Available Online at www.ijcsmc.com

International Journal of Computer Science and Mobile Computing

A Monthly Journal of Computer Science and Information Technology

ISSN 2320–088X

IJCSMC, Vol. 4, Issue. 10, October 2015, pg.69 – 78

SURVEY ARTICLE

A SURVEY ON NEURAL NETWORK MODELS FOR DIAGNOSTIC PROBLEMS

[1]Deepak Pal, [2]Dinesh Kumar Sahu, [3]Anil Rajput

[1]M.Tech Scholar, Department of CSE, Sri Satya Sai College of Engineering, Bhopal M.P., India

[2]Ph.D Scholar, Department of Computer Science, Barkatullah University, Bhopal M.P., India

[3]Professor, Department of Mathematics, CSA, Govt. P.G. College, Sehore M.P., India

Deepak21cs@gmail.com, dinesh_sahu20@yahoo.com

Abstract— Neural network plays an important role in Health care. It really helps to predict the disease based on collated data, Diagnosis in the medical field is a complicated task that should be performed with accuracy and efficiency. A diagnosis performed by a physician for a single patient may differ significantly if the same is examined by the other physicians or by the same physicians at different times to that single patient. Now a days, automated medical analysis are used to help this contribution reviews shortly the application of neural network methods to medical problems and characterizes its advantages and problems in the context of the medical background. Successful application examples show that human diagnostic capabilities are significantly worse than the neural diagnostic systems. Then, paradigm of neural networks is shortly introduced and the main problems of medical data base and the basic approaches for training and testing a network by medical data are described.

Additionally, the problem of interfacing the network and its result is given and the optimal Back propagation algorithm approach is presented. Finally, as case study of neural rule based diagnosis septic shock diagnosis is described, on one hand by a growing neural network and on the other hand by a rule based system.

Keywords- Neural Networks – ANN , Optimal Back Propagat-tion algorithm, Diagnosis, Neural Networks

I. INTRODUCTION

Almost all the physicians are confronted during their formation by the task of learning to diagnose. Here, they have to solve the problem of deducing certain diseases or formulating a treatment based on more or less specified observations and knowledge. Certainly, there is the standard knowledge of seminars, courses and books, but on one hand medical knowledge outdates quickly and on the other hand this does not replace own experience. For this task, certain basic difficulties have to be taken into account:

• The basis for a valid diagnosis, a sufficient number of experienced cases, is reached only in the middle of a physician's career and is therefore not yet present at the end of the academic formation.

• This is especially true for rare or new diseases where also experienced physicians are in the same situation as newcomers.

• Principally, humans do not resemble statistic computers but pattern recognition systems. Humans can recognize patterns or objects very easily but fail when probabilities have to be assigned to observations.

These principal difficulties are not widely known by physicians. Also studies who revealed that about 50% of the diagnoses are wrong do not impede the self conscience of some physicians. It is not by chance that the disease AIDS which manifests by a myriad of infections and cancer states was not discovered directly by treating physicians but by statistical people observing the improbable density

Of rare cancer cases at the U.S. west coast. An important solution for the described problem lies in the systematic application of statistical instruments. The good availability of computers ameliorates the possibilities of statistically inexperienced physicians to apply the benefits of such a kind of diagnosis:

• Also physicians in the learning phase with less experience can obtain a reliable diagnosis using the collected data of experienced colleagues.

• Even in the case of rare diseases, e.g. septic shock, it is possible to get a good diagnosis if they use the experience of world-wide networked colleagues.

• New, unknown diseases can be systematically documented even if this induces complex computations which are not known to the treating physician.

• Also in the treatment of standard diseases a critical statistical discussion for the use of operation methods or medical therapies may introduce doubts in the physicians own, preferred methods as it is propagated by the ideas of evidence based medicine EBM[6].

Neural Networks are ideal in recognizing diseases using scans since there is no need to provide a specific algorithm on how to identify the disease. Neural networks learn by example so the details of how to recognize the disease are not needed. What is needed is a group of examples that are representative of all the variations of the disease. This paper describes, an algorithm, the feed forward neural network (optimal Back propagation algorithm) constructive algorithm for medical diagnosis. It begins network design in a constructive fashion by adding nodes one after another based on the performance of the network on training data.

II. MEDICAL DIAGNOSIS PROBLEMS

The major task of medical science is to prevent and diagnose the diseases. Here our focus is the second task, which as mentioned before, is not a direct and simple task at All. In 2001, Brause highlighted that almost all the physicians are confronted during their formation by the task of learning to diagnose [4]. Here, they have to solve the problem of deducing certain diseases or formulating a treatment based on more or less specified observations and knowledge [7]. Below some certain difficulties of medical diagnosis that have to be taken into account are listed:

• The basis for a valid diagnosis, a sufficient number of experienced cases, is reached only in the middle of a physician's career and is therefore not yet present at the end of the academic formation.

• This is especially true for rare or new diseases where also experienced physicians are in the same situation as newcomers.

• Principally, humans do not resemble statistic computers but pattern recognition systems. Humans can recognize patterns or objects very easily but fail when probabilities have to be assigned to observations [9].

• The quality of diagnosis is totally depends on the physician talent as well as his/her experiences.

• Emotional problems and fatigue degrade the doctor's performance.

• The training procedure of doctors, in particular specialists, is a lengthily and expensive one. So even in developed countries we may feel the lack of MDs.

• Medical science is one of the most rapidly growing and changing fields of science. New results disqualify the older treats, new cures and new drugs are introduced day by day. Even unknown diseases turn up every now and then. So a physician should always try hard to keep his/herself up to date [4,7,9].

Regarding problems above and also many others, the question would be how computers can help in medical diagnosis. Since decades ago, computers have been employed widely in the medical sector. From local and global patient and medicine databases to emergency networks, or as digital archives, computers have served well in the medical sector. Meanwhile, in the case of medical diagnosis, regarding the complexity of the task, it has not been realistic yet to expect a fully automatic, computer-based, medical diagnosis system. However, recent advances in the field of intelligent systems are going to materialize a wider usage of computers, armed with AI techniques, in that application. A computer system never gets tired or bored, can be updated easily in a matter of seconds, and is rather cheap and can be easily distributed. Again, a good percentage of visitors of a clinic are not sick or at least their problem is not serious, if an intelligent diagnosis system can refine that percentage, it will set the doctors free to focus on nuclear and more serious cases.

III. WHAT IS AN ARTIFICIAL NEURAL NETWORK

Artificial neural networks are developed based on brain structure. Like the brain, artificial neural networks can recognize patterns, manage data and learn. They are made by artificial neurons which implement the essence of biological neurons.

- It receives a number of inputs (from original data or from output of other related neurons). Each input comes via a connection, which is called synapses and which has a weight (coefficient of connectivity). A neuron also has a threshold value. If the sum of the weights is bigger than this value, than the neuron is activated.
- The activation signal produces the output of the neuron. This output can be the result of the problem or can be considered an input for another neuron. To create an artificial neural network is necessary to put together a number of neurons. They are arranged on layers. A network has to have an input layer (which carries the values of outside variables) and an output layer (the predictions or the result). Inputs and outputs correspond to sensory and motor nerves from human body there also can be hidden layer(s) of neurons, which play an internal role in the network. All these neurons are connected together.

IV. HOW ARE USED NEURAL NETWORKS IN MEDICINE

ANNs are very useful for analyzing complex problems where the relationships between input and output data are not very well known, such as pattern and speech recognition, machine vision, robotics, signal processing and optimization. They are also useful in fields where there is a high degree of uncertainty, such as market analysis, analysis and control of industrial processes and medical diagnosis

In the case of civil engineering, the ANNs have already begun to be used in problems of structural Diagnosis or work programming. The present work describes the preliminary results of a research effort aimed at investigating the potential of ANNs in the interpretation of data from Nondestructive Testing (NDT),

Artificial neural networks could be used in every situation in which exists a relationship between some variables that can be considered inputs and other variables that can be predicted (outputs). The most important advantages using artificial neural networks are that this kind of system solves problems that are too complex for conventional technologies, do not have an algorithmic solution or the solution is too complex to be used. These characteristics have often appeared in medicine. Artificial neural networks have been successfully applied on various areas of medicine, such as: diagnostic systems, biomedical analysis, image analysis, drug development. Using artificial neural networks, it can be monitored a lot of health indices (respiration rate, blood pressure, glucose level) or can be predicted the patient response to a therapy. Artificial neural networks have a very important role in image analysis, too, being used together with processing of digital image in recognition and classification. They are used in pattern recognition because of their capacity to learn and to store knowledge. The medical image field is very important because it offers a lot of useful information for diagnosis and therapy [6]. There are also a lot of applications that use neural networks connected with Bayesian statistics

The increasing interest in the use of ANNs can be justified by the successful implementation experiences recorded in different areas. The method has been used in economics, medical and technical research, geology, physics and other fields, for solving and classification problems. The successes obtained derive from some of the very interesting intrinsic properties of a neural net, such as the possibility of non-linear modeling, and from the simple architecture that favorably distinguish them from other analysis methods [2]. An ANN works as a solid massive parallel processor, which is constituted by several simple units and has a natural propensity to store experimental knowledge and use it to create non-linear relationships between inputs and outputs [3]. In other words, an ANN is a highly interconnected network made of many simple processors. Each processor in the network maintains only one piece of dynamic information and is capable of only a few simple computations.

An ANN performs computations by propagating changes in activation between the processors [4]. Using the ANN we can acquire, store and use the knowledge extracted from experts or experiments. The knowledge is kept in a steady state net of relationships between individual neurons and can be updated automatically using some kind of learning Algorithm.

A net contains many paths, which are activated, to a certain degree, by the input vector. The signals generated are propagated and combined trough the various layers of the ANN, stimulating the various neurons, and finally generating the output signals [1].

The basic anatomy of an ANN could be divided into seven basic parts [5]:

- The set of individual processing units, or neurons
- The state of activation of a processing unit;
- The function used to compute the output of a processing unit
- The pattern of connectivity among the processing units

- The rule of propagation employed;
- The activation function for each individual processing unit

V. RELATED WORKS

The current work focuses on the fuzzification and defuzzification of patient data[7]. Since data from the patient are nothing but physiological measures, they are subjected to noise and uncertainty. The data from the patient such as height or weight data cannot always be trusted as they are subjected to the quality and accuracy of measuring units and the skill of the technician. Moreover, based on a single data, it would be highly uncertain to make an accurate decision about the future physiological state of the patient. So the patient data has been fuzzified with the objective of transformation of periodic measures into likelihoods that the Body Mass Index, blood glucose, urea, creatinine, systolic and diastolic blood pressure Of the patient is high, low or moderate.

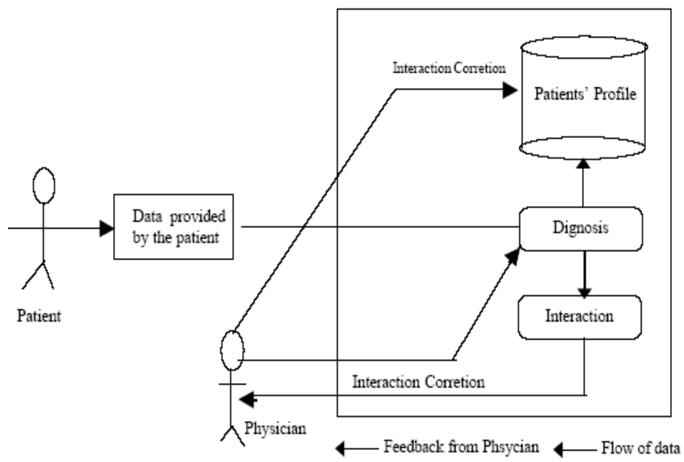

Fig.(1) Patient Diagnosis System

VI. PROPOSED APPROACH

The first stage of this study was decoding the real and typical doctor-based diagnosis procedure by interviewing some experts. A few physicians assisted us to provide the diagnostic flow diagrams of some illnesses, and a list of symptoms that are known to be essential in each case. Fig. 1 Shows two of the diagnosis flow diagrams obtained. Those Diagrams were all double checked by another expert before The selecting of symptoms which was carried out next. In this section, the theoretical background of the Optimal Back propagation learning algorithm pertaining to our study is reviewed.

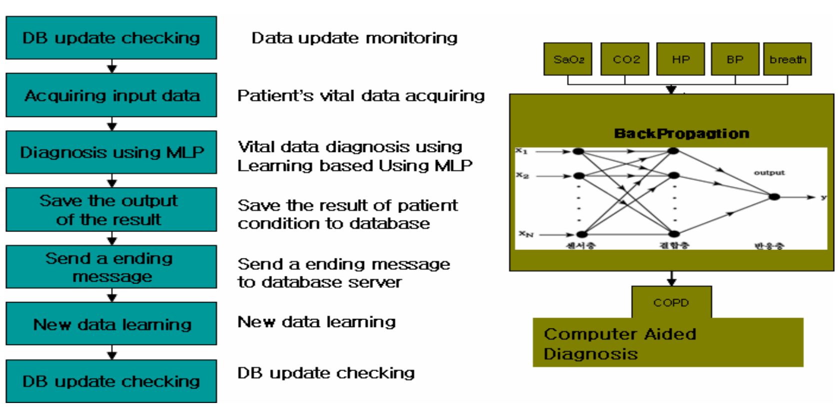

Fig. 2 Computer Aided Diagnosis framework

A. ***Standard Back Propagation Algorithm***

Back propagation neural networks employ one of the most popular neural network learning algorithms, the Back propagation (BP) algorithm. It has been used successfully for wide variety of applications, such as speech or voice recognition, image pattern recognition, medical diagnosis, and automatic controls. The back propagation algorithm trains a given feed-forward multilayer neural network for a given set of input patterns with known classifications. When each entry of the sample set is presented to the network, the network examines its output response to the sample input pattern. The output response is then compared to the known and desired output and the error value is calculated. Based on the error, the connection weights are adjusted. The back propagation algorithm is based on Widrow-Hoff delta learning rule in which the weight adjustment is done through mean square error of the output response to the sample. The procedure of algorithm is following

1. Initialize all the connection weights W with small random values from a pseudorandom sequence generator.

2. Repeat until convergence (either when the error E is below a preset value or until the gradient vE(t)/vW is smaller than a preset value).

2.1 Compute the update using

$$\Delta W(t) = -\eta \frac{\partial E(t)}{\partial W}.$$

2.2 Update the weights with

$$W(t+1) = W(t) + \Delta W(t).$$

2.3 Compute the error E(t+1).

Where t is the iteration number, W is the connection weight, and h is the learning rate. The error E can be chosen as the mean square error (MSE) function between the actual output yj and the desired output dj:

$$E = \frac{1}{2}\sum_{j=1}^{n}(d_j - y_j)^2.$$

An incremental strategy is more efficient than batch training strategy and also faster for systems with large training samples, as random disturbances can be induced to help the system to escape from a local minimum point. (Shigetoshi et al., 1995; Jang et al., 1997)

The BP algorithm described above has some shortcomings. If the learning rate is set small enough to minimize the total error, the learning process will be slowed down. On the other hand, a larger learning rate may speed up learning process at the risk of potential oscillation. Another problem is that, partial minimal points or stable stages on error surface are often encountered during the learning process (Baba, 1989).

Using a momentum term is the simplest method to avoid oscillation problems during the search for the minimum value on the error surface. The weight update in BP algorithm with a momentum term a is defined as follows:

$$\Delta W(t) = -\eta \frac{\partial E(t)}{\partial W} + \alpha \Delta W(t-1).$$

The adaptive learning rate can also be adopted to speed up the convergence of the algorithm. For batch training strategy, the learning rate can be adjusted as follows

$$\eta(t) = \begin{cases} \beta\eta(t-1) & \text{if } E(t) < E(t-1). \\ \theta\eta(t-1) & \text{if } E(t) > kE(t-1). \\ \eta(t-1) & \text{otherwise.} \end{cases}$$

where h(t) is the learning rate at the tth iteration, and β , θ and k are chosen as such that $\beta > 1, 0 < \theta < 1, and\ k > 1.$ While for the incremental training strategy, learning rate can be updated using

$$\eta(t) = \eta_0 + \lambda E(t-1).$$

The learning algorithm with forgetting mechanics is an algorithm that can 'forget' unused connections (Takeshi, 2001). With this forgetting mechanism, the weights that are not reinforced by learning will disappear. The obtained network, thus, has a skeletal structure that reflects the regularity contained in the data, useful to improve the convergence and the network accuracy. In general, the updating of connection weights with forgetting mechanics term is given by

$$\Delta W'(t) = \Delta W(t) - \varepsilon \operatorname{sgn}(W(t)).$$

Where ε is the amount for the forgetting, and sgn(x) is the sign function. The absolute value of connection weight is set to decrease by ε due to the second term on the right-hand side). In practice, some optimization algorithms are often used to improve the network convergence (Gill et al., 1981), such as the steepest descent method, the Newton method, In practice, some optimization algorithms are often used to improve the network convergence (Gill et al., 1981), such as the steepest descent method, the Newton method, the Quasi–Newton method, and the conjugate gradients method. In this study, the conjugate gradients method is adopted, as it has a low computation cost and exhibits good results (Polak, 1971). The connection weights thus can be expressed by:

$$W(t+1)) = W(t) + \eta(t)d(t).$$
$$d(t) = -\nabla E\left[W(t)\right] + \beta(t)d(t-1).$$
$$d(O) = -\nabla E\left[W(O)\right].$$

Where ∇ E is the gradient, d(t) is conjugate gradient, h(t) is the step wide, b(t) is determined given by Polak–Ribiere function shown

*B. **Optimal Back Propagation Algorithm***

In this section, the adjustment of the new algorithm OBP will be described at which it would improve the performance of the BP algorithm. The convergence speed of the learning process can be improved significantly by OBP through adjusting the error, which will be transmitted backward from the output layer to each unit in the intermediate layer. In BP, the error at a single output unit is defined as:

$$\delta^0{}_{pk} = (Y_{pk} - O_{pk}).$$

Where the subscript "P "refers to the pth training vector, and "K "refers to the kth output unit. In this case, Ypk is the desired output value, and Opk is the actual output from kth unit, then δpk will propagate backward to update the output-layer weights and the hidden-layer weights. While the error at a single output unit in adjusted OBP will be as:

$$New\delta^{0}{}_{pk} = (1 + e^{(Y_{pk} - O_{pk})^2}) \quad , if(Y_{pk} - O_{pk}) \geq zero.$$

$$New\delta^{0}{}_{pk} = -(1 + e^{(Y_{pk} - O_{pk})^2}) \quad , if(Y_{pk} - O_{pk}) < zero.$$

Where New δ˚pk is considered as the new proposed in the OBP algorithm. An OBP uses two forms of Newδ˚pk, because the exp function always returns zero or positive values (and the adapts operation for many output units need to decrease the actual outputs rather than increasing it). This New δ˚pk will minimize the errors of each output unit more quickly than the old δ˚pk, and the weights on certain units change very large from their starting values.

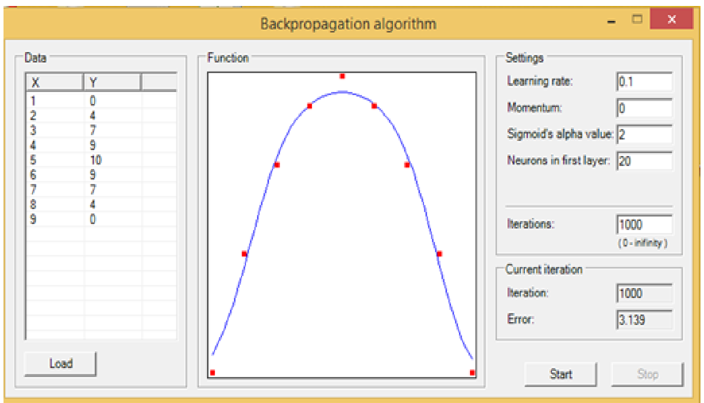

Fig. 3 A Optimal Back Propagation Algorithm

C. *Steps of an OBP*

1. Apply the input example to the input units.

2. Calculate the net-input values to the hidden layer units.

3. Calculate the outputs from the hidden layer.

4. Calculate the net-input values to the output layer units.

5. Calculate the outputs from the output units.

6. Calculate the error term for the output units, but replace Newδ˚pk with δ˚pk.

7. Calculate the error term for the output units, using Newδ˚pk, also.

8. Update weights on the output layer.

9. Update weights on the hidden layer.

10. Repeat steps from step 1 to step 9 until the error (Ypk – Opk) is acceptably small for each training vector pairs.

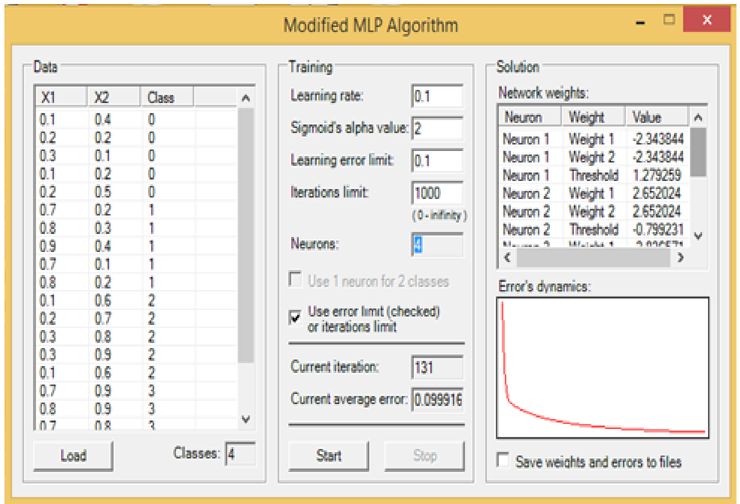

Fig. 4 A Modified MLP Algorithm

VII. Conclusion

In this paper, ANNs have been widely used to solve many problems, they are often we have presented a medical decision support system based on the neural network architecture for Medical diagnosis. The system is trained by employing an improved BP algorithm. The hidden layer of a neural network plays an important role for detecting the relevant features. Due to the existence of irrelevant and redundant attributes, by selecting only the relevant attributes, higher predictive accuracy can be achieved. For a particular input, any (or few) feature(s). This can be extended to other diseases also.

REFERENCES

[1] Long, W. J., Naimi, S., & Criscitello, M. G. (2005). Development of a knowledge base for diagnostic reasoning in cardiology. Computers in Biomedical Research, 25, 292–311.

[2] Azuaje, F., Dubitzky, W., Lopes, P., Black, N., & Adamson, K. (2007). Predicting coronary disease risk based on short-term RR interval measurements: A neural network approach. Artificial Intelligence in Medicine, 15, 275–297.

[3] Tsai, D. Y., & Watanabe, S. (1998). Method optimization of fuzzy reasoning by genetic algorithms and its application to discrimination of myocardial heart disease. Proceedings of IEEE Nuclear Science Symposium and Medical Imaging Conference ,1756–1761.http://www.ifcc.org/ejifcc/vol14no2/140206200308n.htm

[4] Fuster, V., Alexander, R. W., & O'Rourke, R. A. (2001). Hurst's the heart. USA: McGraw Hill Professional.

[5] S. Moein, S. A. Monadjemi, & P. Moallem (2008) "A Novel Fuzzy-Neural Based Medical Diagnosis System", World Academy of Science, Engineering and Technology 37 ,2008.

[6] Tkacz, E. J., & Kostka, P. (2000). An application of wavelet neural network for the Annual International Conference on IEEE Engineering in Medicine and Biology ,1391–1393. classification patients with coronary artery disease based on HRV analysis. Proceedings of

[7] Reategui, E. B., Campbell, J. A., & Leao, B. F. (1997). Combining a neural network with case-based reasoning in a diagnostic system. Artificial Intelligence in Medicine, 9, 5–27.

[8] D. W. Opitz and J. W. Shavlik, "Actively searching for an efective neuralnetwork ensemble," Connection Science, vol.8, no.3-4, pp.337-353, 1996.

[9] M. P. Perrone and L. N. Cooper, "When networks disagree: ensemblemethod for neural networks," in Artificial Neural Networks for Speechand Vision, R. J. Mammone, Ed. New York: Chapman & Hall, 1993.pp.126-142.

[10] Yang Guo, Guohuo BAi and Yan Hu. 2012. "Using Bayes Network for Prediction of Type – 2 Diabetes", 7th International Conference for Internet Technology and Secured Transactions(ICITST), London, UK.

[11] Reza Entezarin-Maleki, Arash Rezaei and Behrouz Mimaei-Bidgoli. "Anaesthetic Risk Prediction", International Journal of Recent Trends in Engineering, Vol. 1, No. 2, May.

[12] Boris Milovic and Milan Milovik. 2012. "Prediction and Decision Making in Healthcare Using Data Mining" Kuwait Chapter of Arabian Journal of Business and Management Review, Vol. 1, No. 12, August.

[13] Vidya R. and Dr. Nasira G.M. 2015. "A Novel Medical Support System for the Social Ecology of Cervical Cancer: A Research to Resolve the Challenges in Pap Smear Screening and Prediction at Firm Proportion", Advances in Natural and Applied Sciences, Vol. 9, No. 6, pp. 633-638.

[14] Vidya R. and Dr. Nasira G. M. 2014. "A Theoretical study on cervical cancer Assessment & Need of Computerization in Cervical Cancer Early Detection Program", IJETCCT, ISSN NO-2348-4454, Vol-1, pp. 109-113.

[15] Karpagavalli S., Jamuna K. S. and Vijaya M.S. 2009. "Machine Learning Approach for Preoperative Comparison of Classification methods Based on the type of and Sample Size", Journal of Convergence Information Technology(JCIT).

[16] Vidya R., Jaia Priyankka R. P. and Nirmal Kumar G. 2014. "A system learning of connection with humans by online social networking - A rapport by means of creating usable customer Intelligence from Social media Data", ICSEMR.

[17] Chung-Lang and Chih-Hao Chen. 2009. "Applying Decision Tree And Neural Network to Increase Quality of Dermatological Diagnosis", Expert System with Applications, Vol. 3, pp.4045-4041. [9] Eleni-Maria Theodoraki, Stylianos Katsaragakis and Christos Koukouvinos Christina Parpoula. 2010. "Innovative Data Mining Approaches for Outcome Prediction of Trauma Patients", J.Biomedical Science and Engineering, V

