Maximizing Accuracy of Electricity Load Forecasting with Deep Learning

Kavyashree KR[1] Kamalakshi Naganna[2] Ashapurna Marndi[3]
[1]P.G. Student [2]Assistant Professor [3]Junior Scientist
[1,2]Department of Computer Science & Engineering
[1,2]Sapthagiri College of engineering, Bangalore, India [3]Council of scientific and industrial research, Bangalore, India

Abstract— Electricity is one such kind of energy that cannot be stored for a longer duration. The excess production than what is required can cause wastage where as the limited production can lead to scarcity. Thus, it is important to have a balanced production and consumption of electricity. Predicting the consumption that could occur in advance can help in this regard. The work involves generating more accurate predictions with the aid of deep learning. Initially, the neural network is made to learn from the historical data based on which it is expected to produce predictions for a new data set. The number of hidden layers and the hidden neurons is adjusted so as to get the minimum error. The accuracy of prediction is measured in terms of root mean square error (RMSE) and correlation coefficient. The number of hidden layers is increased gradually and accuracy of prediction is measured and compared with different network configurations.

Key words: Deep learning, Electricity load forecasting, Neural networks, RMSE

I. INTRODUCTION

In recent days, neural networks have gained wide importance in solving problems related to pattern recognition, classification, digital signal processing and many more. Even though concept of neural networks was introduced decades ago, it has now come to limelight due to many successes.

The concept of neural networks is based on the idea of making a machine learn from the previous historical data, observe patterns and work efficiently based on the prior experience when exposed to new data. The actions to be taken are adjusted depending on the patterns observed. A feed forward neural network which has more number of hidden layers is an example of deep architecture.

Deep learning can be defined as a branch of machine learning which comprises of visible layers of input and output and more than one hidden layer [1]. Deep neural networks require a huge set of data to work well. Thus it can be used for forecasting the electricity load because of the availability of historical data of load. This kind of historical data is categorized as time series data as it consists of a sequence of data points, which is measured successively over a regular interval of time.

II. RELATED WORK

The important characteristic of a neural network is its ability to learn from the environment and improve the performance through the experience gained from the learning process. Deep learning is said to add on to the accuracy by allowing the network to learn in depth through its multiple layers of hidden units. There has been an extensive research carried out in various areas such as weather forecasting, wind prediction, stock prediction and many more.

Xiao Ding et al [2] adopted deep learning method for stock market prediction. The work presents extracting events from news text and the training is carried out. It is claimed that this system is more capable than the previously reported ones in making profits.

James N.K Liu et al [3] apply deep learning to process massive weather data that involved millions of atmosphere records. The results of the work show that the deep neural network is able to give better predictions and thus deep neural networks can be the potential tools for time series problems.

Thomas Unterthiner et al [4] used deep learning for toxicity prediction. The major goal of the work was to identify toxicophores which are the sets of steric and electronic properties that combine to produce toxicological effect. The work proved that the deep neural networks outperformed all other traditional approaches.

Mladen Dalto et al [5] present the application of deep neural networks for ultra short term wind prediction over various locations. The results of the work prove that the deep neural networks provide increased efficiency over shallow neural networks.

III. METHODOLOGY

The work presented here is intended to know what accuracy a neural network can produce for forecasting electricity load data with different configurations of the network. The deeper networks are expected to produce greater accuracy of prediction. The methodology adopted for carrying out the work is as shown in Fig. 1.

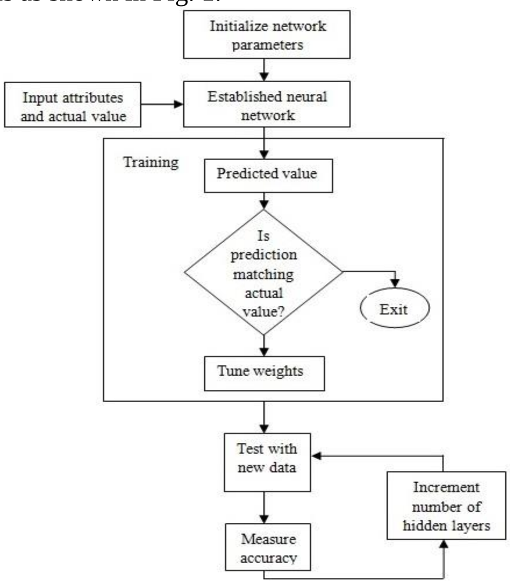

Fig. 1: Proposed Methodology

A. Data Collection

The electricity load history data from The Electric Reliability Council of Texas which consists of historical electricity load data [6] for years from 2002 to 2015 is considered for experimenting. Since holidays and weekends can affect the usage of electricity, the list of holidays and weekends is obtained from the US calendar [7].

B. Neural Network Establishment

The neural network is constructed as a standard multilayer perceptron having one input layer, one output layer and one or more hidden layers. The network parameters such as learning rate, weights, number of neurons and layers are initialized. The starting point for the number of hidden neurons to begin with is considered depending upon the following thumb rules [8]:

The number of hidden units should be between the number of input units and output units.

The number of hidden units 2/3rd the sum of input and output layer size.

The number of hidden units should not exceed twice the size of input layer.

At the later stages, the number of hidden units is varied in order to get the minimum value of RMSE.

C. Training

The network is trained by feeding the historical data set of electricity with the help of Backpropagation training algorithm. Each hidden layer receives a vector of input from the previous layer and converts it to its output vector through a linear transformation followed by a nonlinear activation.

D. Testing

On completion of training, the network is made to work on a new set of data for testing. The prediction is observed from the output and is compared with the original statistics in order to measure the accuracy of prediction.

The number of layers in the network are then increased and tested again with the new data and the accuracy is measured.

Measuring accuracy

The accuracy is measured in terms of RMSE. Lesser the value of RMSE more is the accuracy. In order to calculate RMSE, the residuals should to be determined as a first step. Residuals are nothing but the difference between the actual values and predicted values.

$$\mathrm{RMSE} = \frac{1}{\mathrm{N}}\sqrt{\sum_{\mathrm{i}=1}^{\mathrm{n}}(\mathrm{X_i}-\mathrm{Y_i})^2} \qquad (1)$$

where X_i is the actual value and Y_i is the predicted value for i^{th} observation. RMSE will have the same unit as that of the quantity plotted on the vertical axis.

Another measure of accuracy is correlation coefficient denoted by “r” which measures the direction and strength of linear relationship between two quantities.

$$r = \frac{1}{n-1}\left(\frac{\sum_x\sum_y(X-\bar{X})(Y-\bar{Y})}{S_xS_y}\right) \qquad (2)$$

where X is the actual value and Y is the predicted value. S_x and S_y are the standard deviations of X and Y respectively which are calculated as,

$$S_x = \sqrt{\frac{\sum(X-\bar{X})^2}{n-1}} \qquad (3)$$

The correlation coefficient, r takes the values in the range [-1, +1]. A perfect correlation of ± 1 is said to occur when all the data points lie on a straight line exactly. A correlation above 0.8 is normally considered as strong, whereas a correlation below 0.5 is taken as weak.

IV. EXPERIMENTAL RESULTS

The experiment is conducted with different configurations of the network by varying the number of hidden nodes and layers in the network. The network established has seven input nodes for accepting attributes such as year, month, day, weekend, holiday etc and one output node for giving out the prediction. The number of input and output nodes and layers remains constant throughout the experimentation and only the number of hidden nodes and layers is tuned. The initial number of hidden nodes is set to 5 in order to have it as 2/3rd of the sum of the number of input and output nodes [8].

Then the number of hidden nodes and layers is increased and the RMSE and correlation coefficient are calculated at each stage in order to determine which configuration gives less RMSE thus leading to greater accuracy of prediction. Table 1 gives the summary of various network configurations for which the experiment is conducted and their respective RMSE.

Number of hidden layers	Number of nodes in hidden layer 1	Number of nodes in hidden layer 2	Number of nodes in hidden layer 3	Number of nodes in hidden layer 4	RMSE
1	5	0	0	0	1552.4729
1	6	0	0	0	1250.6551
1	7	0	0	0	1241.1005
1	8	0	0	0	1232.6297
1	9	0	0	0	1332.9542
1	10	0	0	0	1321.2639
2	8	1	0	0	1345.0772
2	8	2	0	0	1270.0621
2	8	3	0	0	1226.7985
2	8	4	0	0	1817.2596
2	8	5	0	0	1300.2312
2	8	6	0	0	1526.55

					39
3	8	3	1	0	1841.3759
3	8	3	2	0	1636.2671
3	8	3	3	0	1820.3268
3	8	4	4	0	1182.2543
3	7	4	1	0	1279.2214
3	7	4	2	0	1285.2204
3	7	4	3	0	1249.8711
3	6	5	4	0	1362.1324
3	6	5	3	0	1352.5375
4	8	4	4	1	1207.2011
4	8	4	4	2	1335.6196
4	8	4	4	3	1260.0506
4	8	4	4	4	1175.9736

Table 1: RMSE for different network configurations

Initially, the network is trained and tested with one hidden layer; the number of hidden nodes being tuned to get a lower value of RMSE. The least RMSE recorded for 1 hidden layer is 1232.6295. Then experiment is repeated for a 2-layer network and the lowest RMSE here is found to be 1226.7985. It is observed that the RMSE noted for a 2-layer network is lesser than that of 1-layer network.

As a further step, the network is implemented with 3 hidden layers, each layer having 8, 4 and 4 hidden units respectively. The minimum RMSE recorded for this configuration is found to be 1182.2543 which is lesser than that for 1-layer and 2-layer network. For a 4 hidden layer configuration, RMSE falls down to 1175.9736.

Number of hidden layers	RMSE	Correlation coefficient
1	1232.6295	0.8763
2	1226.7985	0.8865
3	1182.2543	0.8928
4	1175.9736	0.8956

Table 2: Comparison of accuracy measures

From the table 2, it can be seen that the RMSE decreases with the increase in the number of hidden layers in the network. The value of correlation coefficient increases with increase in the number of hidden layers indicating strong correlation between the original and predicted values of electricity load. So, it can be said that the deeper network provides greater accuracy of prediction.

V. Conclusion

As an outcome of the work, a model is built for forecasting the electricity load which is able to predict loads on new examples with a decent accuracy of prediction. Further, the number of hidden layers in the network is increased at each stage in order to make the network deeper and the accuracy is compared for each increment in the number of layers. After this experimentation, the network with more number of hidden layers is found to be yielding greater accuracy.

VI. References

[1] Dandan Mo, “A survey on deep learning: One small step toward AI”, Year: 2012.
[2] Xiao Ding, Yue Zhang, Ting Liu, Junwen Duan, “Deep learning for even driven stock prediction”, Proceedings of the Twenty-Fourth International Joint Conference on Artificial Intelligence. Year: 2015.
[3] James N.K. Liu, Yanxing Hu, Jane Jia You, Pak Wai Chan, “Deep neural network based feature representation of weather forecasting”, Hong Kong Polytechnic University.
[4] Thomas Unterthiner, Andreas Mayr, Gunter Klambauer, Sepp Hochreiter, “Toxicity prediction using deep learning”, Year: 2015.
[5] Mladen Dalto, Jadranko Matuˇsko, Mario Vasak, “Deep neural networks for ultra-short-term wind forecasting”, IEEE, 978-1-4799-7799-4/15, Year:2015.
[6] http://www.ercot.com/gridinfo/load/load_hist/ Date: 14-12-2015, Time: 16:40.
[7] http://www.timeanddate.com/holidays/us/2000 Date: 20-03-2016, Time: 13: 20.
[8] Gaurang Panchal, Amit Ganatra, Y P Kosta and Devyani Panchal, “ Behaviour Analysis of Multilayer Perceptrons with Multiple Hidden Neurons and Hidden Layers”, International Journal of Computer Theory and Engineering, Vol. 3, No. 2, ISSN: 1793-8201, April 2011.
[9] Jiawei Han, Micheline Kamber, “Data Mining: Concepts and Techniques”, 3rd Edition, Morgan Kaufmann Series, Year: 2011.
[10] Navjot Kaur, Er. Amrit Kaur, “Electricity demand prediction using artificial neural networks in data mining”, International Journal For Technological Research In Engineering Volume 2, Issue 8, Year: April 2015.