Weather Sensitive Medium Term Load Forecasting using Artificial Neural Network

Isha Arora
Electrical and Instrumentation Engineering Department,
Thapar University
Patiala

Manbir Kaur
Electrical and Instrumentation Engineering Department
Thapar University
Patiala

***Abstract*— Anticipation of future load patterns is very significant for optimal decision making in power system operation and planning. The medium term load forecasting (MTLF) is used for the annual maintenance scheduling, fuel supplies scheduling, load dispatch, planning of generation shifting, planning and expansion of transmission and distribution system etc. In this study, artificial neural network (ANN) approach has been used for medium term load forecasting, in which both structure learning and parameter learning procedures are implemented. The input data is comprised of historical weather sensitive data i.e. temperature, humidity, wind speed, hour of the day, type of the day (weekday, weekend, holiday), month of the year and hourly load data.**

For structure learning, a comparative study on the multi-layer feed forward networks and recurrent networks has been performed. The performance of the network architectures is estimated on the basis of mean square error and training time. For the optimally selected network, parameter learning is performed using supervised learning and the results obtained are reported.

Keywords*— ANN; forecasting; load; mean square error (MSE); network architecture; weather parameters.

I. INTRODUCTION

Forecasting is necessary and important function in virtually every industry. Electric utilities run the power grid to deliver power to consumers all around the globe. Load forecasting, mainly referring to forecasting electricity demand and energy, is being used throughout all segments of the electric power industry, including generation, transmission, distribution.

Load forecasting is integral part of power system planning and operation [1]. It is useful in making important decisions related to load management, economical planning, unit commitment, spinning reserve allocation. Utilities always confront the challenge of meeting the increasing load demands while maximizing their short-term and long-term operational efficiency. While load forecasting provides a key input to reliable and economic systems operations, inaccurate load forecasts can lead to equipment failures or even system-wide blackout.

As the cost involved in the operation and management of power system is greatly affected by the load demand, considerable savings can be made by conduction of accurate load forecasts [2]. Presently, with the promotion of smart grid technologies, load forecasting is of even greater importance due to its applications in the planning of demand side management, distributed energy resources, etc.

Load forecasting approaches can be divided as short term forecasting, medium term forecasting, long term forecasting based on the time horizon. Medium term load forecasting which forms the basis of this paper presents the forecast of electric load one day ahead of time.

A large variety of statistical and artificial intelligence techniques have been developed for load forecasting, which include the methods employing regression models, time series, similar day approach, neural networks, expert systems, fuzzy logic models.

A naïve multiple linear regression model used for load forecasting incorporates various qualitative and quantitative factors such as customer count, weather variables like hourly temperature data, hourly load data, and calendar variables such as hour, day, and month [3]. Traditional methods works well only on linear data but ANN performs well on both linear and non-linear data. A detailed survey and comparison of distinct neural network architectures being used by researchers for rainfall forecasting is presented [4]. Different approaches of rainfall forecasting are categorized based on following features : region, training and testing period, rainfall predicting variable, types of neural network, number of input, hidden and output layers, activation function used, accuracy measure.

Nazir A. has compared different neural networks used for intrusion detection based on various performance parameters like classification rate, mean square error, training time etc [5]. ANN based short term load forecasting model has been implemented that incorporates weather related variables, historical load, seasonal variables and other special events such as holidays, weekends etc. as inputs to the neural network. Further, it has been reported that ANN model produces accurate load predictions under wide variety of power system operating conditions and efficiently integrates load pattern with weather and random effects that disturb the normal pattern of the load [6]. The feed forward neural network produced good results to forecast the load has been reported by Sharif et al[7]. A set of independent feedforward neural networks are used to forecast the load of each hour during a day. The range of training data set is adjusted to reduce the error of forecasted hourly loads of the upcoming hours.

Due to the uncertainty of input data such as weather variables, linear and non-linear regression models are unable to predict the load demands accurately. So, artificial neural networks being data-driven are used as they are capable of solving problems where the input-output relationship is neither well defined nor easily computable. They possess the additional advantage of being able to approximate any non-linear function.

The paper is organised as follows. In the second section, we introduce the concept of artificial neural networks, discuss neural network, namely, Multilayer Feed Forward Neural Network, Layer Recurrent Neural Network. Section 3 describes the details of back propagation algorithm. Section 4 addresses the performance evaluation criteria considered, compare different neural network architectures and cite the results. The conclusions of the work presented are given in Section 5.

II. ARTIFICIAL NEURAL NETWORKS

Neuron is the basic building block of an artificial neural network that functions on the pattern similar to the reasoning and learning of brain. ANNs have large number of highly inter connected processing elements (nodes or units or neurons) that usually operate in parallel and are configured in regular architectures. The connections (weights) hold the knowledge. The neural network, through a training process, learns the functional relationship between the network inputs and outputs [8]. Neural networks are simple, powerful and flexible tools for forecasting, provided that there are sufficient data patterns for training, suitable selection of the data samples, an adequate number of hidden nodes [9].

ANNs mainly comprise of three different layers namely input, hidden and output layer, each of which consists of definite number of neurons. The nodes present in the input layer carry forward the input patterns to rest of the network, without any processing. The processing of information occurs in the nodes present in hidden and output layers.

ANNs are made up of some major components like weighting factors, summation function, transfer (activation) function, output function, error function and back propagated value, learning function. After summation each unit has to output a value as a function of its net through activation or transfer function. Commonly used activation functions are step function, ramp function, sigmoid function, hyperbolic tangent function. Learning rules for ANN are classified into three categories: supervised learning, reinforced learning, and unsupervised learning.

A. *Multilayer Feedforward Networks*

Multilayer feed forward network (FFD) distinguishes itself from the single layer feed forward network by the presence of one or more hidden layers, whose computational nodes are correspondingly called hidden neurons. The function of hidden neuron is to intervene between the external input and the network output in some useful manner. By adding more hidden layers, the network is enabled to extract higher order statistics. The network is said to be fully connected if every output from one layer is connected to every node in the next layer. The input signal is applied to the neurons in the second layer. The output signal of second layer is used as inputs to the third layer, and so on for the rest of the network. Multilayer feed forward networks are the best known and most widely used kind of neural network.

B. *Recurrent networks*

When outputs are directed back as input to same or preceding layer nodes, the network is a feedback network. A recurrent neural network has at least one feedback loop. A recurrent network may consist of a single layer of neurons with each neuron feeding its output signal back to the inputs of all the other neurons. Self-feedback refers to a situation where the output of a neuron is fed back into its own input. Unlike feed forward neural networks, RNNs can use their internal memory to process arbitrary sequences of inputs. In these networks, self-loops and backward connections between the nodes are allowed. The benefit of recurrent networks is that smaller networks may provide the functionality of much larger feed forward networks [10].

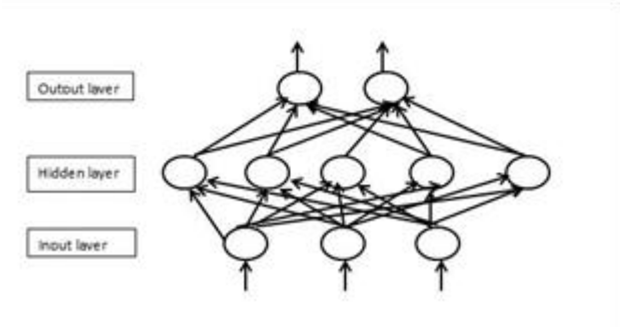

Fig.1: Multilayer feed forward Network Structure

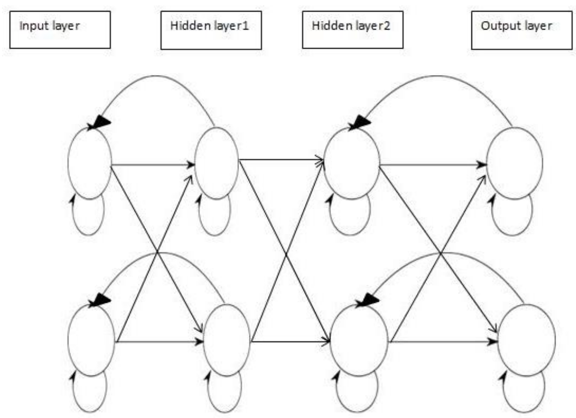

Fig.2: Recurrent Network Structure

III. BACK PROPAGATION ALGORITHM

The back propagation algorithm is applied to networks consisting of units with continuous differentiable functions and is quite good at generalization. This learning algorithm performs two main steps. First, the input patterns are promoted from the input layer to the output layer, producing an actual output, causing forward flow. The output neuron's error which is the difference between target and actual output is then typically propagated into the learning function of another node. Normally, this back propagated value, after being scaled by the learning function, is multiplied against each of the incoming connecting weights to change them before the next learning cycle begins.

The basis of the weight updation in each layer is the gradient descent method with differentiable units.

For a M layer network, m=1, 2……..M

$net_i^{k,m}$ - Net input from i^{th} unit in m^{th} layer for k^{th} pattern,

$y_i^{k,m}$ - Output of the i^{th} unit in m^{th} layer for k^{th} pattern,

w_{ij}^{m} - Connection weights between i^{th} node of m^{th} layer to j^{th} node of m-1th layer for k^{th} pattern.

Input: A set of training pattern vector x^k with the target output d^k for k=1...p.

Step1: Initialisation: Select learning rate η≥0 and maximum tolerable error E_{max}. Set k=1, where k represents k^{th} input pattern.

Step2: Forward flow: The signal is propagated forward through the network using

$$y_i^{k,m} = a(net_i^{k,m}) = a(\sum_j w_{ij}^m y_i^{k,m-1}) \tag{1}$$

for each node (i=1,...n) and layer until the actual outputs of the M^{th} layer, $y_i^{k,M}$ have all been attained, where *a(.)* denotes the activation function of corresponding layer.

Step3: Output Error: Determine the error value and error signals δ_i^M for the M^{th} layer for i=1,....n,

$$E = \frac{1}{2}\sum_{i=1}^{n}(d_i^{k,M} - y_i^{k,M})^2 \tag{2}$$

$$\delta_i^M = (d_i^{k,M} - y_i^{k,M})a'(net_i^M) \tag{3}$$

Step 4: Error back propagation: Propagate the errors backwards to update the weights w^{ij} and obtain the error signals for preceding layers.

$$\Delta w_{ij}^m = \eta \delta_i^{m-1} y_j \tag{4}$$

For m^{th} layer,

$$w_{ij}^{new} = w_{ij}^{old} + \Delta w_{ij}^m \tag{5}$$

Step 5: One Epoch: Check whether the whole set of training patterns have been presented once. If k ≤ p, then k=k+1, go to step 2; otherwise, go to step 6, where *p*= total number of training patterns.

Step 6: Evaluation of total error : Evaluate the total error, If E≤ E_{max}, then terminate the training procedure and output the final updated weights; otherwise, set E=0, k=1 and increment the epoch count and go to step 2.

IV. SIMULATION AND RESULTS

Load forecasting using ANN is carried out in different stages:

1. Selection of optimal architecture
2. Training of patterns
3. Validation and testing.

A. Selection of Network Configuration

This paper aims at estimating the impact of different combinations of multilayer networks, hidden layers and hidden nodes on the forecasting errors. We seek to analyze the relative accuracy of different kind of neural network architectural combinations.

Number of neurons in hidden layer has an influence on the learning ability of the model and the complexity of neural model varies with it. Number of neurons in the hidden layers should be carefully chosen, as too many neurons resulting in the problem of over fitting, that leads to loss of generalizing capability [11]. With too many trainable parameters, the network fails to learn the training data and performs very poorly on the test data. Whereas, if the number of neurons in hidden layer are not enough, it may be difficult for the network to train according to the historical data.

The input data is based on respective base values. Following variables are used as input neurons:

1. Temperature of hour of the day,
2. Windspeed of hour of the day,
3. Humidity of hour of the day,
4. Hour of the day,
5. Day of the week,
6. Month of the year,
7. Load of hour of the day.

Neural network model developed has one input layer with number of neurons equal to number of input variables. As the target output is forecasted hourly load, model has one output that represents forecasted load of a particular hour.

Day type categorizes days into three, namely, working day, weekend and holiday [12].

Status of days considered are as:

Weekday input neurons - 1,

Weekend input neurons - 0.5,

Holiday input neurons - 0.1.

Hour variable denotes hour of the day as load varies during the day from one hour to another [13].

The size of hidden layer is commonly determined using hit and trial method. The number of hidden neurons is started off as a fraction, that is, *2/3* the size of the input layer [14]. If the network fails to converge to a minimum error or a particular solution, more hidden neurons are needed. If the network begins to converge, addition of few more neurons is tried, and finally settles on the size based on overall performance of the network. The number of hidden layer neurons can be altered for different performance of the network. Topology of network can be changed depending on number of hidden layers, number of neurons in hidden layer.

Activation Function

Continuous differential function is being used. *tansig* has been used as the activation function for different networks considered. It is sigmoidal transfer function that gives value between -1 to +1(bipolar) or 0 to 1 (unipolar).

$$a(f) = \frac{2}{1+e^{-\lambda f}} - 1 \tag{6}$$

$$a(f) = \frac{1}{1+e^{-\lambda f}} \tag{7}$$

The performance of forecasting model is evaluated using some accuracy measure. A good load forecasting system should meet the requirement of fast speed and good accuracy. Mean squared error (MSE) and training time are taken as evaluation criteria.

Mean Squared Error: Mean Square Error (MSE) is the squared prediction error. Lesser the MSE the better the prediction of the network, this means less number of false prediction.

$$E = \frac{1}{n}\sum_{i=1}^{n}(d_i - y_i)^2 \tag{8}$$

Training time: Training time is the time required to train the network according to the parameters set for training. It is measured in seconds.

So, to begin with selection of appropriate network design, a comparative study has been considered on two different neural network configurations, multilayer feedforward network (FFD) and recurrent network (LRT).

Topology selection is started with 5 neurons in hidden layer. Several combinations are evaluated that include network with one or two hidden layers and hidden layer with 5, 10, 15 and so on. Each network configuration undergoes 1000 epochs.

The results of the different network topologies compared on the basis of performance evaluation criteria are tabulated in Table 1, where L represents number of hidden layers and N represents number of hidden nodes in each hidden layer in respective arrangement:

Table-1: MSE and Time Elapsed for Different ANN Structures

Topology No.	Type	L	N	MSE	Time (sec)	Structure
T1	FFD	1	5	3.3106	22	
T2	FFD	1	10	0.35568	20	
T3	FFD	1	15	0.34368	27	
T4	FFD	1	21	0.84788	30	overfit
T5	FFD	2	5	0.99204	25	overfit
T6	FFD	2	10	0.08473	34	overfit
T7	FFD	2	15	4:5489	60	overfit
T8	FFD	2	21	11.4569	111	overfit
T9	LRT	1	5	1.2772	20	
T10	LRT	1	10	0.72239	34	overfit
T11	LRT	1	15	1.5765	103	overfit
T12	LRT	1	21	1.6541	234	overfit
T13	LRT	2	5	0.79658	24	
T14	LRT	2	10	0.73174	148	overfit
T15	LRT	2	15	10	610	overfit
T16	LRT	2	21	3.8736	1560	overfit

It is observed that as the selection of number of hidden neurons is nearly three times the number of input neurons, the system begins to overfit. The study of different network architectures is restricted to 21 hidden neurons.

Fig. 3 and 4 depicts the performance of proposed network topologies in terms of Mean Squared Error (MSE) and training time when trained over 1000 epochs.

So, from the above results it is concluded that the multilayer feedforward network consume lesser time in training of the data as compared to recurrent networks.

B. *Parameter Learning*

Selected multilayer feedforward network in the preceding section is used for parameter learning. Error back propagation algorithm is used to train the network for hourly and weekly load forecasting.

Data used to carry out forecasting load demand is collected from Jodhpur State Load Dispatch and Communication Centre, Rajasthan Vidyut Parasaran Nigam (JVN) from March, 2012 – April, 2012.

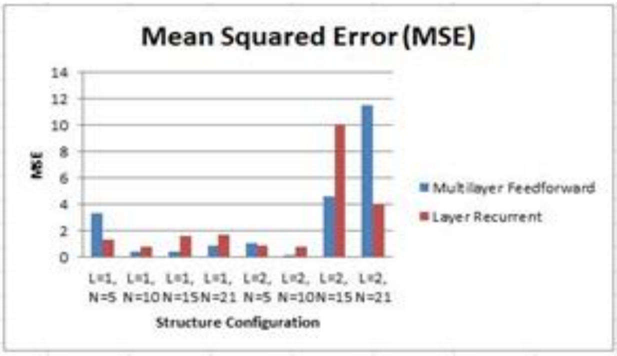

Fig.3: MSE for different Network combinations

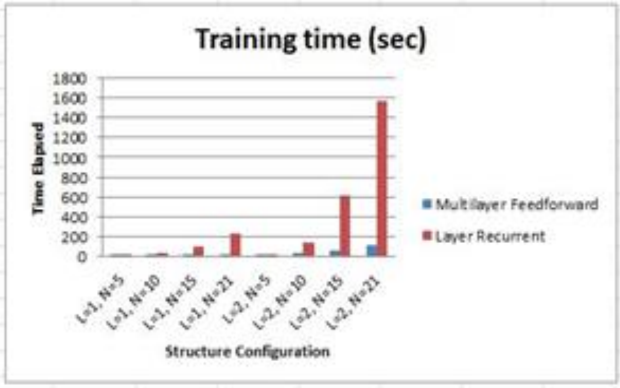

Fig.4: Training time for different Network combinations

<u>Case 1</u>

The parameter learning is performed on the optimal chosen network to forecast hourly loads. The nomenclature of 7 input neurons considered is tabulated in Table 2.

Table-2: Nomenclature of Input Neurons

Neuron No.	Parameter	Notation
1	Temperature	T
2	Wind Speed	W
3	Humidity	H
4	Month	M
5	Day type	d
6	Hour Variable	t
7	Hourly load	L(t,d)

<u>Case 2</u>

In test case 2, the parameter learning is carried out to forecast weekly loads by considering 9 input nodes namely, hourly load at a particular hour, in a week before and after, in addition to weather, load and calendar variables, as tabulated in Table 3.

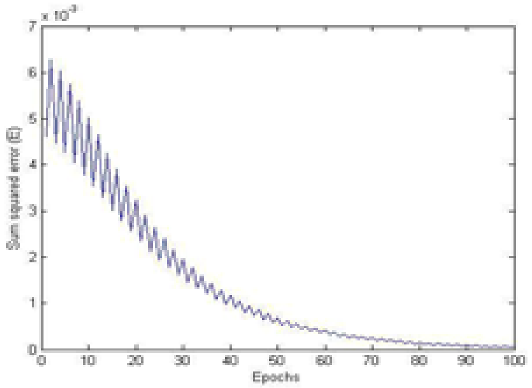

Fig.5: Mean Squared Error (MSE)

Table-3: Nomenclature of Input Neurons

Neuron No.	Parameter	Notation
1	Temperature	T
2	Wind Speed	W
3	Humidity	H
4	Month	M
5	Day type	d
6	Hour Variable	t
7	Hourly load	L(t,d,w)
8	Hourly load, week before	L(t,d,w-1)
9	Hourly load, week after	L(t,d,w+1)

The parameter learning is performed considering different number of neurons in hidden layer that are, 15, 18 and 21.

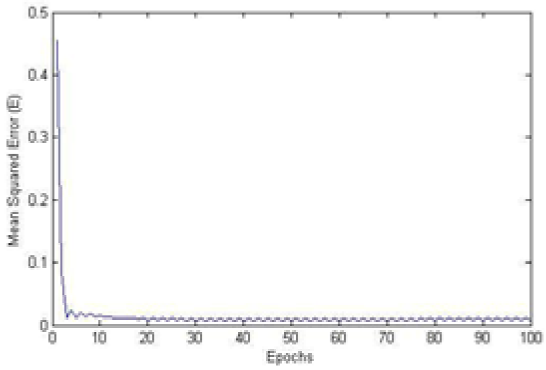

Fig.6: Mean Squared Error (MSE) with 15 hidden neurons

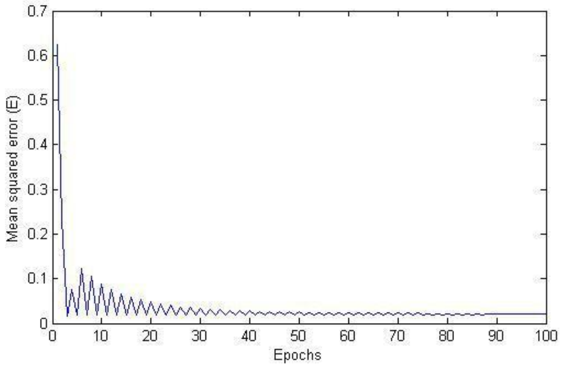

Fig.7: Mean Squared Error (MSE) with 18 hidden neurons

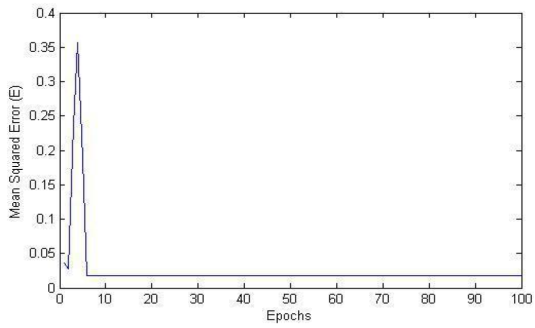

Fig.8: Mean Squared Error (MSE) with 21 hidden neurons

For 100 epochs, MSE obtained for different number of hidden neurons is as tabulated in Table 4.

Table-4: MSE obtained for 100 epochs

Case	Number of neurons	MSE
1	15	0.006
2	18	0.0214
3	21	0.0181

The multilayer feed forward network with 15 hidden neurons is reported as the optimal network due to the least Mean Squared Error (MSE).

Table-5: Nomenclature of Input Neurons

Neuron No.	Parameter	Notation
1	Temperature	T
2	Wind Speed	W
3	Humidity	H
4	Month	M
5	Day type	d
6	Hour Variable	t
7	Hourly load	L(t,d,w)
8	Temperature, week before	T(t,d,w-1)
9	Temperature, week after	T(t,d,w+1)
10	Wind Speed, week before	W(t,d,w-1)
11	Wind Speed, week after	W(t,d,w+1)
12	Humidity, week before	H(t,d,w-1)
13	Humidity, week after	H(t,d,w+1)
14	Hourly load, week before	L(t,d,w-1)
15	Hourly load, week after	L(t,d,w+1)

Case 3

In this test case, the parameter learning is applied to forecast weekly loads by considering 15 input nodes namely, temperature at a particular hour, in a week before and after, humidity at a particular hour, in a week before and after, windspeed at a particular hour, in a week before and after, hourly load at a particular hour, in a week before and after, in addition to previously considered weather, load and calendar variables, as tabulated in Table 5.

The parameter learning is performed considering different number of neurons in hidden layer that are; 15, 25, 30, 35 and 45.

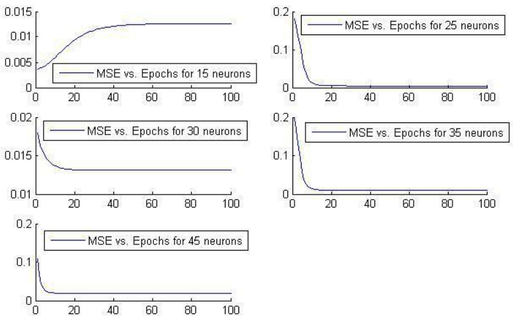

Fig.9: Mean Squared Error (MSE) for different number of hidden neurons

For 100 epochs, MSE obtained for different number of hidden neurons is as tabulated in Table 6.

Table-6: MSE obtained for 100 epochs

Case	Number of neurons	MSE
1	15	0.0126
2	25	0.004
3	30	0.0131
4	35	0.01
5	45	0.0189

Therefore, it is inferred from fig. 9 and Table 6 that minimum Mean Squared Error (MSE) is obtained with 25 hidden neurons.

The testing of the network is conducted on 10% of the training data. The results obtained verify the network configuration as optimal for the case considered in this work.

V. CONCLUSION

This paper presents artificial neural network (ANN) approach to forecast hourly loads. Different ANN structure topologies are compared to attain better results depending on two performance measures considered; which are, time elapsed in training and mean squared error obtained. Based on this evaluation, the multilayer feedforward network with a single hidden layer is considered as the best option.

For load forecasting, parameter learning using error back propagation algorithm is implemented. The obtained results of Mean square error (MSE) for 100 epochs confirm that the proposed ANN structure is effective in forecasting hourly loads.

REFERENCES

[1] Senjyu T., Takara H., Uezato K., and Funabashi T. ”One-Hour-Ahead Load Forecasting Using Neural Network”. IEEE Transactions on power systems, vol. 17, no. 1, February 2002.
[2] Osman Z. H., Awad, M. L., Mahmoud, T. K. ”Neural network based approach for short-term load forecasting”. IEEE/PES Power Systems Conference and Exposition, 2009.
[3] Hong T., Wang, P., Lee Willis, H. ”A Nave Multiple Linear Regression Benchmark for Short Term Load Forecasting”. Power and Energy Society General Meeting,IEEE, 2011.
[4] Darji M. P., Dabhi V.K., Prajapati H.B. ”Rainfall Forecasting Using Neural Network : A Survey”, International Conference on Advances in Computer Engineering and Applications (ICACEA), 2015.
[5] Nazir, A.” A Comparative Study of Different Artificial Neural Networks Based Intrusion Detection Systems”. International Journal of Scientific and Research Publications, vol. 3, issue 7, July 2013.
[6] Papalexopoulos A.D., Hao S., Peng T.M., ”An Implementation of a Neural Network based Load Forecasting Model For The EMS ” IEEE Transactions on Power Systems, vol. 9, no. 4, November 1994.
[7] Sharif S. S., Taylor, J. H. ”Real-Time Load Forecasting by Artificial Neural Networks”, IEEE, 2000.
[8] Yalcinoz T., Eminoglu U. ”Short term and medium term power distribu-tion load forecasting by neural networks”,ELSEVIER, Energy Conver-sion and Management, 2005.
[9] Zarezadeh M., Naghavi A., Ghaderi S.F. ”Electricity price forecasting in Iranian electricity market applying Artificial Neural Networks” IEEE Electrical Power and Energy Conference, 2008.
[10] Lin C.T., Lee C.S.G. Neural Fuzzy Systems, 1996.
[11] Adepoju G.A., Ogunjuyigbe S.O.A., and Alawode K.O., ”Application of Neural Network to Load Forecasting in Nigerian Electrical Power System” The Pacific Journal of Science and Technology, vol. 8, no. 1, May 2007.
[12] Panchal G., Ganatra A., Kosta Y.P., Panchal D., ”Behaviour Analysis of Multilayer Perceptrons with Multiple Hidden Neurons and Hidden Layers ” International Journal of Computer Theory and Engineering vol. 3, no. 2, April 2011 .
[13] Lu Y. ”Short- Term Load Forecasting Method Based on Structural Neural Network” IEEE 2008.
[14] Nahi Kandil N., Wamkeue R., Saad M., Georges S., ”An Efficient Approach for Short term Load Forecasting using Artificial Neural Networks” IEEE ISIE 2006