Use of Modular Neural Network for Heart Disease

Harsh Vazirani, Rahul Kala, Anupam Shukla, Ritu Tiwari
Indian Institute of Information Technology and Management Gwalior, Gwalior, India

Citation: H. Vazirani, R. Kala, A. Shukla, R. Tiwari (2010) Use of Modular Neural Network for Heart Disease, *International Journal of Computer and Communication Technology* 1(2-4), 88-93.

Final Version Available At: http://interscience.in/SpIss_ijcct_accta_2010vol1_nol2/Computing_Paper6.pdf

***Abstract*— Abstract: The medical field is very versatile field and one of the interested research areas for the scientist. It deals with many medical disease problems starting with the diagnosis of the disease, preventing from the disease and treatment for the disease. There are various types of medical disease and accordingly various types of treatment methods. In this paper we mostly concern about the diagnosis of the heart disease. Mainly two types of the diagnosis method are used one is manual and other is automatic diagnosis which consists of diagnosis of disease with the help of intelligent expert system. In this paper the modular neural network is used to diagnosis the heart disease. The attributes are divided and given to the two neural network models Backpropagation Neural Network (BPNN) and Radial Basis Function Neural Network (RBFNN) for training and testing. The two integration techniques are used two integrate the results and provide the final training accuracy and testing accuracy. The modular neural network with probabilistic product method gave an accuracy of 87.02% over training data and 85.88% over testing accuracy and with probabilistic product method gave an accuracy of 89.72% over training data and 84.70% over testing accuracy, which was experimentally determined to be better than monolithic neural networks.**

Keywords- Medical Expert System, Modular Neural Network, Diagnostic System, Back Propagation Neural Network, Radial Basis Function Network, Breast Cancer.

I. INTRODUCTION

The past research in the medical field shows the huge requirement of the expert system for the diagnosis of the disease [1]. One problem is the right treatment to the patient is only provided if the diagnosis of disease is done accurately and on the right time. Another problem is that there are various types of disease such as viral fever, dry cough, and stomach pain etc. for which the treatment is easy and required less time for the diagnosis. The other are more complicated such as cancer and heart disease which takes time to diagnosis. The mortality rates due to these diseases are very high around the world. The approach use in this paper is to develop a decision support system which can train and test the dataset of higher dimensionality in a computational efficient manner.

The artificial neural network is one way to diagnosis the disease the multilayer-perceptron model is one of the basic for of neural network and its network architecture forms by the neurons consists in the input, hidden and output layer. The numbers of neural network models are available from which we use in this paper the BPNN and RBFNN.

The BPNN architecture consists of input, hidden and output layer. In RBFNN the weights according the error at the output layer cause by the difference in the actual output and target output.

In the RBFNN similar kind of architecture as of the BPNN is formed. It also has the three layers.

The ability of the monolithic neural network is less due to problem in handling the large dataset. Due to this we use the modular neural network divide the attributes among the different modules. The modular neural network gives better performance as compare to single neural network model many cases.

The RBF/MLP modular structure [2] is used with three or more networks for microwave design modeling. The artificial neural network training is done by the electromagnetic-neural network. The number of problems in designing the microwave is solved by the modular neural network. But there are limitations of ENN as it use the large amount of data which making it complex and a high dimensionality problem. This problem is solved by the use of modular neural network as it divides the data among the different model to transfer the overall load.

A new method [3] is used named as expert in one thing and good at many for developing modular neural network. The whole training set id divided into subsets which can any of them can select to perform training with different performance.

In [4] they concern about the problem of predicting the real prices of the tomato and green in the US because they shows the complex fluctuation in time. They choose the neural network model to provide the better accurate results. They compare the different neural network models and training algorithms to know which provides the better result and they also compare with the statistical approach. Response integration is done by using several methods to integrate the results.

A new method [5] of response integration is used for the modular neural network by using type-2 fuzzy logic. Face, fingerprint and voice are the three biometric characteristics which are used. The responses from the different modules are integrated using the type-2 fuzzy logic. The working of the modular neural network is divided among the local experts and an integration unit. Each module are further divided into sub module and trained.

In [6] this they propose the robot end-effector using autonomous control. To track the movements of a robot at the remote site is done by the use of computer vision technology. The modular neural network outperforms over the single neural network even in the complex situation of occlusion, clustering or various other motions. The working of the modular neural network is to recognize the robot end-effector in very less processing time.

The modular neural network was applied for the speaker identification [10]. The speaker recognition, which can be classified into identification and verification, is the process of automatically recognizing who is speaking on the basis of individual information included in speech waves.

The approach used in [11] is to improve handwritten digit recognition. In this they used two basic kind of modular networks are used. In the first kind the digits 0, 1, 2, 5, 6, 7 are the digits are provided to the seven expert modular networks. The pair of digits 3-8 and 4-9 respectively is provided for the other modular networks.

II. BACKGROUND

In this section background of the paper is discussed. This section is divided into three parts Back Propagation Neural Network (BPNN), Radial Basis Function Network (RBFN), and Modular Neural Network (MNN).

A. *Backpropagation Neural Network*

Backpropagation Neural Network: It is the one of the more focused area of research. The basic architecture consists of input, hidden and output layers. In this the weights are updated according to the error at the output layer due to the difference in between the actual output and target output. The outputs from the input layer are propagated to the hidden layer which further propagated to the output layer. The results are propagated back from the output layer to the input layer according to which the weights are updated.

B. *Radial Basis Function Network*

RBFNN have similar network architecture as of BPNN is shown in the fig. 1. The radial basis function neural network consists of neuron at the input layer and output layer which shows the input vector and desired output at the output layer. The radial basis function is a type of neural network model that is functionally similar with biological neuron. Biological nerves system adjusts to a narrow range of local area of frequencies. The cells in the ear are locally adjusted to particular frequencies of sound. The feed forward architecture consists by the neural network with non-linear units and an output layer with linear units [9, 13].

The classification and function approximation are the problems that are appropriately handle by the RBFNN.

The centers of the radial basis function, learning method use for input-output mapping are the different criteria of the performance of the neural network.

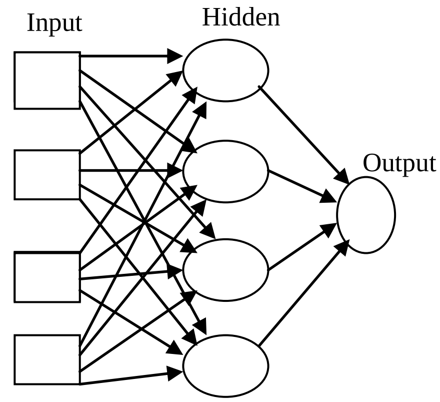

Fig. 1: General Architecture of Radial Basis Function Neural Network

An RBFNN consists of single hidden layer whereas the MLP consists of more than one hidden layer. The RBFNN consist non-linear functions at the hidden layer and linear functions at the output layer. This makes the RBFNN functionally different at the hidden layer and output layer. The selection of number of radial basis functions and there centers are difficult task for this functions.

C. *Modular Neural Network (MNN)*

The overall architecture of proposed system is shown in fig. 3.The analysis of the problem is done to make a solution. The whole problem is big and complex to handle out. The division of the task into number of sub task makes it elementary, simpler and less complex task. There results obtained from different modules are integrated to provide obtain the final results. The working of the neural network models on the different modules are independent that later communicate and integrate to provide the final solution.

The conventional neural network is change into the modular neural network if it is divided into number of modules which works independently and results of them are integrated by the integration unit. The task of integration unit is to integrate the results from the different modules and gives the final output result. Some of the methods use for the integration are winner take all, voting, average, and fuzzy inference system [7].

III. METHODOLOGY

The architecture of working of whole methodology is shown in the fig. 2. The dataset is collected from the uci machine learning repository which is divided into 70% training and 30 % testing dataset. The attributes are divided among the four modules. The attribute 1, 2, 3, 4, 5, 6 and7 represents the age, sex, chest pain type, resting blood pressure, serum cholesterol in mg/dl, fasting blood pressure sugar which is more than 120 mg/dl, and resting electrocardiographic results of patient and attributes 8,9,10,11,12 and 13 represents the maximum heart rate achieved, exercise induced angina, old peak = ST depression induced by exercise relative to rest, the slop of the peak exercise ST segment, number of major vessels (0-3) colored by fluoroscopy, thal: 3 = normal; 6 = fixed defect; 7 = reversible defect[12].

The individual modules are then working upon the attribute set given to them. The module 1 and module 3 are train and test by the BPNN and module 2 and module 4 are train and test by the RBFNN. The attribute division is shown in fig. 4. The integration unit will integrate the results of all the modules to give the final training and testing accuracy.

In this experiment we use the BPNN and RBFNN as the two expert systems which work upon the different modules. The 13 attributes are divided among the four modules. The performance of the individual network may not effectively solve the problem. Hence the dataset is divided two parts of 7 attributes and 6 attributes. The feature space is complex when the whole problem is used as compare to the feature space of the sub-problem which having limited number of attributes which is less complex. All the four modules are work independently and produce the results which are further use by the integration unit. The attribute 1 to 6 is given to the module 1 and module 2. The remaining attributes from 8 to 13 are given to the module 3 and module 4. The two types of method is use for the integration one is probabilistic sum method and other is probabilistic product method. First Probabilistic sum method is use as integration method. The output vectors from all the four modules are use in the equation to obtain the final output. The output vectors from the module 1, module 2, module 3 and module 4 are denoted by a, b, c and d. The final output produce by the integration unit is *O* is shown in equation 1.

$$O = aw_1 + bw_2 + cw_3 + dw_4 \qquad (1)$$

Where all the weights w1, w2, w3, w4 are equal to 0.25. The weights are taken fixed due to the error which comes after the training from each module is similar in nature. The normalization equation is use to find out the final output with probabilistic product method is shown in equation 2.

$$A = anw_1 * bnw_2 * cnw_3 * dnw_4 \qquad (2)$$

If the output O is greater than 0.5 than the heart disease is present and if it is less than 0.5 than the heart disease is absent. This is similar for the output A if it is greater than 0.5 than the heart disease is present and if it is less than 0.5 than the heart disease is absent. This is given by equation.

The probabilistic product is a second method for the integration of results from all the modules. The training and testing results are calculated once again for the all the modules that are further use in the probabilistic product method.

The capabilities of ANN in some cases when used are not satisfactory like in the case of large dataset. The modular neural network is used in such cases which work as a combination of neural networks. The modular neural network has a hierarchical organization comprising multiple neural networks. The combination of estimators may be able to exceed the limitation of a single estimator. The idea also shares conceptual links with “divide and conquer” methodology [7, 8].

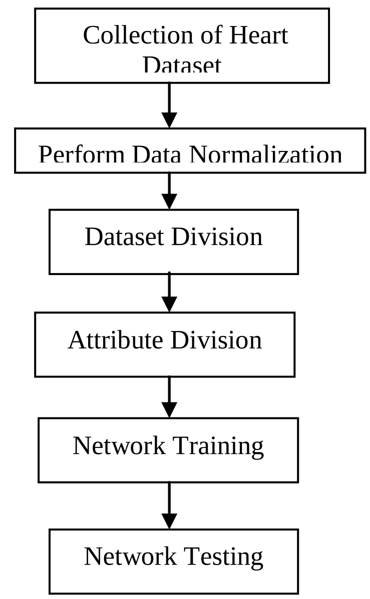

Fig. 2: General Experimental Methodology

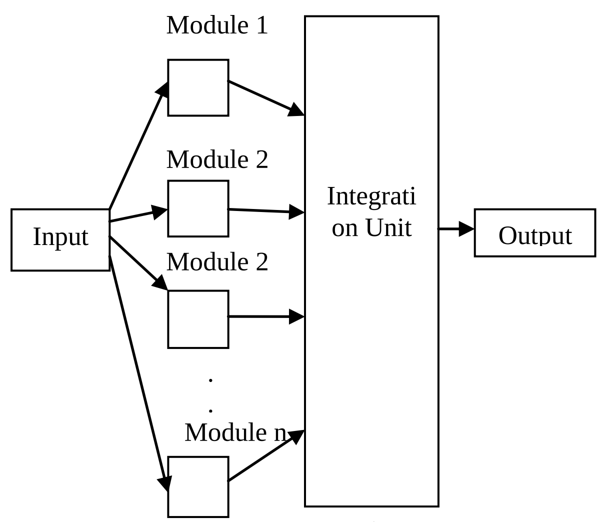

Fig. 3: Overall Architecture of Proposed System

IV. RESULTS

Firstly the dataset [12] is collected from the UCI machine learning repository then the normalization is performed of the database is performed. The attributes are provided to the four modules. The output vectors and the weights are used to calculate the final results. The two neural network models are used for the experiment. The training and testing accuracy of these two models over the whole data is calculated. In the experiment with BPNN first the 15 neurons are use for the whole database. The BPNN is also work upon the module 1 with 20 neurons and module 3 with 25 neurons. The activation functions use for the hidden layer was *tansig* and *purelin*. The training function used was *traingd*. The other parameters were a learning rate of 0.05 and a goal of 10-1. Training was done till 2400 epochs.

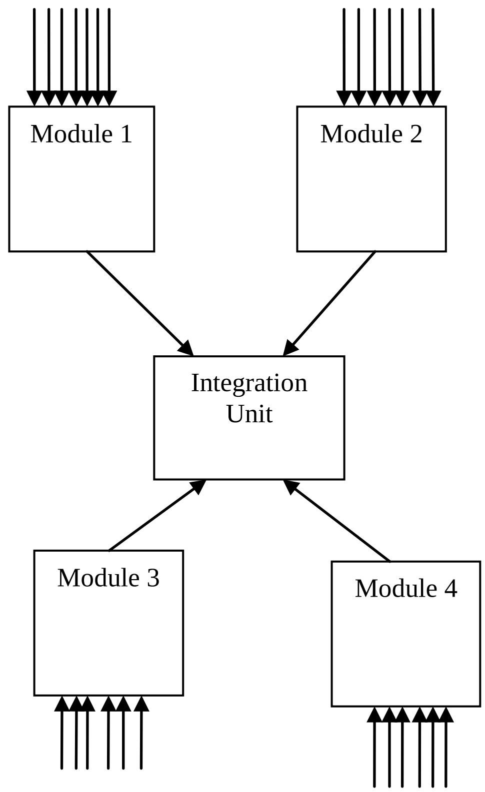

Fig.4: Division of attributes between modules

The RBFNN is another model which is use to work over the module 2, module 3 independently and on the whole data.

The performance of the propose expert system model gives better results with better accuracy with less training time.

The BPNN works upon the module 1 and module3 the training and testing accuracy are found 77.60 % and 76.92 % with elapsed time 9.72 seconds. The module 3 with six attributes is train and test with BPNN and training and testing accuracy are found 84.86 % and 80 % with elapsed time 9.87 seconds. The whole data when train with BPNN gives the training accuracy of 88.10 % and testing accuracy of 82.35 %.

The module 2 is given to RBFNN to train and test and the performance are found 77.83 % and 85.88 % with elapsed time 0.15 seconds and the training and testing accuracy of module 4 with the same neural network model as for the module 2 are 83.78 and 84.70 with the elapsed time .079 seconds. The whole data when train with RBFNN gives the training accuracy of 86.48 % and testing accuracy of 87.05 %.

The performance after the network was trained with modular neural network is 87.02% for the training data and 85.88% for testing data.

The probabilistic sum method is use to integrate the output results from the module 1, module 2, module 3 and module 4. The table 1 shows the summary of result based upon the probabilistic sum method.

Table 1: Experimental Results with probabilistic sum integration

Module	Method	Attributes	Training Accuracy	Testing Accuracy	Time (secs)
1	BPA	1-7	77.60%	76.92%	9.72
2	RBFN	1-7	77.83%	85.88%	0.25
3	BPA	7-13	84.86%	80%	9.87
4	RBFN	7-13	83.78%	84.70%	0.79
-	**MNN**	**1-13**	**87.02%**	**85.88%**	**9.23**
-	BPNN	1-13	88.10%	82.35%	5.58
-	RBFN	1-13	86.48%	87.05%	0.08

The training and testing is done again to calculate the probability of training and testing for the probabilistic product method. In this method the output vectors from all the modules are normalized between -1 and 1 and then the final results is calculated for that the threshold is set. In this case the threshold is set as 0. The training and testing accuracy calculated for the module 1 are 81.08 and 75.29 and for the module 3 are 88.10 % and 80 % with elapsed time 10.28 seconds. In module 2 and module 4 have six attributes. The training and testing accuracy for module 2 calculated are 81.08 % and 75.29 % with elapsed time 0.17 seconds. The module 4 is the last module for which the training and testing accuracy calculated are 83.24 and 84.70 with same elapsed time as in the case of module 2.

The whole data is use for training and testing is done with both the two neural network models. The training accuracy and testing accuracy with BPNN are 90.81 % and 81.17 %.

The two neural network which works on the different modules gives the results which are integrated by the integration unit. The integration of results is done by the probabilistic sum method and probabilistic product method. The integration unit founds the final results.

The attribute division is done to handle the high dimensionality problem. The other benefit of the use mixture of neural network model works upon the different modules independently that lower down the risk involved in working out on the whole problem.

The propose expert system model use the heart dataset that is not a very database as there are less number of attributes and instances. The performance of the propose expert system may be done to get the better results.

Two approaches are use one is the modular neural network that divides the whole system into different modules and other is mixture of expert system that provide working on the problem independently provides the significantly improved results as compare with the monolithic neural network.

Table 2: Experimental Results with probabilistic product integration

Module	Method	Attributes	Training Accuracy	Testing Accuracy	Time (secs)
1	BPA	1-7	81.08%	75.29%	9.25
2	RBFN	1-7	77.29%	82.35%	0.17
3	BPA	7-13	88.10%	80%	10.28
4	RBFN	7-13	83.24%	84.70%	0.17
-	**MNN**	**1-13**	**89.72%**	**84.70%**	**9.24**
-	BPNN	1-13	90.81%	81.17%	9.52
-	RBFN	1-13	91.35%	74.11%	0.12

The training and testing results shows the performance improvement over the single neural network model that is shown by the training and testing accuracy measured with the whole dataset.

The modularity in the network improves the performance and reduces the learning time of the network. The work load of each module is less than as compare with using the monolithic neural network.

The two integration that are use in this paper is work differently as one is based on the probabilistic sum that calculates the sum of the output vectors from the different modules. There are other integration methods which may be use for the integration purpose.

Besides, there is a lot of dependence on the choice and measurement of attributes and the diseases at large. While some methods may look better for the problem of heart disease, it cannot be guaranteed that the results would observe same behavior for other diseases. This again necessitates on the knowledge of both theoretical and practical aspects of the various hybrid methods and an iterative design approach to get a good soft computing system for problem solving.

In this paper we saw the working of modular neural network with two different integration methods for the problem of detection of heart disease.

The database records the presence or absence of diabetes in the Indian female patients. The ultimate aim is to make an expert system that would assist the doctors in diagnosis of the disease.

The overall performance of the modular neural network for the diagnosis of the disease is better than the monolithic neural network.

In a field like bio-medical engineering wrong decisions can greatly harm a patient's health. This imposes a big limitation in the use of these systems for completely autonomous diagnosis of disease. No matter how better our systems are engineered, for a very long time we would require the supervision of doctors for the final verdict. We have a very long way to go before the accuracy can be made 100%, enabling systems to perform autonomously.

The use of the manual system along with the automated diagnosis system may be use for the diagnosis of the disease. This would provide the appropriate results for accurately diagnosis the disease.

In this paper we proposed the use of modular neural network for breast cancer diagnosis. The entire system first divided the attributes into sets of smaller attributes to escape from the curse of dimensionality, which is a known problem in neural computation. Each problem set was solved using two different neural models. Probabilistic sum integration was used to get the final result. The approach was applied over the problem of diagnosis of breast cancer. It was observed that the system so evolved was easily able to solve the problem and gave an accuracy that was much better than the monolithic neural methods. Hence this can be an effective approach for disease diagnosis.

The approach discussed was applied over the problem of breast cancer disease diagnosis. This database is not a very large in terms of the number of attributes and the number of data instances. The experimentation over the other databases with very large number of attributes and number of data instances may be done that may give more efficient results with the use of modular neural networks.

ACKNOWLEDGMENT

This work is sponsored and supported by Indian Institute of Information Technology and Management Gwalior.

REFERENCES

[1] Igor Kononenko, "Machine Learning for medical disease diagnosis: History, state of the art and the perspective", Artificial Intelligence in Medicine, Vol. 23, Issue 1, pp. 89-109, 2001.

[2] Marico G. Passos, Paulo H. da F. Silva and Humberto C.C. Fernandis, "A RBF/MLP Modular Neural Network for Microwave Design Modeling", International Journal of Computer Science and Network Security, Vol. 6, pp. 5A, 2006.

[3] Ken Mcgarry and Stefan Wermter, "Training without data: Knowledge Insertion into RBF Neural Networks", Proceedings of 19th International Joint Conference on Artificial Intelligence, pp. 792-797, 2005.

[4] Patricia Melin, Valente Ochoa, Luis Valenzuela, Gabriela Torres and Daniel Clemente, "Modular Neural Networks with Fuzzy Sugeno Integration Applied to Time Series Prediction", Hybrid Intelligent Systems, Vol. 208/207, pp. 403-413, 2007.

[5] Jerica Urias, Denisse Hidalgo, Patricia Melin, and Oscar Castillo, "A New Method for Respose Integration in Modular Neural Networks using Type-2 Fuzzy Logic for Biometric Systems", Proceedings of International Joint Conference on Neural Networks, Florida, USA, August 12-17, 2007.

[6] Dong-Sun Park, Sok Yoon, YoungBu Kim,"Robot end-effector recognition using modular neural networks for autonomus control", Proceedings of Ineternation Joint conference on Neural Networks, Vol. 3, pp. 2032-2036, 1999.

[7] Vineet R. Khare, Xin Yao, Bernhard Sendhoff, Yaochu, Heiko Wersing, "Co-evolutionary Modular Neural Networks for Automaric Problem Decomposition", Congress on Evolutionary Computation (CEC). IEEE Press, pages 2691-2698, 2005.

[8] J.W.H. Lee, A.K.Y. Lai, Qi-Lun Zheng, "Function-Based and Physics-Based Hybrid Modular Neural Networks for Radio Wave Propagation Modelling", Proceedings of IEEE Antennas Propag Soc AP S INT Symp, 2000.

[9] Bing Yu, Xingshi He, "Training Radial Basis Function Networks with differential evolution", Proceedings of IEEE International Conference on Granular Computing, Atlanta,GA,pp. 369-372, 2006.

[10] Gabriela Martinez, Patricia Melin, and Oscar Castillo, "Optimization of Genetic Algorithms of Modular Neural Networks using Hierarichcal Genetic Algorithm Apllied to Speech Recognition", Prcoceedings of International Jont Conference on Neural Networks, Vol. 3, pp. 1400-1405, 2005.

[11] Claudio A Perez, Patricio A Galdmas, and Carlos A Holzmann, "Improvements On Handwritten Digit Recognition By Cooperation of Modular Neural Networks", IEEE International Conference on Systems Man And Cybernetics, Vol. 5, pp. 4172-4177, 1998.

[12] UCI Machine Leaning Repository, [http://www.ics.uci.edu/~mlearn/MLRepository.html], Available at : http://archive.ics.uci.edu/ml/datasets/Statlog+Heart

[13] N.B. Karayiannis, Gips Randolph,"On the Construction and Training of Reformulated Radial Basis Function Neural Networks", IEEE Transactions on Neural Networks, Vol. 14, pp. 835-836, 2003.