IJARSE
ISSN (O) 2319 - 8354
ISSN (P) 2319 - 8346

Modular Neural Network Approach for Data Classification

[1,]Divya Taneja, [2]Dr. Vivek Srivastava,

M.tech Scholar, Dept. of CSE

Faculty of Engineering, Rama University, Kanpur

ABSTRACT

Classification is a challenging task that has important application in real life and its application are excepted to grow more in future. In this paper, we analyze the effectiveness of Modular Neural Network as a modelling tool for data classification. The MNN classifier outperforms the surveyed nets due to its novel task decomposition and multi-module decision-making techniques. In this paper, we present a MNN architecture for supervised learning. The basic building block of the architecture are multilayer feed forward neural network with back propagation algorithm. MNN is consider as one of the state-of-art system as feature extractors and classifier and are proven to be very efficient in analyzing problem with complex feature space. The aim of this work is achieve by five bench mark problem- Magic Gamma Telescope Data set, Liver Disorder Data set, Balance Scale Data set, Monk's Problem Data set, Yeast Data Set. Experiment describe in this paper show that the architecture is especially useful in solving problems with a large number of input attributes.

Keywords– Classification, Modular neural network, feedforward neural network, back-propagation algorithm.

I. INTRODUCTION

Classification is a task of determining datasets into predefined classes based on the certain kind of their similarities. Classification tasks are an integral part of science, industry, business, and health care system; being such a pervasive technique, its smallest improvement is valuable. Developing more accurate and widely applicable classification models has significant implication in these areas. It is the reason that despite of the numerous classification models [2] available, the research for improving the effectiveness of these models has never stopped [6]. Artificial Neural Network(ANN) is one of the strongest technique used in many disciplines for classification. The ANN technique suffers from drawback such as in transparency in spite of its high prediction power. The choice of Modular Neural Network [1] model for Data Classification, due to their flexibility, adaptive and generalization capability and their easy application in software and hardware devices Specifically, in the field of artificial neural network research, which derives its inspiration from the functioning and structure of the brain, modular design techniques are gaining popularity. The use of modular neural networks for the purpose of regression and classification can be considered as a competitor to conventional monolithic artificial neural networks, but with more advantages. Two of the most important advantages are a close neurobiological basis and greater flexibility in design and implementation. The paper focus on the powerful concept of modularity. It is describing how this concept is deployed in natural neural network on an architectural as well as on functional level.

IJARSE
ISSN (O) 2319 - 8354
ISSN (P) 2319 - 8346

II. PROBLEM STATEMENT

A. Magic Gamma Telescope Data set

The dataset was generated to stimulate registration of high energy gamma particles in a Major Atmospheric Gamma-Ray Imaging Cherenkov (MAGIC) Gamma telescope. The task is to distinguish gamma rays(signal) from hadronic showers (background). There are two classes g-gamma(signal):12332 instances h-hadron (background):6688 with 10 attributes

B. Liver Disorder Data set

Dataset contain the information of blood test which are thought to be sensitive to liver disorder that might arise from excessive alcohol consumption and the number of half print equivalents of alcoholic beverages drunk per for each individual. The task is to select if a given individual suffers from alcoholism. It consists of 345 instances and 6 attributes.

C. Balance Scale Data set

Dataset was generated to model psychological experimental results. Each example is classified as having balance scale tip to the right, tip to the left, or to be balanced. It consists of 625 instance with three class and 4 attributes.

D. Monk's Problem data set

The MONK's Problem was the result of a first international comparison of learning algorithms. The result of comparison is summarized in "The MONK's Problems". There are three monk problem we have chosen one of that.

E. Yeast data set

This data set contains information about a set of Yeast cells. The task is to determine the localization site of each cell among 10 possible alternatives. It consists of 1484 instances and 8 attributes.

III. PROPOSED METHODOLOGY

A. Back-Propagation Neural Network Algorithm

This learning algorithm [7] is applied to multilayer feed forward neural network [10] with different activation function. The training of the BPN [12] is done in three stages-the feed-forward of the input training pattern, the calculation and back-propagation of the error, and updating of weights. The testing of the BPN involves the computation of feed-forward phase only. There can be more than one hidden layer (more beneficial) but one hidden layer is sufficient. Even though the training is very slow, once the network is trained it can produce its output very rapidly.

Algorithm entails three phase-

- training,

During the training phase, the neural network is trained by series of training set in which input features have known classifications. The network adjust itself continuously to minimize the difference between the neural networks output and the known classifications. This is known as 'training cycle'. When the network adjust itself to minimize error, this is called an "epoch". An epoch can occur after every training cycle or after several training cycles to get a net error adjustment of several training set. The number of training cycles necessary to fully train a network may vary from a few hundred to may hundreds-of-thousands.

IJARSE
ISSN (O) 2319 - 8354
ISSN (P) 2319 - 8346

- recall (weight updat) and

After the network has been tentatively tarried, the recall phase can be entered. During the recall phase the same training set with which the network was trained are again presented. However, the network does not adjust itself, but simply generates an output, which the user compares with the desired outcomes. In this way, one can tell how well the network has learned the pattern on which it has been trained.

- testing.

After the training and recall phase have concluded, the testing phase can commence. In the testing phase the network is presented with testing pattern it has not encountered before, though they are of the same origin as the training set, to determine how well the network can interpolate patterns it has not been before. Once a satisfactory neural network has been developed it can be considered for deployment on unknown patterns.

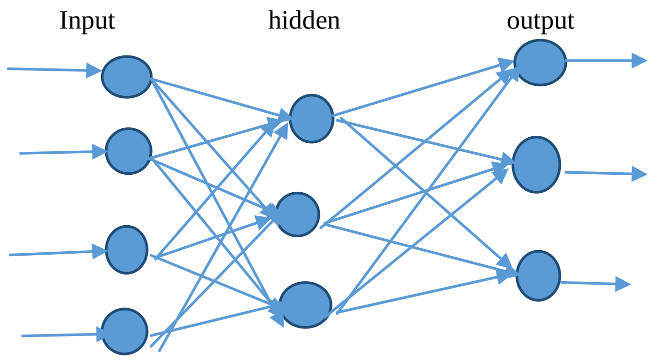
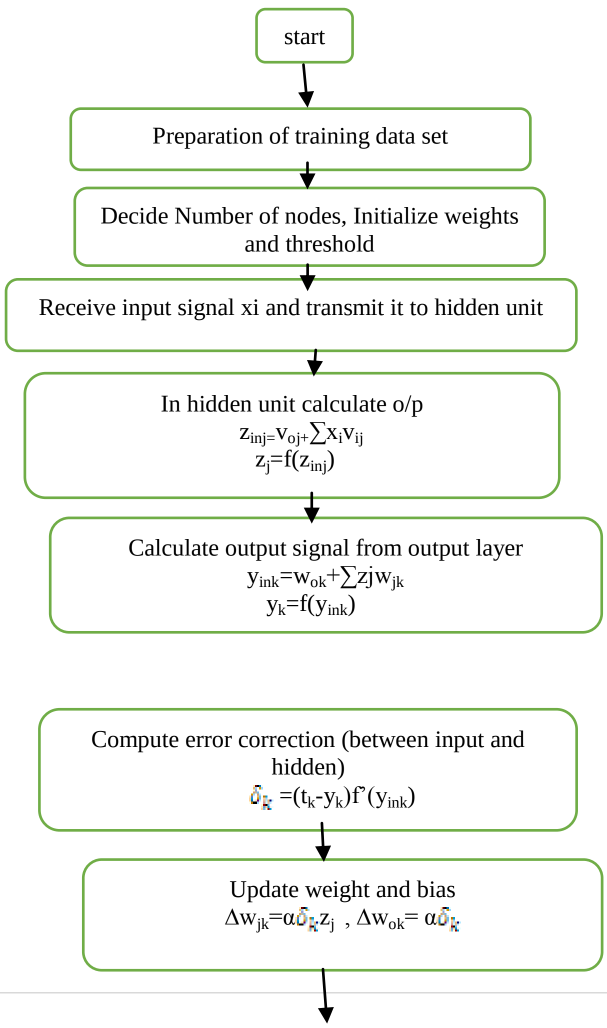


Fig1. Architecture of multilayer feedforward neural network.



IJARSE
ISSN (O) 2319 - 8354
ISSN (P) 2319 - 8346

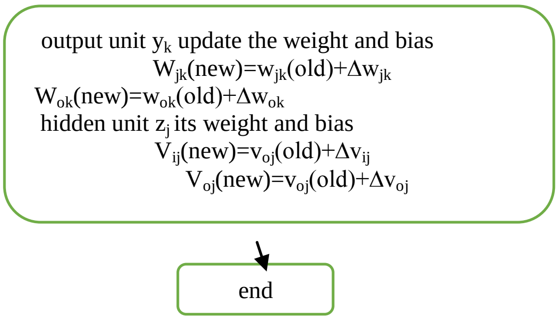


Fig2. Flowchart of Training algorithm

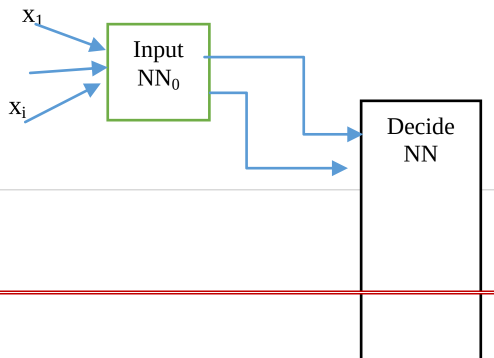
B. Modular Neural Network

One of the major drawbacks of the current neural network [11] generation is the inability to cope with the increase of size/complexity of classification tasks. Modular neural network classifiers attempt to solve this problem through a "divide and conquer" approach. However, the performance of the modular neural network classifiers is sensitive to efficiency of the "task decomposition" technique and the "multi-module decision-making" strategy.

One idea of a modular neural network architecture [4] is to build a bigger network by using modules as building blocks. All modules are neural network. The architecture of a single module is simpler than the sub networks are smaller than a monolithic network. Due to structural modifications the task the modules has to learn is in general easier than the whole task of the network. This makes it easier to train a single module. The modules are independent to certain level which allows the system to work in parallel.

The proposed model can transform a classification problem data set [5] into set of their output class, each of which is solved is solved by single neural network [8]. The result of all neural network are combined to form a final classification decision. As shown in figure 3. Combining several models or using hybrid model has become a common practice in order to overcome the deficiencies of single models and can be an effective way of improving upon their predictive performance, so here we used each neural network to classify one class of data set and other class by other neural network.

Due to the principle of divide-and-conquer used in the proposed architecture, the modular neural network can yield both good performance and significantly faster training. The proposed architecture has been applied to several classification problem and has achieved satisfactory results.



IJARSE
ISSN (O) 2319 - 8354
ISSN (P) 2319 - 8346

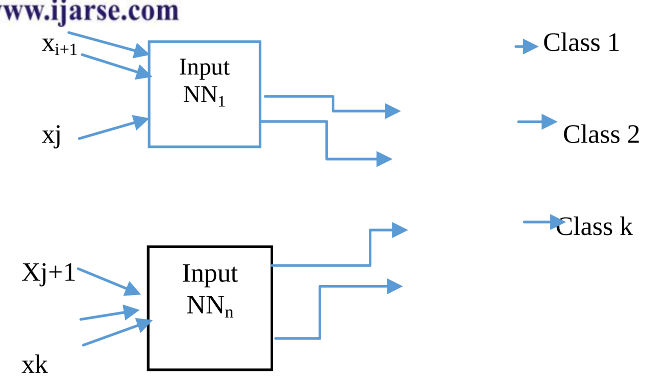


Fig3. Architecture of Modular neural network

IV. EXPERIMENT AND RESULT

We have taken five benchmark problem from UCI Data Repository and perform classification using Back-propagation algorithm [9]. So comparative analysis of five data set is shown in Table 1

TABLE 1: COMPARATIVE ANALYSIS OF DATA SET FOR CLASSIFICATION TASK

Dataset	Instances (Training + Testing)	Attributes	class
Magic	19020(9510+9510)	10	2
Liver	345(173+172)	7	2
Balance	625(213+212)	4	3
Monk	556(124+432)	6	2
Yeast	1484(742+742)	8	10

Input layer receives the training set pattern. BPN propagates the input pattern set from layer to layer until the output layer results are generated [3]. Then, if the output layer results differ from the expected, an error is calculated, and then transmitted backwards through the network to the input layer. In this process values for the weights are adjusted to reduce the error encountered. This mechanism is repeated until a terminating condition is achieved. So for each data set we have different architecture and in result they have different iteration to achieve termination condition.

A. Magic Gamma Telescope Data set

In classification of data set neural network architecture is 10-h-2 i.e. 10 input neuron, one hidden layer with h (5,6,7) neuron that vary to observe change in M.S.E and 2 neurons in output layer.

IJARSE
ISSN (O) 2319 - 8354
ISSN (P) 2319 - 8346

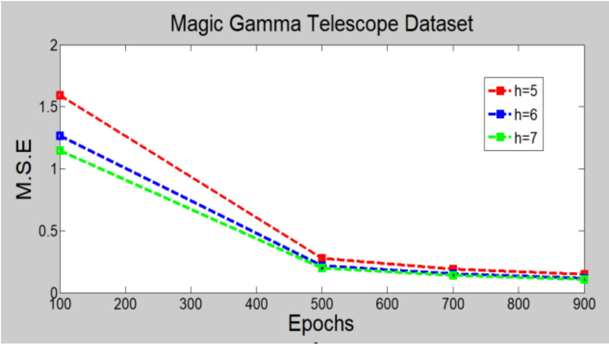


Fig 4. For 900 Epochs vs. M.S.E

B. Liver Disorder Data set

In classification of data set neural network architecture is 7-h-2 i.e. 7 input neuron, one hidden layer with h (10,15,20) neuron that vary to observe change in M.S.E and 2 neurons in output layer.

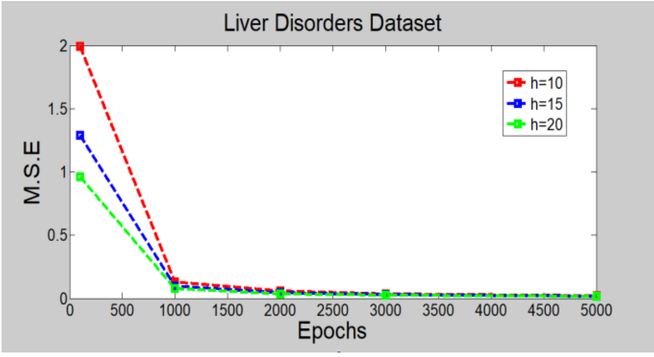


Fig 5. For 5000 Epochs vs. M.S.E

C. Balance Scale Data set

In classification of data set neural network architecture is 4-h-3 i.e. 4 input neuron, one hidden layer with h (15,20,25) neuron that vary to observe change in M.S.E and 3 neurons in output layer.

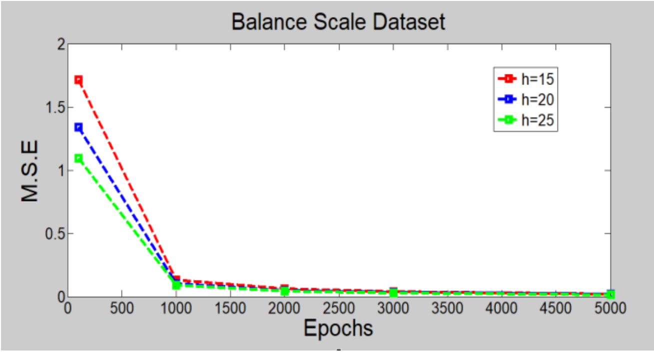


Fig 6. For 5000 Epochs vs. M.S.E

D. Monk' Problem Data set

IJARSE
ISSN (O) 2319 - 8354
ISSN (P) 2319 - 8346

In classification of data set neural network architecture is 6-h-2 i.e. 6 input neuron, one hidden layer with h (10,15,20) neuron that vary to observe change in M.S.E and 2 neurons in output layer.

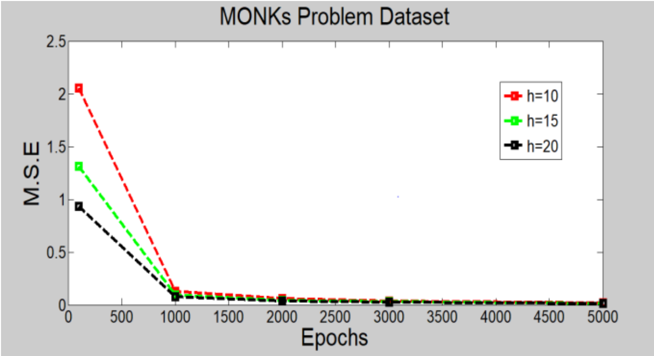


Fig 7. For 5000 Epochs vs. M.S.E.

E. Yeast Data set

In classification of data set neural network architecture is 10-h-2 i.e. 10 input neuron, one hidden layer with h (10,11,14,15,19,20) neuron that vary to observe change in M.S.E and 2 neurons in output layer.

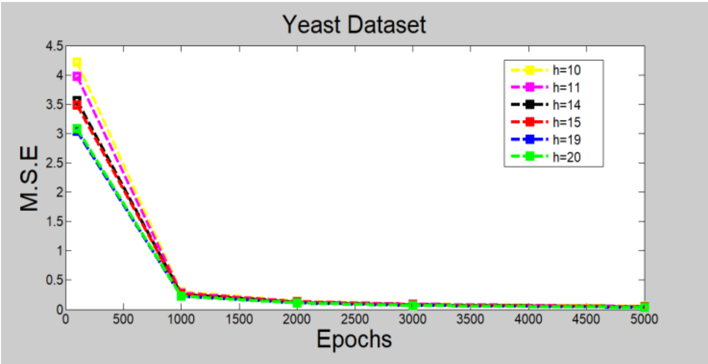


Fig 7. For 5000 Epochs VS M.S.E

Classification accuracy

It shows how well classifier correctly identify the actual class. It can be calculated either as an overall accuracy or as a class specific accuracy

Overall classification accuracy= (C/Ct) *100%

Class specific classification accuracy= (Cn/Cm) *100%

Where C is the number of correct classification in the whole set Ct of classes, and Cn is the number of correctly classified class of class Cm.

TABLE 2: TESTING PERFORMANCE

IJARSE
ISSN (O) 2319 - 8354
ISSN (P) 2319 - 8346

MAGIC	LIVER	BALANCE	MONK	YEAST
99.8%	100%	100%	100%	85.7%

Testing accuracy of each dataset observe at theta 0.1 on different hidden layer. Classification performance depend upon characteristic of dataset to be classified. So we observe different testing accuracy on each dataset.

V. CONCLUSION

This paper aimed to evaluate the artificial modular neural network in predicting classes of different dataset. The feedforward backpropagation neural network with supervised learning is proposed to classification. The reliability of the proposed neural network method depends on data collected that's why we have chosen different domain of dataset. Backpropagation learning algorithm is used to train the feedforward neural network to perform a given task of classification using intelligent machine learning. We achieved relatively better accuracy as compared to conventional neural network. It is shown that for different real world data sets the training is much easier and faster with a modular architecture. Due to the independence of the modules in the input layer parallel training is readily feasible.

REFERENCES

[1] P.Poirazi, C. Neocleous, C.S. Pattichis, C.N. Schizas," Classification capacity of Modular Neural Network implementing neutrally inspired architecture and training rule", IEEE Transaction on Neural Network, vol. 15, pp. 597-612,2004.

[2] Ragasekar Venkatesan, Meng Joo Er.," Multi label Classification method based on extreme learning machines", 13th International Conference on Control Automation Robotics and Vision, pp. 619-624, 2014.

[3] Migel D. Tissera, Mark D. McDonell, "Modular Expansion of hidden layer in single layer feed-forward Neural Network", International Joint Conference on Neural Network, pp. 2939-2945, 2016.

[4] Diogo S. Severo, Everson Verissimo, George D.C. Cavalcanti, Tsang Ing Ren, "Modular Neural Network architecture that selects a different set of feature per module" International Joint Conference on Neural Network, pp. 1370-1374, 2014.

[5] Yogesh Kumar Meena, K.V. arya, Rahul Kala," Classification using reduandant mapping in modular neural network", Second World Congress on nature and biologically inspired Computing, vol.10, pp.554-559,2010.

[6] Josep Garner, Albert Mestres, Eduard Alarcon, Albert Caballor," Machine learning based network modelling: Artificial Neural Network model vs a theoretical inspired model", Ninth International Conference on Vbiquitous ad Future network, pp.522-524,2017.

[7] Swapna C, R.S. Shaji," A survey on evolutionary Machine Learning Algorithm for multi-dimensional data classification", International Conference on Control, Instrumental, Communication and Computational Technology, pp.781-785,2015.

IJARSE
ISSN (O) 2319 - 8354
ISSN (P) 2319 - 8346

[8] Timothy Bender, V. Scott Gordon, Michael Daniels," Partitioning Strategies for modular neural networks", International joint conference on neural network, pp. 296-301, 2009.

[9] Zhong-Qui Zhao, "An evolutionary modular neural network for unbalance pattern classification", IEEE Congress on evolutionary Computation, 2007, pp. 1662-1669.

[10] Murillo Guimaraes Carnerio, Liang Zhao, "Organizational Data Classification Based on the importance concept of complex networks", IEEE Transaction on Neural Network and Learning System, pp.1-13, 2017.

[11] Guanglei Zhang, Lei Che, Yongsheng Ding," A multi-label classification model using convolutional neural networks", 29th Chinese Control and Decision Conference, pp.2151-2156, 2017.

[12] Javier A. Cruz-Lopez, Vincent Boyer, Didier El-Baz," Training many neural network in parallel via Back-Propagation", IEEE International Parallel and Distributed Processing Symposium Workshops, pp. 501-509, 2017.

[13] url -http://archive.ics.uci.edu/ml/datasets