Original Research Article

J. Computing & Int. Systems (2020) 93–98

Journal of Computing and Intelligent Systems

Journal homepage: www.shcpub.edu.in

SACRED HEART RESEARCH PUBLICATIONS

ISSN: 2456-9496

Combined Fuzzy and Projection based Learning in META-Cognitive Neural Network for MAMMOGRAM classification

S. Padma #1, **R. Pugazendi**#2

Received on 17th JAN 2020, Accepted on 31st JAN 2020

Abstract *— Breast cancer is the life-threatening disease nowadays especially for the women. Several methods have been developed in early identification of this disease. The traditional RBF network when combined with fuzzy and projection based learning performs better in benchmark datasets. This concept is combined with the meta cognitive principles. The McNN has two components namely the cognitive component and the meta cognitive component. The cognitive component hold the normal RBF and the meta cognitive the copy of the cognitive component along with learning strategies. In this article the MIAS mammogram dataset is taken and several trials were taken to prove the betterment of the proposed method compared to support vector machine (SVM) and self-regulatory resource allocation network (SRAN).*

Keywords – Nural Network , Breast cancer, Correlation

1. INTRODUCTION

Breast cancer is the most widely recognized malignant growth among women around the globe. Regardless of tremendous medicinal advancement, cancer malignant growth has still remained the subsequent driving reason for death around the world in this manner, its initial finding significantly affects diminishing mortality. Be that as it may, it is frequently hard to analyze cancer variations from the norm [1]. Various instruments have been created to screen cancer malignant growth. Along these lines, the PC enables radiologists to distinguish breast irregularities by utilizing image processing and man-made brainpower apparatuses [2]. Cancer malignancy is the most widely recognized disease and the subsequent driving reason for death among women around the world. (i) Breast malignant growth happens when the cell tissues of the cancer become irregular and wildly separated. These unusual cells structure an enormous chunk of tissues, which thus turns into a tumor. (ii) It was accounted for that 1.7 million instances of cancer disease were distinguished on the planet in 2012 and given in figure 1.1. Cancer malignant growth is the second reason for disease demise with the institutionalized death pace of 12.9 percentage of 100,000 and its frequency has expanded over the years. Along these lines, it is of significance to have proper strategies for screening the soonest indications of cancer.

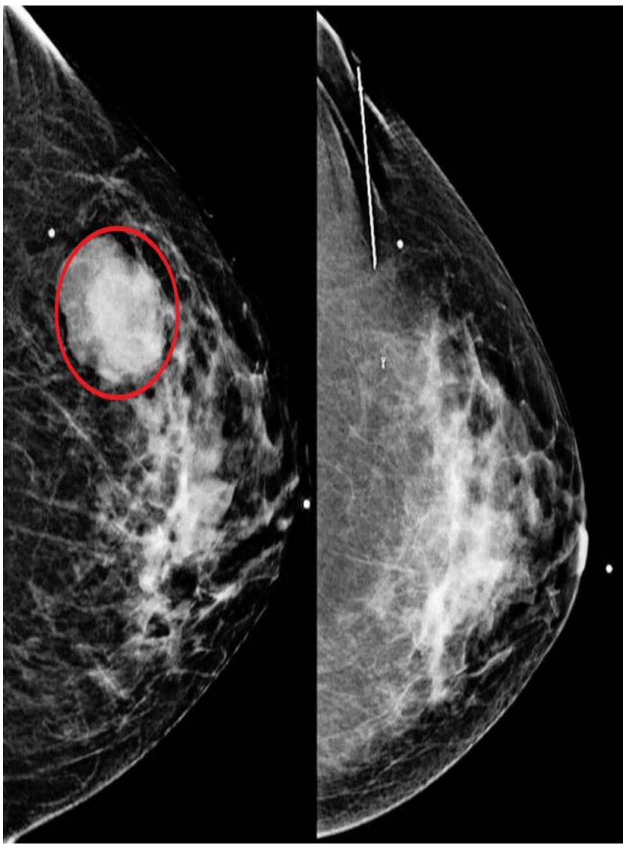

Figure 1.1 Images of Breast Affected by Cancer and Normal Breast

* Corresponding author: E-mail: padmaselvaraj25@gmail.com, Pugazendi_r@rediffmail.com

[1]Research Scholar, Research and Development Center, Bharathiyar University Tamilnadu, India.

[2]Assistant Assistant Professor & PG and Research Department of Computer Science& Government Arts College (Autonomous),Salem-7, Tamilnadu, India.

There are different imaging techniques accessible for the screening and determination of breast cancer detection. Mammography is one of the most significant early finding strategies for cancer disease. Since mammography is not exceptionally effective for thick cancers, ultrasound or analytic sonography procedures are found and uncertainty is not present that assessment of patient information and master judgment is the most significant factor in picture based conclusion in any case, there are numerous components influencing this sort of analysis [3]. The normal and the cancer affected breast are given in figure 1.1. The encircled part is the cancer affected part. Among the variables influencing picture based finding are the nearness of clamor in pictures, the radiologist's visual observation capacity, deficient lucidity, poor difference, and the less experienced radiologist [4]. The recognizable proof of cancer and the order of masses on mammographic pictures are not inconsequential assignments for thick cancers, and it is a test for man-made brainpower and example acknowledgment [5]. Due to inherent problems related with a picture, including poor difference, commotion, and absence of the acknowledgment with the eye, instruments have been created to make and create picture handling. As of now, restorative picture preparing is one of the quickest developing regions in the human services division. The reason for the picture handling is to utilize methods for making legitimate pictures of the human body, which are dependable for use in the determination and treatment processes. The conventional RBF organize when joined with fluffy and projection based learning performs better in benchmark datasets. This idea is joined with the meta-intellectual standards. The cognitive component holds the normal RBF and the meta-cognitive the copy of the cognitive component along with learning strategies [6]. In this research the MIAS mammogram dataset is taken and several trials were taken to prove the betterment of the proposed method SRAN.

2. RELATED WORKS

Neural Network are extensively applied for classification and employed to solve real word problems]. The classification through neural system should be possible either by successive learning or bunch learning. In group learning, the whole preparing informational collection isolated into clusters and took into account learning, the clump learning techniques require the whole preparing information must be executed progressively number of times so as to lessen the estimate blunders. This kind of learning requires more memory and time as the whole preparing information is required for execution. Single Layer Feed-forward Network (SLFNs) arbitrarily picks info loads and concealed layer predispositions. The shrouded hubs in SLFN can be called arbitrary concealed hubs. Every one of the parameters concerning SLFN should be tuned. To conquer these different consecutive calculations were characterized. Successive learning calculations can likewise be utilized for ordering the information. The example enters the system engineering one-by-one and the examples are disposed of the system after the learning procedure is over.

Sequential learning methods are used for better classification. Goal of the arrangement issue is to get familiar with the evaluation surface that superbly maps an info highlight to a yield highlight of class marks. In a successive learning structure the preparation tests arrives individually and are disposed after the learning process. The successive calculation in the neural system recovers learning about the data contained in the surge of information by utilizing every one of the examples in the preparation informational index. The current consecutive calculation works on conveying consistently the example in the information space. By adapting every one of the examples in the preparation informational index may result in preparing the example in the thickly populated locale of the info space. The learning process can be made effective when the learners do self-regulation in the learning using Meta-cognition. Radial Basis Function Neural Network shown in figure 1.2 has been generally utilized for taking care of arrangement issues and capacity approximations [7].

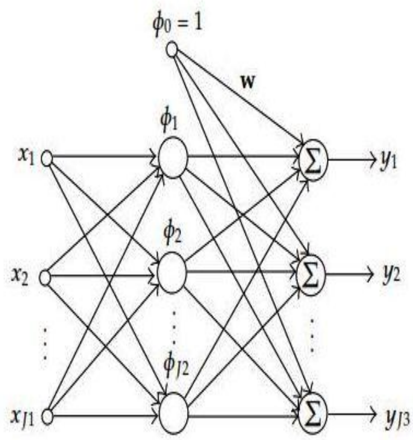

Figure 1.2 Architecture of the RBF Network

The figure 1.2 has three layers. The first layer contains the inputs received from the dataset which can be a normalized one. The second layer is the hidden layer where the required weights are updated to the input. The third layer contains the initiation work for giving exact yield. Clustering is a data analysis tool for describing the dissemination of an informational index and is normally utilized for deciding the RBF focuses. The preparation set is assembled into suitable bunches whose models are utilized as RBF focuses. The quantity of groups can be specified or decided naturally relying upon the bunching calculation. The presentation of the grouping calculation is essential to the efficiency of RBF system learning. Clustering for example, the C-implies is prominent for bunching RBF focuses. RBF focuses dictated by directed bunching are typically more efficient for RBF system learning than those controlled by unaided grouping, since the circulation of the yield examples is likewise considered. At the point when the RBF system is prepared for classification, the LVQ1 calculation is well known for grouping the RBF focuses. Any unaided or directed grouping calculation can be utilized for clustering RBF centers.

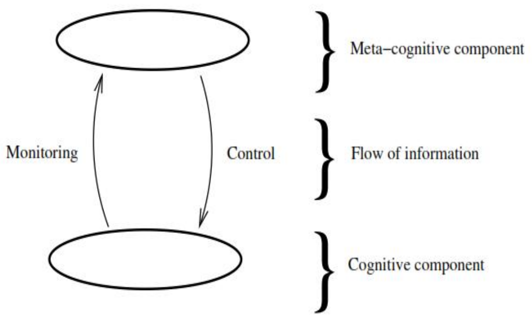

Figure 1.3 Nelson and Narens Model of Meta-cognition

Learning of the network can be improved by incorporating the human learning techniques. Cognition is a cluster of mental activities like attention, understanding, analyzing etc. It includes mental representations like learning, processing it through reasoning, executing by means of problem solving and decision making techniques. Among several meta-cognition models Nelson and Narens [8] proposed a simple model based on meta-cognition represented in figure 1.3 which covers the cognitive part, meta-cognitive or meta-intellectual part and both are connected through controls.

These cognitive principles when included in the learning methods of samples give better results. The cognitive component has three layers, the info layer maps all highlights to the concealed layer without doing any changes. Gaussian enactment capacity is utilized in the concealed layer while straight actuation capacity is utilized in the yield layer for approximating the choice capacity. The meta-intellectual segment utilizes the accompanying measures to recover learning from the preparation tests. The measures are class name estimation, greatest pivot blunder, certainty of classifier and class-wise criticalness. Utilizing these measures together with the standards of self-controlled human learning the preparation tests are found out. The learning procedures are test erase system, neuron development methodology, parameter update technique and test save procedure.

The drawbacks of the existing algorithms are:

- Allocation of new hidden neuron center without knowing the previous center which leads to misclassification.
- Minimize the error function in estimating output weights during classification in RBF.
- Reducing the time taken for allocation of center for Gaussian function.

Hence there is a need to develop a learning method which identifies the center and width of the Gaussian function automatically for better classification. Also the proposed method is applied to classify medical images for better diagnosis in a minimum time. The proposed method is incorporating the fuzzy concepts to identify the center of the Gaussian function and projection based learning to update the parameter weight.

These features while included in the meta-cognitive neural network perform well when compared with the traditional support vector machines and with self-regulatory resource allocation network.

3. Methodology

3.1Fuzzy Radial Basis Function Network

Integrating the concepts of FCM with RBF develops an effective model for various real time problems. Determining the centers and weight of the hidden layer is the primary task of the three layered RBF network. The centers maybe identified either randomly or by clustering or through a learning procedure. Random selection of centers may lead to undesirable performance if the data is large. If the selections of centers are not satisfied then random selection must continue until it gives desired performance. The learning process can be one phase, two phases and three phases for the selection of parameters like center, width and weight. Identification of centers through clustering gives a desirable performance when compared to other methods.

The Fuzzy C-means calculation is an iterative calculation that discovers bunches in unlabeled information which is joined with the hypothesis of fluffy participation as an option of passing on a pixel to a solitary group as far as comparative and every pixel will have divergent enrollment esteems on each bunch. Fluffy grouping has been connected effectively in different fields, including geological looking over, money or promoting. It is done dependent on minimization of the target capacity is pursues as,

$$J_m = \sum_{j=1}^{c} \quad \sum_{i=1}^{n} \mu_{ij} d_{ij} \quad ...(1)$$

$\mu_{ij}$is membership function, $d_{ij} = |o_i - z_j|$ is the centroid distance $z = \{z_1, z_2, ..., z_n\}$ and data points$O = \{O_1, O_2, ... O_n\}$. The centroids are obtained using the equation

$$z_j = \frac{\sum_{i=1}^{n} \mu_{ij}^{m} o_i}{\sum_{i=1}^{n} \mu_{ij}^{m}} \quad ... (2)$$

membership function is updated as μ_{ij}, i = 1, 2... n; j = 1,2,..., c by the use of

$$\mu_{ij} = \frac{1}{\sum_{k=1}^{c} (\frac{d_{ij}}{d_{ik}})^{\frac{2}{m-1}}} \quad ... (3)$$

Where$d'_{kj}$is the distance in the patterns vector through the centre$v_k$of the kth class.

In the conventional RBF, the exchange capacity of a concealed hub is demonstrated by Gaussian circulation work, which enables a shrouded hub to deliver non-zero reaction, notwithstanding when the information design vector does not coordinate with the relating group focus. The nonzero reaction relies on the fluctuation of the Gaussian appropriation work. Since the Gaussian appropriation move capacity of each concealed hub is its neighborhood property, the yield of each shrouded hub can be processed locally. In the fluffy spiral premise capacity organize, the goal is to play out a fluffy dividing of the information in the concealed layer.

In other words, the destinations to register the participation estimation of any example to a class relating to any covered up node equation 2 without utilizing a Gaussian dispersion type move capacity, and joining the reactions of the shrouded hubs in the yield layer. Be that as it may, the participation estimation of any example to any group relies on the separations of the example from every current bunch equation 3. Subsequently, if the engineering of the Fuzzy RBF is actually equivalent to that of a regular RBF, at that point it is beyond the realm of imagination to expect to process the fluffy enrollment estimations of an example locally [62][70].To do the local computation, equation 3 a modified scheme for Fuzzy RBF is used.

Equation 3 can be written as

$$u_{kj} = \frac{h_k(j)}{\sum_{i-1}^{c} h_i(j)} \quad ...(4)$$

Where

$$h_i(j) = \left(\frac{1}{d_{ij}}\right)^{\frac{2}{m-1}} \quad ...(5)$$

The activation of each output node in the output layer is given as

$$y_1^{(p)} = \sum_{j=1}^{c} w_{ij} h_j(p) \quad ...(6)$$

where$y_i^{(p)}$is considered as the response for the ith output node when$X_p$is the present input

From equations 5 and 6, $y_i^{(p)}$ is re-written as

$$y_i^{(p)} = \frac{1}{H_a^{(p)}} \sum_{j=1}^{c} W_{ij} \,?_j^{(p)} \quad ...(7)$$

Where

$$H_a^{(p)} = \sum_{i=1}^{c} ?_i^{(p)} \quad ...(8)$$

Equations 7 and 8 uncover the way that can be processed locally in the shrouded hubs and the initiation of the yield hubs can be registered from the concealed hub enactments with an extra standardization by the absolute yield in the shrouded layer H. For this reason, it is presented as an assistant concealed hub in the Fuzzy RBF to process the absolute enactment in the shrouded layer, and feed it to the yield layer. The focuses are determined utilizing the FCM technique.

3.2 Projection Based Learning

The Projection-Based Learning algorithm chips away at the standard of minimization of blunder capacity and finds the ideal system yield loads for which the mistake is least.

The error function for $i^{t?}$ sample is defined as

$$J_i = \sum_{j=1}^{n} \quad (y_j^i - \sum_{k=1}^{k} \quad w_{k_j} ?_k^i)^2 \quad ...(9)$$

For t training samples, the overall error function is defined as

$$J(W) = \frac{1}{2}\sum_{i=1}^{t} \quad J_i = \frac{1}{2}\sum_{i=1}^{t} \quad \sum_{j=1}^{n} \quad (y_j^i - \sum_{k=1}^{K} \quad w_{k_j} ?_k^i)^2 \quad ...(10)$$

$$W* = \arg \min_{W \in ?^{K \times n}} J(W) \quad ...(11)$$

3.3 Projection Based Learning Fuzzy Meta-cognitive Neural Network (PBLF-McNN)

Radial Basis Function neural system is a neural system model that executes with the overall estimation. Data layer contains the estimations of the data. The covered layer contains the incitation work which goes about as commitment for the yield layer. The yield layer solidifies the contribution with its initiation capacity and conveys reasonable contribution with included loads. By utilizing bunching procedures for choosing focus and width of concealed neurons in RBF, which reflects in increasingly precise dispersion in the information focuses. Fluffy C-Means is a fluffy bunching calculation which is utilized to locate the inside and width of the Radial Basis Function as opposed to picking haphazardly.

Radial Basis Function performs a typical nonlinear kind of mapping from the input area R^n to the outer space R^m. R^n is vector for the input and it is denoted as x_i and R^m is the vector for the output and it is denoted as y_i. The j^{th} neuron that is considered as hidden in the Radial Basis Function, that calculates the value for Gaussian function as follows

$$Z_i(x) = exp(-\frac{\|x-c_j\|}{2\sigma_i^2})j = 1,2,\ldots.m \quad ...(12)$$

Here x is taken as input feature n dimension. Vector$c_j$terms to be the of Gaussian vector of i and σ_i and the width.The hidden node's center position is then calculated using the unsupervised learning such as Fuzzy C-Means as followed by equation 12.

The width $\sigma_i$is computed by

$$\sigma_j = \sqrt{\frac{1}{m_j}\sum_{i=1}^{m_j} d^2(c_j - x_i)} \quad ...(13)$$

When the sample was made used for the updation in the weight of the output, the equation becomes,

$$\frac{\partial J(W_K^t)}{\partial w_{pj}} = \frac{\partial J(W_K^t)}{\partial w_{pj}} + \frac{\partial J_t(W_K^t)}{\partial w_{pj}} = 0,$$

$$p = 1, \ldots..K; \; j = 1, \ldots, n \quad ...(14)$$

With intended to 0 and by equating 1^{st} partial derivative and re-ordering the

$$(A^{t-1} + (h^t)^T W_k^t - (B^{t-1} + (h^t)^T (y^t)^T) = 0 \quad ... (15)$$

By substitute $B^{t-1} = A^{t-1} W_K^{t-1} A^{t-1} + (h^t)^T h^t = A^t$ and adding or subtracting the term $(h^t)^T h^t W_K^{t-1}$ on both sides of equation 15 is reduced to

$$W_K^t = (A^t)^{-1} \left(A^t W_K^{t-1} + (h^t)^T (y^t)^T - h^t W_K^{t-1}\right) \quad ... (16)$$

At the end , the output weight obtained is

$$W_K^t = W_K^{t-1} + (A^t)^{-1} (h^t)^T (e^t)^T \quad ...(17)$$

The meta-psychological segment contains duplicate of the subjective part. At the point when another preparation test arrives, the meta-psychological segment of McNN predicts the class name and gauge the information present in the new preparing test as for the intellectual part. In light of this data, the meta-psychological segment chooses an appropriate learning system, for the present example. Along these lines, tending to the three central issues in learning process: (a) what-to-learn, (b) when-to-learn and (c) how-to-learn. To begin with, it present the subjective segment and next it feature different learning methodologies of the meta-psychological part.

3.4 Performance Measures

The MIAS Digital Mammogram Database contains 322 pictures Mediolateral Oblique speaking to 161 reciprocal Mammogram sets. The database is isolated into seven classes, for example, smaller scale calcifications, encircled masses, estimated injuries, badly characterized masses, engineering contortion and awry densities of ordinary and strange Mammograms. Sensitivity is to measure the correctly classified content. Specificity is to measure the actual negatives that the classifier done.

$$Sensitivity = TP/(TP+FN) \ldots (18)$$
$$Specificity = TN/(TN+FP) \ldots (19)$$

Where TP = True Positive, TN = True Negative, FP = False Positive and FN = False Negative. Area Under Curve (AUC) to check the performance of the classifiers.

4. Results and Discussions

The dataset is divided into training and testing dataset. The classification is executed in 10 combinations in different classifiers namely SVM, SRAN and PBLF-McNN shown in table 1.1 and figure 1.4. The experimental outputs from the simulated platform are projected in figures 1.5, 1.6 and 1.7. The results prove that the sensitivity and the specificity of the proposed ROC- PBLF-McNN method have produced greater value when compared with that of the standard SVM and SRAN.

Table 1.1 Combinations in different classifiers

Combinations	SVM (%)	SRAN (%)	PBLF- McNN(%)
C1	80.7	92	95.71
C2	84.1	92	95.71
C3	88.3	93	96.9
C4	86.4	93	96.67
C5	87.5	92	95.71
C6	88.6	91	94.86
C7	89.6	91	94.83
C8	88.3	92	95.06
C9	87.7	91	96.67
C10	88.3	91	97.03

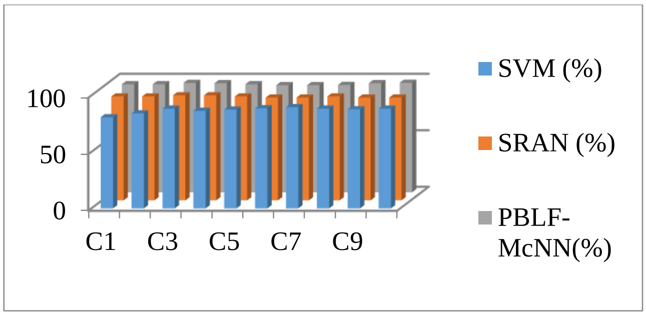

Figure 1.4 10 Combinations Trials of MIAS Dataset

The result on every trial clearly shows that the proposed method performs better when compared to the existing one. The experiments were simulated and the results are shown in the table 1.2 and figures 1.5, 1.6, 1.7 and 1.8. The specificity was observed in three different approaches namely the ROC- SVM, ROC-SRAN and ROC- PBLF-McNN. The threshold probability was set to 0.001. The experiments were conducted in three criteria values. It is seen that the proposed ROC- PBLF-McNN shows greater specificity.

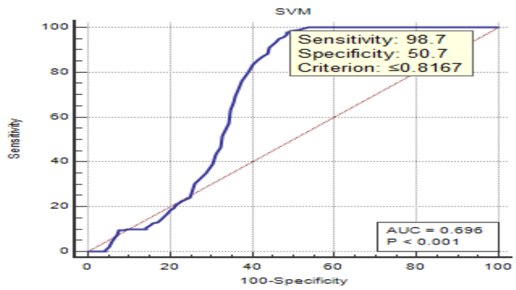

Figure 1.5 Specificity Measures of ROC- SVM

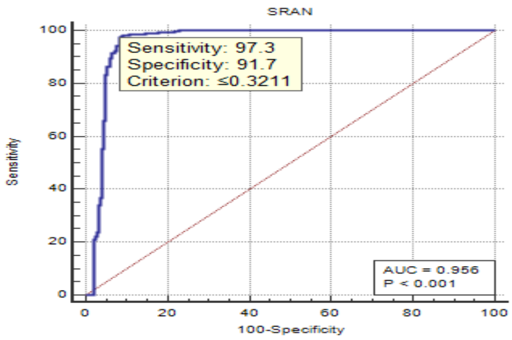

Figure 1.6 Specificity measures of ROC-SRAN

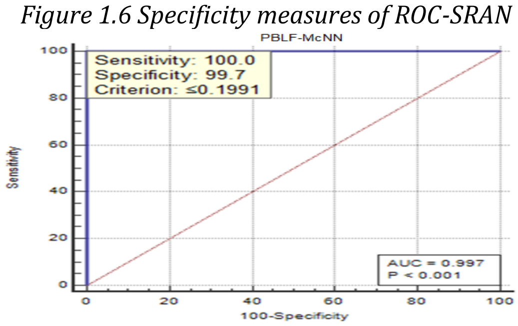

Figure 1.7 Specificity measures of ROC - PBLF-McNN

Table 1.2 Performance Measures of Various Classifiers

Classifiers	Specificity	Sensitivity
SVM	98.7	50.7
SRAN	97.3	91.7
PBLF-McNN	100	99.7

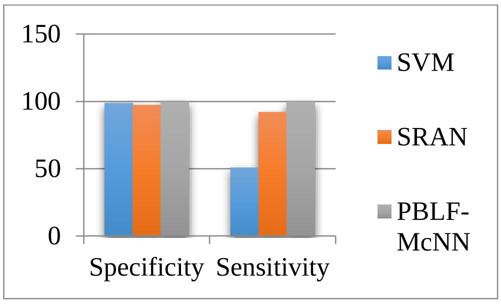

Figure 1.8 Specificity and Sensitivity Analysis of Classifiers

5. Conclusion

This paper gives a clear view that the self-regulatory concepts implied in the machine learning concepts works well. The identification of center and width using the fuzzy and projection based learning concepts improves the generalization of the classifier. The meta- cognitive component works on the principles of what-to-learn, when-to-learn and how-to-learn gives an additional force in the performance of classification. Breast cancer is a life threatening disease which can be identified at the initial stage through such efficient classifier.

REFERNECE

[1] Tingting Mu, Asoke K. Nandi, “Breast Cancer Diagnosis and Prognosis using Different Kernel-Based Classifiers”, Biosignals, Vol. 2, 2008, pp 342-348.

[2] C. J. Vyborny, “Can computers help radiologists read mammograms?”, Radiology, Vol 191,1994, pp: 315-317.

[3] C. L. Chen, W. C. Chen, and F. Y. Chang, “Hybrid learning algorithm for Gaussian potential function networks”, IEEE Proceedings D, Vol 140, Issue 6, November 1993, pp. 442 – 448.

[4] Igor Kononenko, “Machine learning for medical diagnosis: history, state of the art and perspective”, Artificial Intelligence in Medicine, Vol.23, Issue 1, August 2001, pp. 89-109.

[5] Javad Salimi Sartakht, Makthar A S , “Hepatitis disease diagnosis using a novel hybrid method based on Support Vector Machine and Simulated Annealing”, Computer Methods and Programs in Biomedicine, Vol.108, Issue 2, November 2012, pp. 570-579.

[6] PadmaS. and ArunK.M.,“Classification of Escherichia Coli Bacteria using Meta-Cognitive Neural Network ”, International Journal of Computer Science & Engineering Technology, Vol. 5, Issue 2, February 2014,pp. 141 – 149.

[7] Esmaeili, A., Mozayani, N., “Adjusting the parameters of radial basis function networks using particle swarm optimization”, 2009 IEEE International Conference on Computational Intelligence for Measurement Systems and Applications, IEEE, pp. 179-181.

[8] T. O. Nelson and L. Narens, Metamemory: A theoritical framework and new findings. Boston, USA: Allyn and Bacon, 1992.