Parkinson's Disease Prediction Using Machine Learning Approaches

Gokul.S
PG Student M.E (Applied Electronics)
Department of EEE
Kongu Engineering College
Perundurai, India
gokuls.12ae@kongu.edu

Sivachitra.M
AssistantProfessor(SLG)
Department of EEE
Kongu Engineering College
Perundurai, India
sivachitra@kongu.ac.in

Vijayachitra.S
Professor
Department of EIE
Kongu Engineering College
Perundurai, India
svijayachitra@kongu.ac.in

Abstract—**This paper proposes the application of a Fully Complex-Valued Radial Basis Function network (FC-RBF), Meta-Cognitive Fully Complex-Valued Radial Basis Function network (Mc-FCRBF) and Extreme Learning Machine (ELM) for the prediction of Parkinson's disease. With the help of Unified Parkinson's Disease Rating Scale (UPDRS), the severity of the Parkinson's disease is predicted and for untreated patients, the UPDRS scale spans the range (0–176). The FC-RBF network uses a fully complex valued activation function *sech*, which maps $c^n \rightarrow c$. The performance of the complex RBF network depends on the number of neurons and initialization of network parameters. The implementation of the self-regulatory learning mechanism in the FC-RBF network results in Mc-FCRBF network. It has two components: a cognitive component and a meta-cognitive component. The meta-cognitive component decides how to learn, what to learn and when to learn based on the knowledge acquired by the FC-RBF network. Extreme learning mechanism uses sigmoid activation function and it works with fast speed. In ELM network, the real valued inputs and targets are applied to the network. The result indicates that the Mc-FCRBF network has good prediction accuracy than ELM and FC-RBF network.**

***Keywords*-Neural Networks, Parkinson's Disease (PD), Fully Complex-Valued Radial Basis Function Networks (FC-RBF), Meta-Cognitive Radial Basis Function Networks (Mc-FCRBF), Extreme Learning Machine (ELM).**

I. INTRODUCTION

Parkinson's disease is a neurological disorder effect and is a second most common neurodegenerative disorder after Alzheimer's [1]. In North America alone more than one million people are affected by PD [2]. Parkinson's disease affects the movements, including speaking and writing. It involves the malfunction and death of vital nerve cells in the brain called Neurons. The dying neurons develop a chemical substance called dopamine, which sends messages to the part of the brain that controls movements and coordination and also act as a messenger between two brain areas. Substantia nigra and corpus striatum are the two brain areas used to produce the smooth controlled movements. In the brain the amount of dopamine produced decreases and it makes the person unable to control the movements normally [3]. Due to increase in population, number of PD patients is expected to raise. Although medication is available, there is no complete treatment for PD. So the early diagnosis is important to help patients and to improve the quality of their life. PD is characterized by tremor of the limbs, muscle rigidity, slowness of the movement, difficult with walking, balance and coordination, vocal impairment and mood disturbances [4].

Currently, the Parkinson's disease severity can be measured using a widely accepted metric, the unified Parkinson's disease rating scale (UPDRS). The UPDRS value ranges from 0-176, with 0 represents healthy condition and 176 represents total disability condition. The UPDRS scale is based upon the following three factors 1. Mood, Mentation and behavior 2. Daily living activities 3. Motor. Out of these three factors the third factor ranges from 0-108, with 0 denotes symptom free condition and 108 denotes severe motor condition [1]. With the help of the UPDRS scale the severity of the Parkinson's disease is predicted.

Recently the research in the field of neural networks has taken diversion towards processing data in the complex domain and it is applied in the areas like communication engineering, adaptive array signal processing and medical imaging which uses complex-valued signals [7]. The major drawback in the real valued neural network is that it could not process maximum information. In an ordinary neural network there will be information loss, if the real valued data is processed as a complex-valued data the information loss will be overcome. Several complex-valued neural network information are given in [12], [13], [14], [15], [16], [17], [18], [19], [20]. In this paper a FC-RBF network, Mc-FCRBF network and Extreme learning machine are used for the prediction of Parkinson's disease.

978-1-4799-3448-5/13/$31.00 ©2013 IEEE

II. FULLY COMPLEX-VALUED RADIAL BASIS FUNCTION NETWORK (FC-RBF)

A FC-RBF Network uses complex-valued weight, complex-valued activation function and uses gradient descent learning algorithm [5]. The complex-valued activation function $\sec h(.)$ has Gaussian like characteristics which maps $c^n \rightarrow c$. The FC-RBF network has input layer, hidden layer and output layer. The input layer has *m* neurons, a hidden layer has *K* neurons and a output layer has *l* neurons.

The neurons in the hidden layer use $\sec h$ activation function, which is the basic functional unit of the hidden neuron for complex domain. For the given training data set $\{(x_1^t,y_1^t), (x_2^t,y_2^t),\ldots. (x_m^t,y_n^t)\}$, where $x^t \in c^m$ are the inputs and $y^t \in c^n$ are the targets. The training data set uses N number of samples and the output of *l*-th neuron of the network is given as:

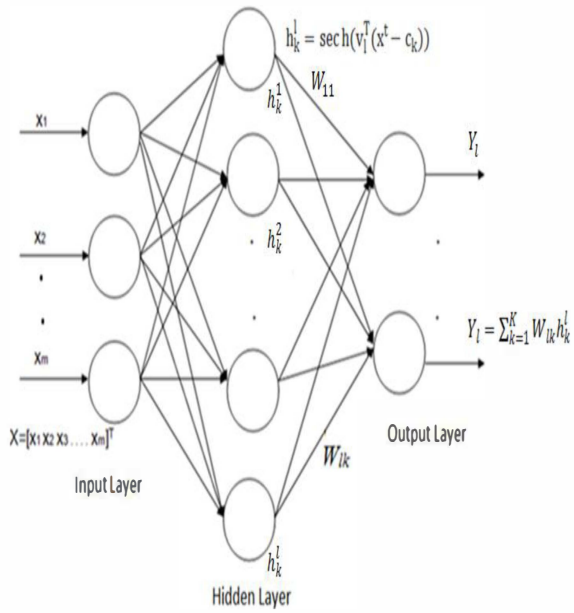

Figure 2.1 Architecture of FC-RBF Network

$$Y_l = \sum_{k=1}^{K} W_{lk}\, h_k^l \qquad (1)$$

Where the W_{lk} are the complex-valued weights and the h_k^l are the response of the network in k-th hidden neuron. The activation function used in hidden neurons is:

$$h_k^l = \sec h(v_l^T(x^t - c_k)) \quad l = 1,2\ldots k \qquad (2)$$

Where v_k is the scaling factor and c_k is the center of the k-th neuron. The network parameters (learning rate, weight) are updated based on gradient descent learning algorithm[5].

III. META-COGNITIVE FULLY COMPLEX-VALUED RADIAL BASIS FUNCTION NETWORK (MC-FCRBF)

Meta-cognitive Fully complex-valued radial basis function network (Mc-FCRBF) has been developed based on the Nelson and Narens model [6]. It has two components: a cognitive component and a meta-cognitive component [8], [21], [22]. A cognitive component contains the FC-RBF network and a meta-cognitive component has a dynamic model of the cognitive component.

A FC-RBF network with a fully complex-valued activation function is explained in the section II. A meta-cognitive component decides how to learn, what to learn, and when to learn when a sample is presented [27], [28], [39]. The learning process is based on the knowledge acquired by the FC-RBF network and also the information contained in the sample. The learning mechanism enable the samples with higher information content to be learned first and then the samples with lower information content to be learned in the later stages of the training process. The sample which contains similar information is deleted during the training process [23], [24], [25], [26]. So, the meta-cognitive component avoids learning the similar type of samples in every epoch of the batch learning process, which improves the performance of FC-RBF network.

The strategies used in the Mc-FCRBF network are sample deletion, sample learning and sample skipping.

S***ample deletion*****:** It is used to delete the samples which contain similar information in the training data set. If $M_t^e < E_d^M$ & $\phi_t^e < E_D^\phi$, where E_d^M is the delete magnitude threshold and E_D^ϕ is the delete phase threshold, the sample *t* is deleted from the training data set.

Sample learning**:** It is used to update the network parameters in the current epoch. The sample learning condition is as follows: If $M_t^e \geq E_l^M$ or $\phi_t^e \geq E_l^\phi$, where E_l^M and E_l^ϕ are the self-regulated learning parameters. The initial values of the learning parameters are set to be high which facilitates the participation of more samples in the initial epochs. The self-regulation of the magnitude and phase learning parameters occur based on the following conditions: If $M_t^e \geq E_l^M$, $E_l^M = \delta E_l^M - (1-\delta)M_t^e$

If $\phi_t^e \geq E_l^\phi$, $E_l^\phi = \delta E_l^\phi - (1-\delta)\phi_t^e$

Where δ is the slope at which the thresholds are self regulated. The parameter E_l^M and E_l^ϕ can be choosen between the range (0.1 to 0.4) and (0.00005 to 0.005) respectively. During learning, the network parameters

are updated based on the gradient descent learning algorithm for the FC-RBF network [5].

***Sample skipping*:** It is used to skip the sample from learning in the current epoch and to retain the sample, this happens when a sample does not satisfy the sample deletion or sample learning condition in the current epoch. So the same sample might be used for learning in subsequent epochs.

IV. EXTREME LEARNING MACHINE (ELM)

In general, feed forward neural networks are slower and it may take several time to train neural networks. In feed forward networks gradient descent-based learning algorithms are most widely used. This learning method is very slow due to the improper learning steps and also it may easily converge to local minima. To obtain a better learning performance these algorithms need several learning steps iteratively. Inorder to overcome the above mentioned disadvantage Extreme Learning Machine (ELM) network is used [30], [31].

ELM is a simple tuning-free three step algorithm and it is a single-hidden-layer feed forward neural network (SLFNs) [9]. The learning speed of ELM is extremely fast compared to other traditional methods [32], [33]. SLFNs randomly choose the input weights and it determines the output weight analytically. The hidden node parameters are not only independent of the training data but also of each other. It could generate the hidden node without considering the training data. ELM works for all nonlinear piecewise continuous activation functions [34], [35].

The three-step learning model for ELM is summarized as follows:

Given a training data set $\aleph = \{(x_i, t_i) | x_i \in R^n, t_i \in R^m, i = 1, \ldots, N\}$, the output function of a hidden node is *G(a, b, x)* and L denotes the number of hidden nodes ,

step1: Generate random hidden nodes (a_i, b_i), i=1,…,L.

step2: Determine the hidden layer output matrix H.

step3: Find the output weight vector β: $\beta = H^+T$.

Where H^+ denotes the Moore-Penrose generalized inverse of matrix H.ELM is efficient in batch learning, sequential learning and incremental learning. It has been recently extended to kernel learning. The prediction accuracy of ELM is better than Back Propagation (BP) and similar to support vector machines in many regression applications.

V. PREDICTION OF UPDRS SCALE USING FC-RBF, Mc-FCRBF, ELM

The Parkinson's disease data set is taken from UCI machine learning repository [11] and it was created by Athanasios Tsanas and Max Little of the University of Oxford, which is collaborated with ten medical centers in the US. The data set is composed of a range of biomedical voice measurements from 42 people with early stage Parkinson's disease recruited to a six month trial, and it consists several attributes such as subject, age, test time, motor UPDRS, total UPDRS, jitter, shimmer, noise-to-harmonics ratio , harmonics-to-noise ratio, recurrence period density entropy, detrended fluctuation analysis, pitch period entropy [1]. The prime important attribute considered here is total UPDRS. The total number of samples present in the dataset is 804. Out of 804 samples, 575 samples are used for training and 229 samples are used for testing. The dataset (training and testing) contains four input features and one target. The four input features of a sample represent the total UPDRS value measured for consecutive four weeks and the target represents the UPDRS value measured on the fifth week.

All the samples are real valued data and it is converted into a complex-valued data. The number of input neurons used is four. It has fifteen hidden neurons and one output neurons. The hidden neuron selection is based on the incremental-decremental procedure presented in [10]. These networks are trained for 1000 epochs and the results obtained are compared. Testing and validation RMSE of Fc-RBF, ELM and Mc-FCRBF are given in the table I and table II.

A plot of time Vs UPDRS value as it is predicted by FC-RBF for the training data and testing data are presented in the figure 5.1 ,5.2, 5.3 and 5.4. UPDRS (Magnitude and Phase) Vs time prediction graph for Mc-FCRBF has been represented in the figure 5.5, 5.6, 5.7 and 5.8.The validation and testing plots for ELM network has been presented in the figures 5.9 and 5.10. In these curves the solid line represents the actual output value and the dotted line represents the predicted value. The error between these two curves is given by their difference. From these figures it is observed that Mc-FCRBF network predicts the Parkinson's disease accurately as compared with ELM and FC-RBF network because the error obtained is less. Mc-FCRBF network performance is better than FC-RBF due to the self regulatory mechanism in the meta-cognitive component.

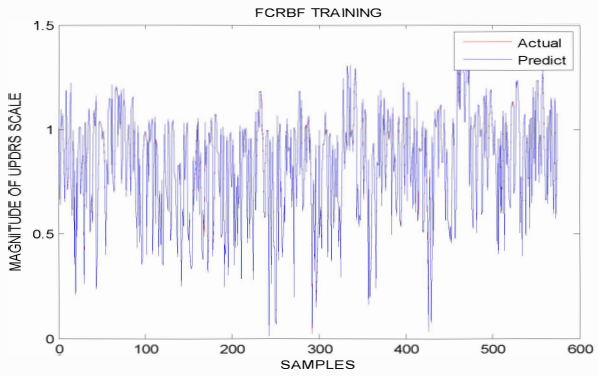

Figure 5.1 FC-RBF: Prediction of magnitude of UPDRS scale for training data

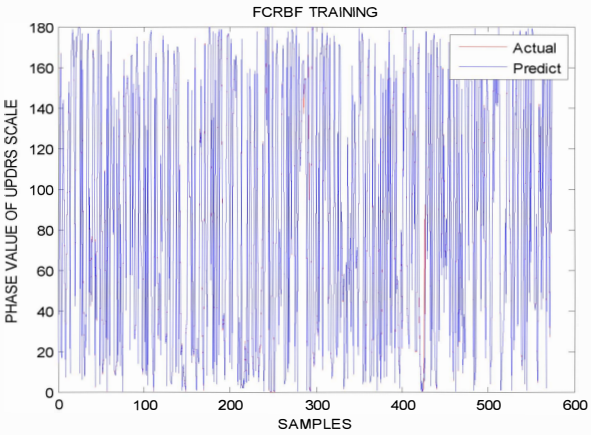

Figure 5.2 FC-RBF: Prediction of phase value of UPDRS scale for training data

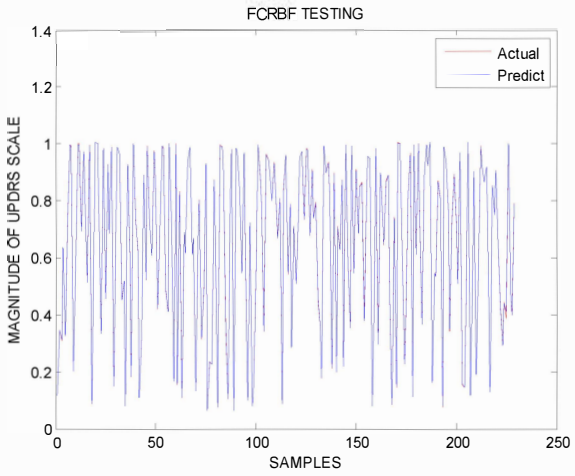

Figure 5.3 FC-RBF: Prediction of magnitude of UPDRS scale for testing data

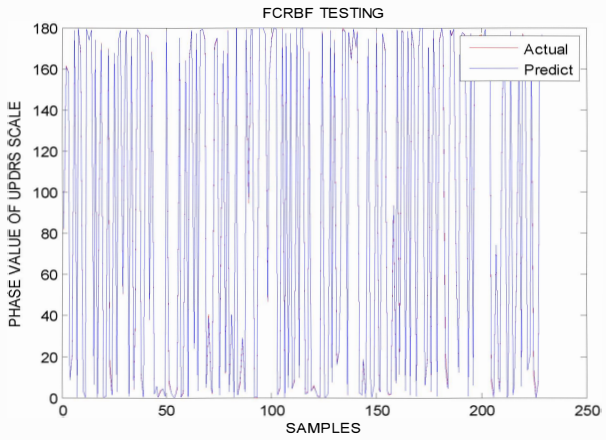

Figure 5.4 FC-RBF: Prediction of Phase value of UPDRS scale for testing data

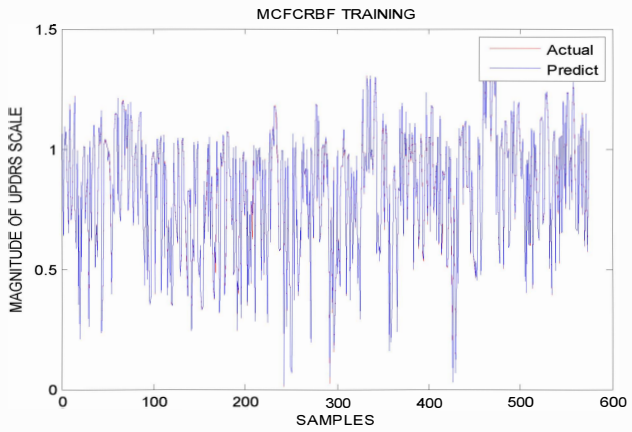

Figure 5.5 Mc-FCRBF: Prediction of magnitude of UPDRS scale for training data

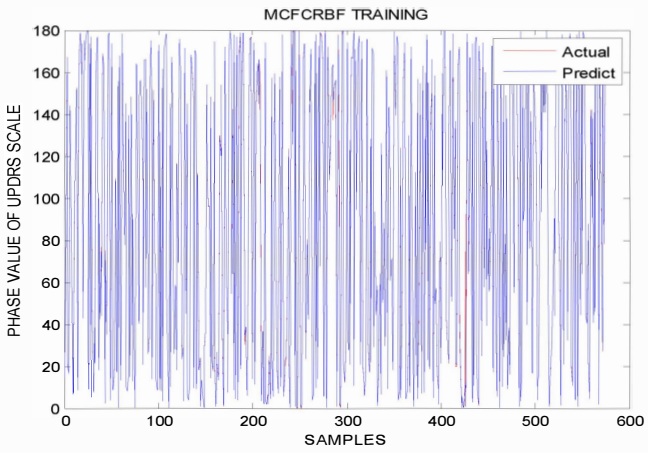

Figure 5. 6 Mc-FCRBF: Prediction of Phase value of UPDRS scale for training data

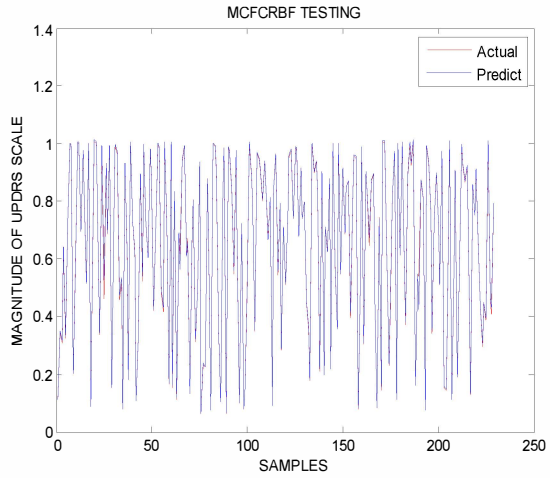

Figure 5.7 Mc-FCRBF: Prediction of magnitude of UPDRS scale for testing data

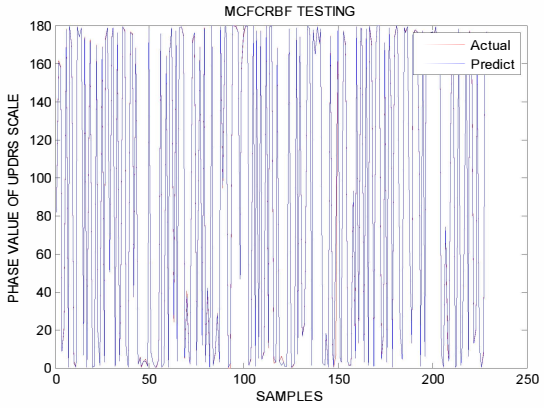

Figure 5.8 Mc-FCRBF: Prediction of Phase value of UPDRS scale for testing data

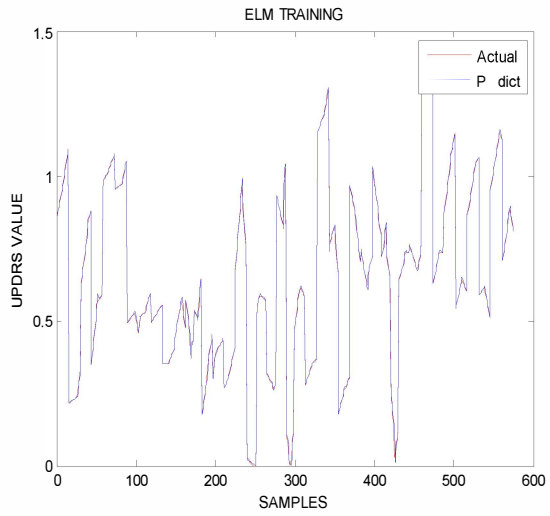

Figure 5. 9 ELM Validation curve

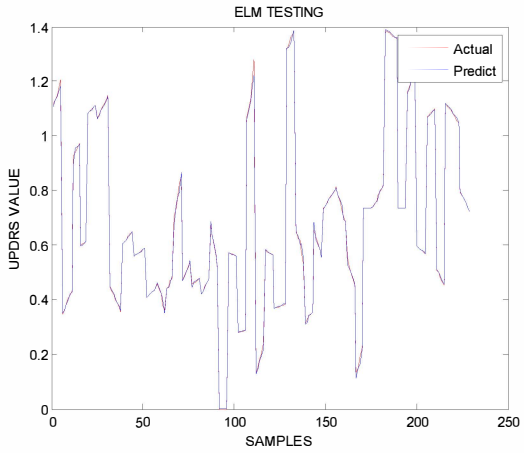

Figure 5.10 ELM Testing curve

TABLE I. UPDRS VALUE PREDICTION- VALIDATION

Network	No of hidden neurons	Root Mean Square Error for Magnitude of UPDRS scale	Root Mean Square for Phase value of UPDRS scale
FC-RBF	15	0.0091	0.2
Mc-FCRBF	15	0.0083	0.2
ELM	15	0.0124	-

TABLE II. UPDRS VALUE PREDICTION- TESTING

Network	No of hidden neurons	Root Mean Square Error for Magnitude of UPDRS scale	Root Mean Square for Phase value of UPDRS scale
FC-RBF	15	0.003	0.2
Mc-FCRBF	15	0.003	0.2
ELM	15	0.0088	-

VI. CONCLUSION

The recently developed Meta-cognitive Fully Complex-valued Radial Basis Function (Mc-FCRBF) network has been applied for predicting the Parkinson's disease. Performance comparison of the meta-cognitive fully complex-valued RBF network (Mc-FCRBF) applied for Parkinson's disease prediction shows better prediction of the disease when compared to a real-valued extreme learning machine and FC-RBF network. The improvement in performance is attributed to the self-regulatory learning mechanism of the meta-cognitive component.

ACKNOWLEDGEMENT

We are grateful to Dr.Suresh Sundaram, Professor Nanyang Technological University, Singapore for comments on early drafts of the paper and valuable suggestions that prompted improvements to the paper.

REFERENCES

[1] Athanasios Tsanas, Max A. Little, Patrick E. McSharry, Lorraine O. Ramig, "Accurate telemonitoring of Parkinson's disease progression by non-invasive speech tests," TBME-00652, 2009.

[2] A.E. Lang, A.M. Lozano, "Parkinson's disease – First of two parts," New England Journal Medicine, 339, 1044-1053, 1998.

[3] G. S. Babu, S. Suresh, B. S. Mahanand, "A novel PBL-McRBFN-RFE approach for identification of critical brain regions responsible for parkinson's disease," Expert System with Applications, 41(2), pp. 478-488, 2014.

[4] Meysam Asgari and Izhak Shafran "Predicting Severity of Parkinson's Disease from Speech," IEEE, 2010.

[5] R.Savitha, S. Suresh, and N. Sundararajan,"A Fully complex-valued radial basis function network and its learning algorithm," International Journal of Neural Systems 19(4), pp. 253-267, 2009.

[6] R. Savitha, S. Suresh, and N. Sundararajan, "Meta-cognitive Learning in a Fully Complex-valued Radial Basis Function Network," Neural Computation, vol.24, no. 5, pp. 1297-1328, 2012.

[7] R.Savitha, S. Suresh, N. Sundararajan and P. Saratchandran, "A new learning algorithm with logarithmic performance index for complex-valued neural networks," Neurocomputing, vol. 72, no. 16-18, pp. 3771- 3781, 2009.

[8] R.Savitha,S.Suresh,N.sundararajan, "A Self-Regulated learning in fully complex-valued Radial basis Function Networks," International joint conference on Neural Networks pp. 1-8, 2010.

[9] Guang-Bin Huang, Q.Y. Zhu, C. k. Siew, "Extreme Learning Machine: Theory and Applications," Neurocomputing 70(1-3)(2006) 489-501.

[10] S.Suresh, S. N. Omkar, V. Mani, T. N.Guru Prakash, "Lift coefficient prediction at high angle of attack using recurrent neural network," Aerospace Science and Technology,vol. 7, no. 8, pp. 595-602, 2003.

[11] Athanasios Tsanas, Max A. Little, "UCI machine learning repository," 2009. URL: http://archive.ics.uci.edu/ml/datasets/Parkinsons+Telemonitoring.

[12] S. Suresh, N. Sundararajan and R. Savitha, Supervised Learning Algorithm in Complex Domain, Springer-Verlag, 2012.

[13] S. Suresh, R. Savitha and N. Sundararajan, Meta-cognitive complex-valued relaxation neural network and its sequential learning algorithm, Recent Trends in Complex-valued Neural Networks, 2012.

[14] S. Suresh, R. Savitha, and N. Sundararajan, "A Projection Based Fast Learning Fully Complex-valued Relaxation Neural Network," IEEE Transactions on Neural Networks and Learning Systems, 24(4), 529-541, 2013. (I.F.: 2.633)

[15] R. Savitha, S. Suresh and N. Sundararajan, "Fast Learning Complex-valued Classifiers for Real-valued Classification Problems," International Journal of Machine Learning and Cybernetics (Accepted, DOI: 10.1007/s13042-012-0112-x), 2012.

[16] R. Savitha, S. Suresh, and N. Sundararajan, Metacognitive learning algorithm for a fully complex-valued relaxation network, Neural Networks, 32, pp. 309-318, 2012

[17] R.V. Babu, S. Suresh, and R. Savitha, Human action recognition using a fast learning complex-valued classifier, Neurocomputing, 89, pp. 202-212, 2012 .

[18] R. Savitha, S. Suresh, and N. Sundararajan, Fast learning circular complex-valued extreme learning machine for real-valued classification problems, Information sciences, 187(1), pp. 277-290, 2012.

[19] R. Savitha, S. Suresh, N. Sundararajan and H. J. Kim, A fully complex-valued radial basis function classifier for real-valued classification problems, Neurocomputing, 78(1), pp. 104-110, 2012.

[20] S. Suresh, R. Savitha and N. Sundararajan, A sequential learning algorithm for complex-valued self-regulatory resource allocation network, IEEE Trans. on Neural Networks, 22(7), pp. 1061-1072, 2011.

[21] R. Savitha, S. Suresh and H. J. Kim, "A Meta-cognitive Learning Algorithm for an Extreme Learning Machine Classifier," Cognitive Computation (Accepted), 2013.

[22] K. Subramanian, S. Suresh, N. Sundararajan, "A meta-cognitive neuro-fuzzy inference system (McFIS) for sequential classification problems, IEEE Trans. on Fuzzy Systems, 2013

[23] G. Sateesh Babu, S. Suresh, Sequential projection based metacognitive learning in a Radial basis function network for classification problems, IEEE Trans, on Neural Networks and Learning Systems, 24(2), pp. 194-206, 2013.

[24] K. Subraminan, and S. Suresh, Human action recognition using meta-cognitive neuro-fuzzy inference system, International Journal of Neural Systems, 22(6), 2012.

[25] K. Subraminan, and S. Suresh, A meta-cognitive sequential learning algorithm for neuro-fuzzy inference system, Applied soft computing, 12(11), 36703-3614, 2012.

[26] G. Sateesh Babu, and S. Suresh, Meta-cognitive RBF networks and Its Projection based Learning Algorithm for Classification problems, Applied Soft Computing, 13(1), pp. 654-666, 2013.

[27] G. Sateesh Babu, and S. Suresh, Parkinson's Disease Prediction using Gene Expression – A Projection based Learning Metacognitive Neural Classifier Approach, Expert System with Applications, 40(5), pp. 1519-1529, 2012.

[28] R. Savitha, S. Suresh, and N. Sundararajan, Metacognitive learning algorithm for a fully complex-valued relaxation network, Neural Networks, 32, pp. 309-318, 2012

[29] G. Sateesh Babu, and S. Suresh, Meta-cognitive neural network for classification problems in a sequential learning framework, Neurocomputing, 81(1), pp. 86-96, 2012.

[30] G.B. Huang, M. Li, L. Chen, C.K. Siew, Incremental extreme learning machine with fully complex hidden nodes, Neurocomputing 71 (4–6) (2008) 576–583.

[31] G.B. Huang, X. Ding, H. Zhou, Optimization method based extreme learning machine for classification, Neurocomputing 74 (1–3) (2010) 155–163.

[32] Vasiliy, S, S. Suresh, R. Savitha and H. J. Kim, "An ensemble of extreme learning machines for Steaganalysis based on risk sensitive hinge loss function," cognitive Computation, 2013 (Accepted)

[33] S. Saraswathi, S. Suresh, N. Sundararajan, M. Zimmermann, and M. N., Hamilton, ICAG-PSO-ELM approach for Accurate Multi-Class Cancer Classification Resulting in Reduced Gene Sets in Which Genes Encoding Secreted Protein are Highly Represented, IEEE/ACM Trans. On Computational Biology and Bioinformatics, 8(2), pp. 452-463, 2011.

[34] S. Suresh, S. Saraswathi and N. Sundararajan, Performance enhancement of extreme learning machine for multi-category sparse data classification problems, Engineering Application of Artificial Intelligence, 23 (7), pp. 1149-1157, 2010

[35] S. Suresh, R. Venkatesh Babu and H. J. Kim, No-reference Image Quality Assessment using Modified Extreme Learning Machine Classifier, Applied Soft Computing, 9(2), 541-552, 2009.