Effects of Hidden Layers on the Efficiency of Neural networks

Department of Computer Science, National University of Computer and Emerging Sciences

Muhammad Uzair, Noreen Jamil
Department of Computing
National University of Computer and Emerging Sciences
Islamabad, Pakistan
i191221@nu.edu.pk, noreen.jamil@nu.edu.pk

2020 IEEE 23rd International Multitopic Conference (INMIC) | 978-1-7281-9893-4/20/$31.00 ©2020 IEEE | DOI: 10.1109/INMIC50486.2020.9318195

***Abstract*— Hidden layers play a vital role in the performance of Neural network especially in the case of complex problems where the accuracy and the time complexity are the main constraints. The process of deciding the number of hidden layers and number of neurons in each hidden layer is still confusing. In this article, we reviewed different impacts of Hidden layers on the network which provides an overview of using three numbers of hidden layers that were found to be optimal in terms of reducing the time complexity and getting the qualified accuracy. The techniques implementing less than three number of hidden layers mostly had a loss in accuracy while the architecture implementing more than three numbers of hidden layers were found not to be optimal in terms of time complexity. Usually implementing three numbers of hidden layers give the optimal performance in terms of time complexity and accuracy. We had a survey on recent work about the Neural network based on the Empirical observations, in which if the number of hidden layers is reduced it has a direct impact on the accuracy of the network as with the complex problem having less number of hidden layers it might be possible that network will not be trained properly. On the other hand when the number of hidden layers cross the optimal number of hidden layers (three layers), time complexity increases in orders of magnitude as compared to the accuracy gain.**

***Index Terms*—Neural network, Hidden layers, Neurons, Accuracy, Time complexity**

I. INTRODUCTION

Neural network is one of the vast domains of 21st century on which a lot of research has been conducted. Human decision-making powers are programmed into the networks so that they will be able to make decision and predict the results based on the trainings. Neural networks [1] are information processing paradigms based on the human beings' biological neural systems. Also, a lot of research and development has been made in the field of Artificial Intelligence with the help of Neural network. Basically, in Neural networks Artificial Intelligence plays a great role as it tries to simulate some major properties of human biological Neural networks [1].

Neural networks have been successfully used for image recognition and speech identification [3]. On the other hand, these are very suitable for pattern recognition and many other real-world problems like signal processing. Appropriate results in pattern recognition can be generated in Forecasting, classification and data analysis. Data Pre-processing [3] Input Layer is a layer in the Neural network which is used to communicate with the outer environment. Once an input is introduced to the input layer of the network, output layer starts producing the resultant patterns. It also defines the conditions on which the network training depends [1, 4, 7].

Output Layer is also connected with the outer environment and represents the output of the input given to the network to the environment. Neurons in the output layers have direct relation with the complexity of the problem or the work given to the network [1].

Hidden layer is basically an intermediate layer between the input and the output layer of the network. Number of hidden layers within the network changes from problem to problem. Some problems need single hidden layer, some need multiples hidden layers. Hidden layers are the collection of small neurons which transfer the data and the training layer to layers [1, 6].

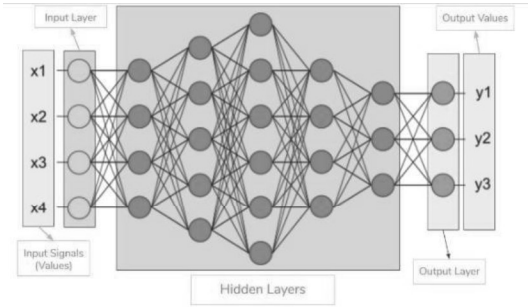


Fig 1: Neural network with Five Hidden layers

Input layer interact with the outer environment to get some data and transfer it to hidden layers to be processed as shown in Fig 1. After processing from the hidden layers this data will show to outer layers which shows the results to the outer environment on the basis of data coming from hidden layers [6].

Selection of hidden layers is very difficult task as in some cases due to the number of hidden layers a condition known as

overfitting and underfitting occurs, which effect the efficiency and time complexity of the network very badly [3, 7, 4].

Overfitting condition occurs when the number of hidden layers become very large as compared to the complexity of the problem due to which a process of overtraining of the network starts and badly effects the time complexity of the network and it mostly arises when network matches the data so closely that it loses its generalization ability over the data to be tested [1, 7].

Underfitting condition occurs when the number of hidden layers in the network becomes less than complexity of the problem as network barely handles such problems [12, 16]. This is also known as undertraining. It effects the efficiency of the network very negatively. In such case the time complexity of the network become very low and produces inefficient results [1, 7].

Basically, the sole purpose of this study is to check the number of hidden layers within the network and effect of these hidden layers on the network.

II. Literature review

A. Related Work

S. Karsoliya [1] proposed a Neural network having three hidden layers. In this approach, use of fourth hidden layers just increase the time complexity but it does not effect on the accuracy. G. Panchal, A. Ganatra, Y. Kosta and D. Panchal [2] discuss about the Neural networks and hidden layers within the network. They concluded that satisfactory training time of the network can be achieved by using just one or maximum two hidden layers. F. Panchal and M. Panchal [3] used Try and Error method by using least number of neurons and hidden layers in BNN. They used a terminology "Epoch" and suggested that the time complexity will increase if the number of epochs increased. Even the results from large dataset were matched with the other researches [5] which verified the methodology and the results produced from it. S. Asthana, A. Pandit and A. Bahrdwaj [4] used approach in which network uses five hidden layers while implementing BPNN for hand written numeral image recognition. They lack the satisfaction of time constraint as network having five hidden layers need a lot of time to be trained. I. Shafi, I. Shah, A. Jameel and F. Kashif [5] introduced Neural network as being trained for giving different blurry spectrogram as input and to execute it as they used single and multiple hidden layers. They proposed to use only one hidden layer for this type of problem having max 40 number of neurons within the hidden layer. S. Xua and L. Chen [6] discuss FNN algorithm which simulates and predict the total taxation revenue for Australia. For their approach they first took this large data set to the network having three hidden layers and then they applied on five hidden layers from which they got an error rate of 11.8 and 5.53 respectively. T. Gupta and K. Raza [7] suggested two hidden layers in FNN for the optimal solution with high rate of accuracy but their methodology lacks the accuracy while applying on the database of MNIST. To optimize Deep Feed Forward Neural Network Architecture, Tabu Search and Gradient decent with momentum back propagation will be applied [7]. G. Huang [8] suggested FNN algorithm with two hidden layers and hidden neurons that can learn any type of sample with small number of errors. He used first hidden layer of maximum size and second hidden layer of narrow size and determined their connection analytically whereas first hidden layer's connection was prefixed with input layer [8]. J. De Villiers, E. Barnad [10] used three and four hidden layers which are not only feasible for only small and linear type problems but these number of layers are also enough to handle large and very complex problems with larger number of data set. S. Seifollahi, J. Yearwood and B. Ofoghi [11] talks about ELM and suggests a binary classifier working on the single Hidden layer by using Radial basis function and sigmoid function in the hidden layer. They also suggest a new weight called "Influence weight" [11].

B. Comparison Study

The common parameters that are available in most of the survey papers are "Time Complexity" and "Accuracy". As these are the most important constraints for neural network from learning time to result production in case of underfitting or overfitting conditions.

1. Time Complexity

Basically, time complexity means that how much time a network takes to learn the whole problem and produces the required results in a required time span [17, 19]. It varies from problem to problem and network to network. Normally a network having large number of hidden layers shows a very large time complexity except in case of overfitting and networks having a smaller number of hidden layers show a very satisfactory time constraints except in under fitting condition.

2. Accuracy

It shows how much the results of the networks are accurate or near to real results. Networks having large number of hidden layers normally show high accuracy even for the large and complex problems. High accuracy can be achieved only if whole problem can be understood by the network without having any overfitting and underfitting conditions [15, 14, 20]. To gain good result in Neural network very large number of hidden layers and neurons are required [14, 21].
Table 1 shows the comparison between accuracy and time complexity on the basis of different hidden layers and neurons.

TABLE 1

Comparison between accuracy and time complexity of different Neural Networks having different number of hidden layers

Sr.#	Study	Number of hidden layers and neurons	Accuracy	Time Complexity
1	"Approximating Number of Hidden layer Neurons in Multiple hidden Layer BPNN Architecture" **Karsoliya,S.[1]**	Three or less than three hidden layers	Approximately more than 90% accuracy will be achieved	Best time complexity in comparison of accuracy and problem will be gained
2	"Behavior Analysis of Multilayer Perceptron with Multiple Hidden neurons and Hidden Layers" **Panchal, G., Ganatra, A., Kosta, Y.,Panchal, D.[2]**	One or maximum two hidden layers	Reasonable accuracy achieved	Less time complexity as just using one or maximum two hidden layers
3	"Review on Methods of Selecting Number of Hidden Nodes in Artificial Neural network" **Panchal, F., Panchal, M. [3]**	Minimum number of hidden layers will be used according to the problem and requirement i.e. One or two for the small problems and three to five for the complex problems	Very good accuracy, sometimes gives 100 percent accuracy as it works according to the problem	Very good and optimal time complexity as it goes with minimum number of hidden layers for the problems due to which network needs very less time to be trained
4	"Analysis of Multiple Hidden layers vs. Accuracy in Performance using Back Propagation Neural network" **Asthana, S., Pandit,A., Bahrdwaj, A.[4]**	Five hidden layers	100% accuracy even in large and complex problems	Not very good time complexity as in case of five hidden layers, network needs a lot of time to be trained
5	" Impact of Varying Neurons and hidden layers in Neural network Architecture for a time Frequency Application" **Shafi, I., Shah, I., Jameel, A., Kashif, F. [5]**	One or maximum two hidden layers with larger number of neurons within the hidden layer	Very good Accuracy obtained as large number of neurons trained the network as good as large number of hidden layers	As compared to hidden layers by increasing number of neurons within a hidden layer accurate time complexity will be gained
6	"A novel approach for determining the optimal number of hidden layer neurons for FNN's and its application in data mining" **Xu, S., Chen, L. [6]**	Five hidden layers	More than 96 % accuracy or can say less than 4% error rate in the results	Inaccurate time complexity will be gained from the approach
7	"Optimizing Deep Neural Network Architecture: A Tabu search Based Approach" **Gupta, T., Raza, K. [7]**	Two hidden layers	Very good accuracy gain in case of sample problem and linear problem but in complex and large problems this shows very bad results	Good time complexity will be gained as just two hidden layers were used to train the network
8	"Learning capability and Storage Capacity of Two-Hidden-Layer Feed Forward Networks" **Huang, G. [8]**	Two hidden layers	Great accuracy with the sample and small data. In case of complex problem under fitting will occur	Best time complexity will be gained if the problem is small and linear
9	"Back propagation neural nets with one and two hidden layers" **Villiers, J., Barnard, E. [10]**	Three hidden layers	Best accuracy will be gained using this approach as after third hidden layer the growth in accuracy shows no big fold except producing many other problems	Best time complexity will be gained with the three hidden layers as more than three hidden layers just increase the time complexity in a very bad manner
10	"Novel waiting in single hidden layer feed forward Neural networks for data classification" **Seifollahi, S., Yearwood, J., Ofoghi, B. [12]**	Single hidden layer with large number of neurons	In simple and linear problems, accuracy shows satisfaction but not in complex problems	Good time complexity will be gained as it works on the number of neurons

TABLE 2
The major advantages and disadvantages of the approaches used in different papers with the outcome section which hopefully seems to be very helpful as it shows the approaches of many researches

Study	Number of hidden layers and neurons	Outcomes	Published Year
"Approximating Number of Hidden layer Neurons in Multiple hidden Layer BPNN Architecture" **Karsoliya,S.[1]**	Three or less than three hidden layers	It will handle any type of problem complex or large –data with best accuracy and reasonable time complexity. If the criteria is to gain 100 percent accuracy then it will not work in any case as it works with the 90 percent accuracy	2012
"Behavior Analysis of Multilayer Perceptron with Multiple Hidden neurons and Hidden Layers" **Panchal, G., Ganatra, A., Kosta, Y.,Panchal, D.[2]**	One or maximum two hidden layers	This methodology works very well with sample data or small giving very good accuracy in a very small time. But are not able to work on complex and large data set as by using this a condition of under fitting occurs which disables to produce satisfactory accuracy	2011
"Review on Methods of Selecting Number of Hidden Nodes in Artificial Neural network" **Panchal, F., Panchal, M. [3]**	Minimum number of hidden layers will be used according to the problem and requirement i-e One or two for the small problems and three to five for the complex problems	Approach uses least number of hidden layers at first and then increases it as per need due to which very a smaller number of tries are required to find the minimum number of hidden layers for small problems. A lot many tries are needed to find least number of hidden layers for the complex and larger data-sets	2014
"Analysis of Multiple Hidden layers vs. Accuracy in Performance using Back Propagation Neural network" **Asthana, S., Pandit,A., Bahrdwaj, A.[4]**	Five hidden layers	This approach works very well for the problems where criteria is totally accuracy as it gives approximate 100% accuracy rate. Where the criteria is not accuracy, this approach fails as by using five hidden layers only a good range of accuracy is gained leaving many problems for the network i.e. time complexity	2017
" Impact of Varying Neurons and hidden layers in Neural network Architecture for a time Frequency Application" **Shafi, I., Shah, I., Jameel, A., Kashif, F. [5]**	One or maximum two hidden layers with larger number of neurons within the hidden layer	This methodology performs very well as neurons transfer data much faster than the actual hidden layer. Condition of overfitting starts occurring when very large number of neurons are used within a network which effects the time complexity very badly	2007
"A novel approach for determining the optimal number of hidden layer neurons for FNN's and its application in data mining" **Xu, S., Chen, L. [6]**	Five hidden layers	This approach works with the problems having very large dataset but the problems having main criteria of time complexity do not work well with this approach as this approach does very bad with the time constraint	2008
"Optimizing Deep Neural Network Architecture: A Tabu search Based Approach" **Gupta, T., Raza, K. [7]**	Two hidden layers	In case of sample data and small data set it works well and produces optimal results with a very low cost but in large data set like MNIST it does not work well and shows many problems in the results	n/a
"Learning capability and Storage Capacity of Two-Hidden-Layer Feed Forward Networks" **Huang, G. [8]**	Two hidden layers	The approach is used with first hidden layer being larger than the second one. Lots of information transfers easily from neuron to neurons which help the network to be trained in a very short time. This approach does not fulfil the requirement of network facing complex problems while training large dataset under fitting occurs and network shows unsatisfactory results.	2013
"Back propagation neural nets with one and two hidden layers" **Villiers, J., Barnard, E. [10]**	Three hidden layers	This proves that the use of three hidden layers are enough as fourth layer just increases the time complexity. Fifth hidden layer might show some different results but no comparison is made between third and fifth hidden layer	1992
"Novel waiting in single hidden layer feed forward Neural networks for data classification" **Seifollahi, S., Yearwood, J., Ofoghi, B. [12]**	Single hidden layer with large number of neurons	This approach works very well as number of neurons are increased while hidden layers remain same. Over fitting occur if the number of neurons become as large as the network is not able to train itself properly	n/a

It can be seen in table 1 and 2 respectively that if complex and large problems are introduced in the networks having a smaller number of hidden layers then no such satisfactory results will be gathered except a reasonable time complexity. So, if we want to get satisfactory results, we have to increase the number of hidden layers until a satisfactory accuracy has been achieved [1].

C. GRAPHICAL REPRESENTATION

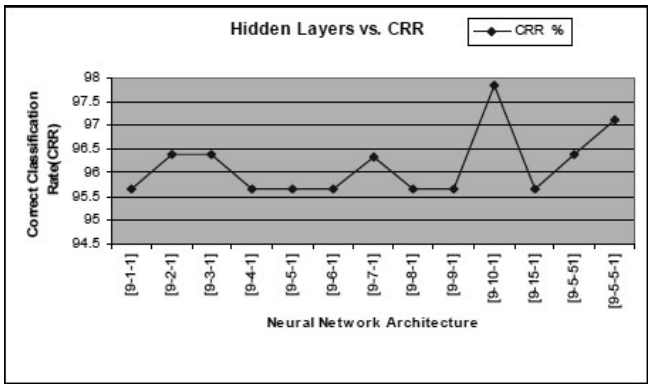
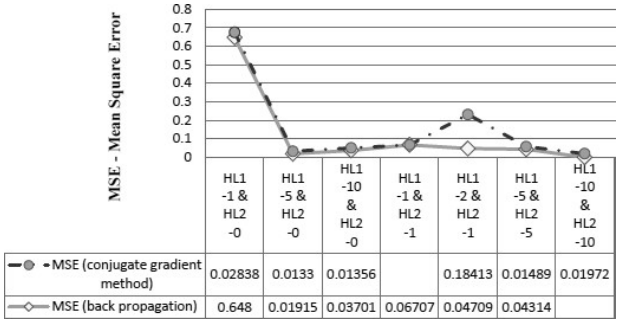


Fig 2[2]: Relationship Between CRR and Hidden layers

From the above fig 2 [2] it is shown that as soon as number of hidden neurons and hidden layers increase, Neural network gives better performance.



	HL1 -1 & HL2 -0	HL1 -5 & HL2 -0	HL1 -10 & HL2 -0	HL1 -1 & HL2 -1	HL1 -2 & HL2 -1	HL1 -5 & HL2 -5	HL1 -10 & HL2 -10
MSE (conjugate gradient method)	0.02838	0.0133	0.01356		0.18413	0.01489	0.01972
MSE (back propagation)	0.648	0.01915	0.03701	0.06707	0.04709	0.04314	

Fig 3[3]: Relationship of Number of Hidden layer and MSE

Fig 3 [3] shows that MSE error of Back propagation method becomes steady after a particular point but in conjugate gradient method the MSE is more fluctuating than back propagation [3]. It clearly shows that if satisfied numbers of hidden layers are used, we get better result with a very less time complexity. On the other hand, if we increase the number of hidden layers, suitable accuracy can be obtained up to great extent, but Neural network becomes complex than the previous methods [3], [5], [1].

Fig. 4 [5] represents the comparisons between number of neurons and the error coverage as the number of neurons within the hidden layer increases, the error coverage of the result comes towards zero. Even at the stage where 40 neurons are used, the network will be able to produce optimal results [5].

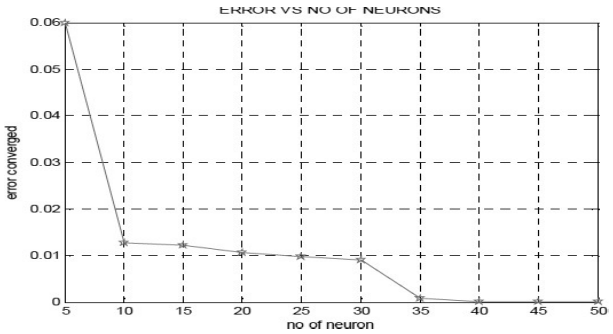


Fig 4[4]: Relationship between Error and number of Neurons

III. METHODOLOGY

These studies were developed on the basis of articles published in "A+" ranked Journals and conferences, which were further examined in the context of the impact factor and their importance in this field. More attention has been paid to recent work. For this study we surveyed various papers and searched the problem related to the topic in detail. Fig 5 shows the flow of methodology followed in this paper.

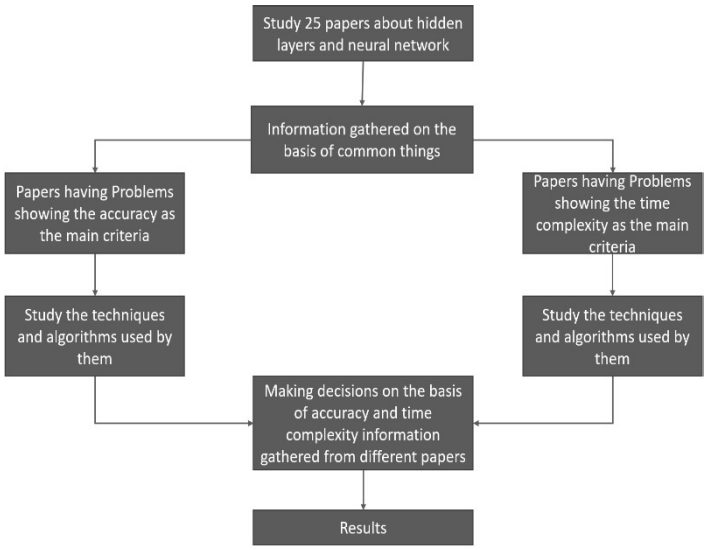


Fig 5: Flow chart for the methodology used in this paper

IV. Conclusion

It is difficult to determine a good network topology only from number of inputs and outputs. Fact is that suitable number of hidden layers can reduce training time with high accuracy. Many approaches are defined to standardize the number of hidden layers that are required for Neural network, but the approximation depends on type of database each time [1, 7, 10].

We conclude that the training process of Neural networks slows down if the large number of hidden layers are used. So, if the criteria of the problem is to get the better accuracy, then large number of hidden layers is the most suitable solution but if the time complexity is the major factor of an application then large number of hidden layers will not work in these types of application. Also, unnecessary increment in the neurons or layer will led to overfitting problem. So it is quiet essential that before designing the Neural network, training database samples must be analyzed so that approximation of number of neurons and hidden layers can be guessed properly [1].

V. FUTURE WORK

Finding the real number of hidden layers is still a very difficult task. Many researchers are still working on this to determine the exact and optimal number of hidden layers. In the future, it would be good to expand the study to include large applications containing training sets of more than 5,000 I / O pairs. Below are some techniques which helps in extending the work related to Hidden layers.

A. Abnormal behavior of the fourth and onward hidden layers

Normally with the increment of any hidden layers accuracy of the network increase with the time complexity of the network but when the fourth hidden layer is added to the network the accuracy factor doesn't show any noticeable change, showing the same fold on the time constraint .Same with the fifth, sixth and the onwards hidden layers.

B. Number of neurons in the initial hidden layer

If the optimize number of neurons will be found for the initial hidden layers then it might be possible that complex problems will also be solved using maximum three hidden layers, this factor is also under observation and many researchers are working on it.

C. No hidden layer network

For small problems, some networks are used which had no hidden layers and gives more or less optimal results. If some more research is done in this area then it might be possible that we will be able to find a network for sample and small data set which produces optimal results without the use of hidden layers .As in those cases the neurons in the input and the output layers will be enough to execute the data.

VI. References

[1] S. Karsoliya," Approximating Number of Hidden layer Neurons in Multiple hidden Layer BPNN Architecture", International Journal of Engineering Trends and Technology Volume 3, issue 6, 2012.

[2] G. Panchal, A. Ganatra, Y. Kosta and D. Panchal, "Behavior Analysis of Multilayer Perceptron with Multiple Hidden neurons and Hidden Layers", International Journal of Computer Theory and Engineering Volume 3, No. 2, 2011.

[3] F. Panchal and M. Panchal, "Review on Methods of Selecting Number of Hidden Nodes in Artificial Neural network", IJCSMC, Vol 3 issue 11, page 455—464, 2014.

[4] S. Asthana, A. Pandit and A. Bahrdwaj "Analysis of Multiple Hidden layers vs. Accuracy in Performance using Back Propagation Neural network", Indian Journal of Science and Technology, Vol. 10 (4), 2017.

[5] I. Shafi, I. Shah, A. Jameel and F. Kashif, "Impact of Varying Neurons and hidden layers in Neural network Architecture for a time Frequency Application", IEEE, 2007.

[6] S. Xua and L. Chen," A novel approach for determining the optimal number of hidden layer neurons for FNN's and its application in data mining". 5th (ICITA), 2008.

[7] T. Gupta and K. Raza, "Optimizing Deep Neural network Architecture: A Tabu search Based Approach", Department of Computer Science, Jamia Milia Islamia, New Delhi, 110025

[8] G. Huang, "Learning capability and Storage Capacity of Two-Hidden-Layer Feed Forward Networks", IEEE, Vol-14, 2, 2013.

[9] J. De Villiers, E. Barnad, "Back propagation neural nets with one and two hidden layers", IEEE, Vol 4, 1, 1992.

[10] S. Seifollahi, J. Yearwood and B. Ofoghi," Novel waiting in single hidden layer feed forward Neural networks for data classification", Computer and Mathematics with applications, vol 64, 128-136.

[11] B.Shah and B. Trivedi, "Artificial Neural network based Intrusion Detection System", International Journal of Computer Applications Vol 39, 2012

[12] Q. Zhang and S. Chun Zhu, "Visual Interpretability for Deep learning", arXiv:1802.00614 V2, 2018

[13] K. Sheela and S. Deepa, "Review on Methods to Fix Number of Hidden Neurons in Neural networks", Hindawi Publishing Corporation Mathematical Problems in Engineering, 2013

[14] S. Awan, M. Riaz and A. Khan, "Prediction of Heart Disease Using Artificial Neural network", VFAST Transactions on Software Enginerring, Vol. 13, 2018.

[15] G. Huang, "Learning Capability and Storage Capacity of Two-Hidden-Layer Feedforward Networks", IEEE Transactions of Neural network, Vol. 14, 2003.

[16] Z. Wu, O. Watts and S. King, "An Open Source Neural Network Speech Synthesis System", ISCA Speech Synthesis Workshop ,2016

[17] D. Xiao, B. Li and Y. Mao, "A Multiple Hidden Layers Extreme Learning Machine ethod and Its Application", Hindawi Mathematical Problems in Engineering, 2017.