weight vector and biases of the hidden nodes. Figure 2 shows the basic internal network structure of ELM.

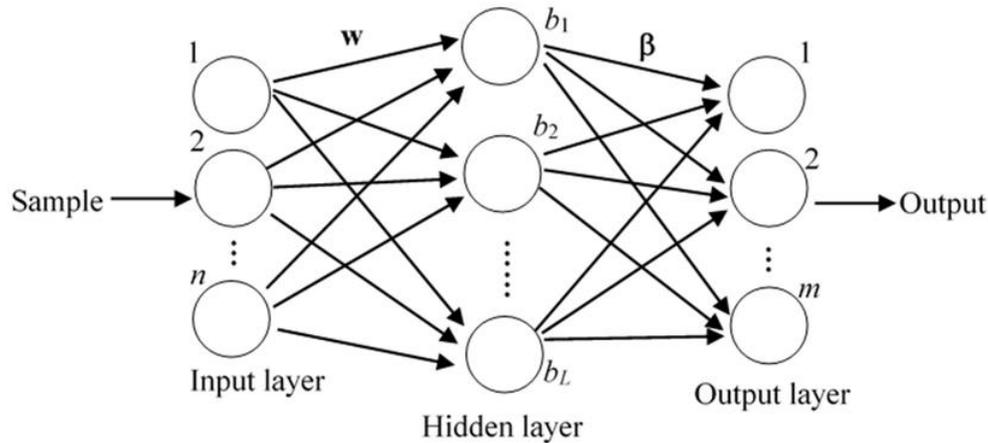


Figure 2 Basic Structure of ELM

The parameters are obtained by a simple computation process of inverse operation. The training procedure is based on the nonlinear transformation activities for generalizing the performance of minimizing the computation process [7][8][9]. The improvements of activation function in every hidden node are a universal approximation to the standard ELM, which is applied to many real-time predictions or classification applications. The parameters of hidden nodes are selected randomly and it becomes a linear system with the inverse operation [10][11][12].

2. ORGANIZATION OF RESEARCH ARTICLE

The proposed research article is arranged as follows: section 3 includes the related research works, section 4 describes the types of ELM, and section 5 provides a discussion about the results and performance measures between them. The conclusion and future research scope is described in section 6.

3. PRELIMINARIES

The new developing algorithm concentrates more on computational time and training accuracy with classifiers. This can be achieved by ELM based on the select factors for training and biases of the layers. In ELM, the factors are determined by mathematical analysis manipulation [13][14]. To improve the performance, the algorithm should be optimized with computational intelligence by using less complex networks [15]. This section includes the discussion on many related existing research works.

Huang et al introduces original ELM, which is used by assigning the weights and biases for the layer. They used mathematical transformation formula to obtain the parameter for every layer. They solved both classification and regression problems and compared many deep learning methods [16].

Rong et al proposed the method of ELM with pruning function. They concentrated on work by classifying the relevant and irrelevant hidden layers and it is removed after the class label [17]. Feng et al concentrates on error minimized ELM to solve the classification and regression problems. They developed an algorithm to automatically detect the hidden nodes [18]. Lan et al proposed two-stage algorithms for ELM. Their network structure includes two different structures to train the model. They have solved regression problems of the repository [19]. Liang et al developed an online sequence method of ELM for convolutional operations. Besides, they solved the time series prediction problems by developing an algorithm [20]. Evolutionary ELM was developed by Zhu et al, which is used to train a model to classify the different items. They introduced the optimization method in the ELM algorithm, which is called as effective ELM. They have solved classification problems with linearity conditions [21]. Cao et al developed the ELM method based on voting. They include real-time datasets from the UCI database. They determined protein information from the resource spot based on the conservative voting method [22]. Deng et al developed a model for regression problem-solving. They also obtained many large-scale regression datasets with artificial content, which is called as ordinal ELM [23]. Li et al developed a complex ELM algorithm with a dual-domain channel. They are extended with real and complex domain for phase channel model [24]. Liu et al proposed an even type ELM to

solve prediction problems. They have developed a synthetic activation function with hidden layer input variables of the sample [25].

4. METHODOLOGIES

In this section, various ELM methodologies for managing hidden nodes are presented.

4.1 Evolutionary ELM

Generally, the ELM process is randomly determined by the hidden neurons. But these hidden neurons are randomly determined due to the computation of input weight vectors and biases. This evolutionary procedure is initiated in the learning process to optimize the existing deep learning process. This will be varied due to the selection of input to output weight factors with the conditions for biases in hidden nodes. The model optimization can be adopted with input weights and biases in hidden layers [19]. The inverse process is implemented here to determine the output weights. This procedure can be called as Evolutionary ELM (E-ELM). This procedure can be implemented to get good broad view performance of complex networks.

4.2 Voting based ELM

The ELM includes the learning parameter of hidden nodes that contained unchanged biases. There are possibilities for misclassification due to boundary classification challenges. The improved version is based on polling samples at the classification section with an extreme learning machine called Voting based ELM (V-ELM). Instead of a single training method, numerous ELM training nodes are created to conduct a voting system for obtaining a final decision based on the intelligence degree of neural networks. [26]. This will provide good accuracy and enhanced classification efficiency performance, which is used to reduce the misclassification problem. This V-ELM is performing well to overcome many classification problems with real-time datasets.

4.3 Ordinal ELM

The ordinal regression is incorporated with the ELM algorithm that provides rapid training speed improvement and good classification ability. The ordinal regression architecture is constructed with three types of encoding with various classifier units named multi-output, multiple, and decomposition-based classifiers. The single multi-output classifier, multiple binary

classifications, and decomposition method are incorporated with ELM and redesigned with three types of datasets [27].

4.4 Fully complex ELM

The construction of a learning algorithm is done in the complex domain, which addresses the complex nonlinear channel problems. For the continuous distribution probability, the ELM is designed and extended for real input weight vector and layer biases. The output weights were computed with iterative tuning method [28]. The Quadrature Amplitude Modulation (QAM) signals are used for equalizing the complex domain from the real domain.

4.5 Symmetric ELM

Li et al conducted the test to modify ELM algorithm with the activation function of input variables. The transformation of original activation is done with symmetric properties. It provides the result with a comparatively faster learning process and higher accuracy in the classification domain. The arbitrary distinct samples are designed by symmetric property with zero error [24]. This model can be extended for improving the performance of ELM as fuzzy ELM, Parallel ELM, and weighted ELM.

4.6 Incremental ELM

The feed forward network is constructed by increased number of randomly added nodes in the network. The hidden layers are frozen based on the output weights of the previously hidden nodes. Besides, the activation functions can be extended and it will design more compact network architectures. It is derived from SLFN with a differentiable function. Based on this phenomenon, it will achieve faster convergence rates in the network architecture [29]. Several complex nodes are extended through numbers, which are called as incremental ELM (I-ELM).

4.7 Pruning ELM

When hidden nodes are increased, the over-fitting problems occur in the complex network domain of classification. This systematic and automatic approach is used to design the ELM network. The removal of some irrelevant hidden nodes will help to obtain the relevance of complex nodes. This change in the architecture of ELM is based on leaving the deactivated hidden nodes, which improves the speed of the learning and execution time.

4.8 Error minimized ELM

The overall error can be minimized in the complex network by deactivating the hidden nodes that are not relevant to the classification process. This can be determined by generalized SLFNs by measuring output weight vector. The computational complexity is reduced during this generalization for obtaining the efficient operation of ELM.

4.9 Dual Stage ELM

Two different algorithms are used at each network stage to provide better performance in a dynamic process. The forward recursive algorithm is executed at the first stage network to select the hidden nodes. The relevant nodes are selected to train the model for reducing network complexity.

4.10 Online sequential ELM

The training data is obtained sequentially for the classification process. The ELM is extended by this phenomenon called OS-ELM with the sequential property. This unified framework linking the input to hidden nodes with randomly bias generation. These output weights are determined by using many hidden nodes. The learning algorithm requires some sort of hidden nodes for improving the performance of ELM. The ensemble networks are also integrated with OS-ELM. This result provides better accuracy and efficiency when compared to the existing learning algorithm [30].

4.11 Classification Process for Variants of ELM

The classification is done through the prediction of many classifiers. The sample image set can be pre-processed to extract the features. The spectral characteristics are arranging the sample images by using several training sets, which are provided to the classifier section after the completion of training process. Figure 3 shows the classification process in ELM.

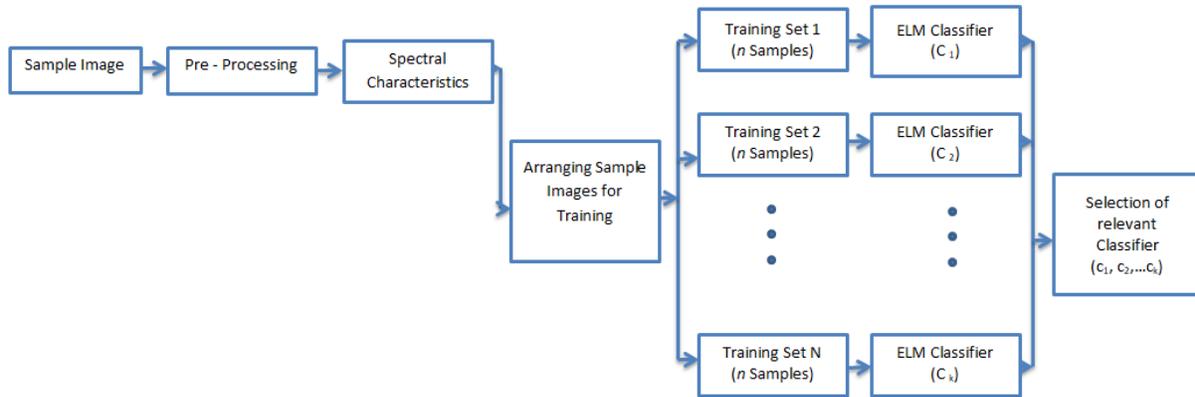


Figure 3 Classification Process in ELM

Initially, the occupied region can index the object dynamically through the changes in the regional environment. This extreme learning machine algorithm is developed for both multi-label and “Uni” label situations based on classification procedure. There are many ECG signal classification in medical images by ELM for obtaining high effective accuracy with the short time learning process [31]. This can be diagnosed through the signal with less computation to investigate any type of heart disease. The appropriate classifier is selected for obtaining the relevant sample images with the reference of unknown sample images shown in figure 4.

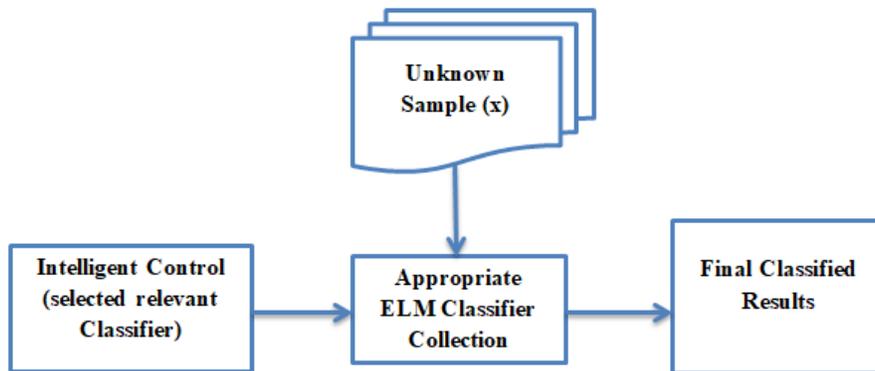


Figure 4 Final Stage of Classification Section

5. RESULTS DISCUSSION

The proposed research work consists of a whole empirical study about the variants of ELM to achieve a generalization performance. These performance measures are investigated

with testing accuracy, computation time, and classification success. The comparison has been performed with statistical features and intelligent-based learning methods to show better generalization performance. The variants of ELM include methods that are based on adjusting the hidden nodes in the neural structure layers. This type of incremental and pruning ELM is generating a faster response in the learning time and it is used to determine the number of active nodes in the layers. The calculation of execution time of a single dataset handling structure is considered in this research work. Table 1 shows the overall comparison results of ELM algorithms in the learning methods.

Table 1 Overall Performance Measures of ELM Classification

S.NO	ALGORITHM	ACCURACY	EXECUTION TIME	CLASSIFICATION SUCCESS	
				Tr	Ts
1	Statistical Feature	85%	> 1 h	75%	69%
2	Intelligent Learning based method	89.2%	37 min	78%	70%
3	Deep Belief Networks	93%	4.3 min	85%	80%
4	ELM based Classification	94.1%	58 s	87%	85%

The adjustment of hidden layer's nodes are performed with error minimized ELM and dual-stage ELM method in the domain. OS-ELM is very faster than the learning algorithm with better accuracy for dynamic or static datasets. Figure 5 shows the overall performance graph of ELM based classification algorithm.

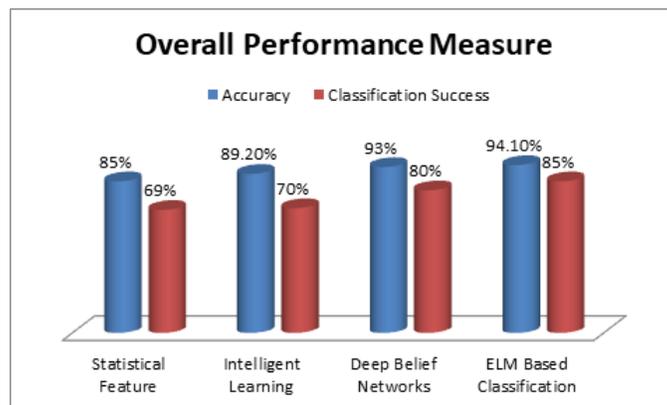


Figure 5 Overall Performance Measures for Classification Agenda

E-ELM algorithm is used to optimize the model structure by adjusting the weights and biases values which are reducing computation time. The fully complex ELM algorithm is used to convert the real domain to complex domain. When compared too other methods, ELM-based methods provide the faster learning speed with good scalability function.

6. CONCLUSION

Thus, the research work has described the variants of ELM and compared its accuracy and executing time. The comparison has been done with various machine learning algorithms successfully. The research problem is discussed in the results and discussion section. During our investigation, we discovered some useful research points that will help us to continue our research in the future, they are.;

1. The determination of an appropriate number of neurons in the hidden layer is a very challenging task with various datasets in the complex domain. This research problem is still open for researchers.
2. The oscillation boundary of the generalization performance measure estimation is still a challenging task in the domain.
3. With the huge amount of dataset, the classification success rate is reduced due to computation and various classifiers that include increased computation time for the procedure.
4. The various applications can be included in the further process of ELM. The generalization ability can be improved with the selection of an appropriate dataset.
5. ELM will be applied in many applications such as pattern recognition, weather forecasting, medical image diagnosis, and any classification structure problems in single framework architecture with the distributed computation procedure.

REFERENCES

- [1] G.-B. Huang, Q.-Y. Zhu, and C.-K. Siew, "Extreme learning machine: theory and applications," *Neurocomputing*, vol. 70, no. 1–3, pp. 489–501, 2006.

- [2] N.-Y. Liang, G.-B. Huang, P. Saratchandran, and N. Sundararajan, "A fast and accurate online sequential learning algorithm for feedforward networks," *IEEE Transactions on Neural Networks*, vol. 17, no. 6, pp. 1411–1423, 2006.
- [3] Y. Lan, Y. C. Soh, and G.-B. Huang, "Ensemble of online sequential extreme learning machine," *Neurocomputing*, vol. 72, no. 13–15, pp. 3391–3395, 2009.
- [4] H.-J. Rong, G.-B. Huang, N. Sundararajan, and P. Saratchandran, "Online sequential fuzzy extreme learning machine for function approximation classification problems," *IEEE Transactions on Systems, Man, and Cybernetics B: Cybernetics*, vol. 39, no. 4, pp. 1067–1072, 2009.
- [5] G. Feng, G.-B. Huang, Q. Lin, and R. Gay, "Error minimized extreme learning machine with growth of hidden nodes and incremental learning," *IEEE Transactions on Neural Networks*, vol. 20, no. 8, pp. 1352–1357, 2009.
- [6] W. Zong, G.-B. Huang, and Y. Chen, "Weighted extreme learning machine for imbalance learning," *Neurocomputing*, vol. 101, pp. 229–242, 2013.
- [7] B. Mirza, Z. Lin, and K.-A. Toh, "Weighted online sequential extreme learning machine for class imbalance learning," *Neural Processing Letters*, vol. 38, no. 3, pp. 465–486, 2013.
- [8] X. Liu, L. Wang, G. B. Huang, J. Zhang, and J. Yin, "Multiple kernel extreme learning machine," *Neurocomputing*, vol. 149, part A, pp. 253–264, 2015.
- [9] Y. Lan, Y. C. Soh, and G. B. Huang, "Two-stage extreme learning machine for regression," *Neurocomputing*, vol. 73, no. 16-18, pp. 3028–3038, 2010.
- [10] C.-T. Lin, M. Prasad, and A. Saxena, "An Improved Polynomial Neural Network Classifier Using Real-Coded Genetic Algorithm," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 45, no. 11, pp. 1389–1401, 2015.
- [11] C. F. Stallmann and A. P. Engelbrecht, "Gramophone Noise Detection and Reconstruction Using Time Delay Artificial Neural Networks," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 47, no. 6, pp. 893–905, 2017.

- [12] E. Cambria, G.-B. Huang, and L. L. C. Kasun, "Extreme learning machines," IEEE Intelligent Systems, vol. 28, no. 6, pp. 30–59, 2013.
- [13] W. Zhu, J. Miao, L. Qing, and G.-B. Huang, "Hierarchical Extreme Learning Machine for unsupervised representation learning," in Proceedings of the International Joint Conference on Neural Networks, IJCNN 2015, Ireland, July 2015.
- [14] W. Yu, F. Zhuang, Q. He, and Z. Shi, "Learning deep representations via extreme learning machines," Neurocomputing, pp. 308– 315, 2015
- [15] J. Wei, H. Liu, G. Yan, and F. Sun, "Robotic grasping recognition using multi-modal deep extreme learning machine," Multidimensional Systems and Signal Processing, vol. 28, no. 3, pp. 817– 833, 2016.
- [16] Huang GB, Zhu QY, Siew CK (2004) Extreme learning machine: a new learning scheme of feedforward neural networks. In: Proceedings of international joint conference on neural networks (IJCNN2004), vol 2, no 25–29, pp 985–990.
- [17] Rong HJ, Ong YS, Tan AH, Zhu Z (2008) A fast pruned-extreme learning machine for classification problem. Neuro computing 72:359–366.
- [18] Feng GR, Huang GB, Lin QP, Gay R (2009) Error minimized extreme learning machine with growth of hidden nodes and incremental learning. IEEE Trans Neural Netw 20(8):1352–357.
- [19] Lan Y, Soh YC, Huang GB (2010) Two-stage extreme learning machine for regression. Neuro computing 73(16–18):3028–3038.
- [20] Liang NY, Huang GB, Saratchandran P, Sundararajan N (2006) A fast and accurate on-line sequential learning algorithm for feedforward networks. IEEE Trans Neural Netw 17(6):1411–1423.
- [21] Zhu QY, Qin AK, Suganthan PN, Huang GB (2005) Evolutionary extreme learning machine. Pattern Recognit 38:1759–1763.
- [22] Cao JW, Lin ZP, Huang GB, Liu N (2012) Voting based extreme learning machine. Inf Sci 185(1, 15):66–77.

- [23] Deng WY, Zheng QH, Lian SG, Chen L, Wang X (2010) Ordinal extreme learning machine. *Neurocomputing* 74(1–3):447–456.
- [24] Li MB, Huang GB, Saratchandran P, Sundararajan N (2005) Fully complex extreme learning machine. *Neuro computing* 68:306–314.
- [25] Liu XY, Li P, Gao CH (2013) Symmetric extreme learning machine. *Neural Comput Appl* 22(3–4):551–558.
- [26] C. Sun, W. He, W. Ge, and C. Chang, “Adaptive Neural Network Control of Biped Robots,” *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 47, no. 2, pp. 315–326, 2017.
- [27] S.-J. Wang, H.-L. Chen, W.-J. Yan, Y.-H. Chen, and X. Fu, “Face recognition and micro-expression recognition based on discriminant tensor subspace analysis plus extreme learning machine,” *Neural Processing Letters*, vol. 39, no. 1, pp. 25–43, 2014.
- [28] G. Huang, “An insight into extreme learning machines: random neurons, random features and kernels,” *Cognitive Computation*, 2014.
- [29] G. B. Huang, Q. Y. Zhu, and C. K. Siew, “Extreme learning machine: theory and applications,” *Neurocomputing*, vol. 70, no. 1–3, pp. 489–501, 2006.
- [30] S. Ding, N. Zhang, X. Xu, L. Guo, and J. Zhang, “Deep Extreme Learning Machine and Its Application in EEG Classification,” *Mathematical Problems in Engineering*, vol. 2015, Article ID 129021, 2015.
- [31] B. Qu, B. Lang, J. Liang, A. Qin, and O. Crisalle, “Two hidden-layer extreme learning machine for regression and classification,” *Neurocomputing*, vol. 175, pp. 826–834, 2016.