Electricity Load Forecasting Methods-A Review

Ajay Kumar[1], Rishabh Jain[2], Awinash Pandey[3], Shivani Prakash Gupta[4], Shivangi Garg[5]
Department of Computer Science & Engineering
JSS Academy of Technical Education
Noida, India
(E-mail: ajay.itcs@gmail.com)

Abstract—Electricity load forecasting plays quite an important role in smart grids system stations. Due to large variations in power & demand supply, the electricity load prediction plays an important role in efficient power management system. To solve this issue various researchers have adopted various methodologies. This paper presents the six approaches of electricity load forecasting using Recurrent Neural Networks and its variants, Support Vector Machines, Auto-regressive methods and K-means method. In this paper, we also classify all the approaches on the basis of their advantages and shortcomings. Besides all this, different important points are presented which need more stress while doing load prediction when the surrounding is competitive and not regulated.

Keywords—*load forecasting; recurrent neural network; machine learning; support vector machine; k-means.*

I. INTRODUCTION

Nowadays, the electric load prediction is a vital process with applications that are used in various fields because of large demand of power supply and consumption in an electric power houses and in households respectively. The reasons for accurate electricity load forecasting are: purchasing and producing electric power, transmitting, transferring and distributing electric power, managing and maintaining the electric power sources, managing the daily electric load demand, financial and marketing planning.

Buildings are identified as a major energy consumer worldwide, accounting for 20%-40% of the total energy production [16-18]. In addition to being a major energy consumer, buildings are shown to account for a significant portion of energy wastage as well [19]. As energy wastage poses a threat to sustainability, making buildings energy efficient is extremely crucial. Therefore, in making building energy consumption more efficient, it is necessary to have accurate predictions of its future energy consumption.

Further, demand or load forecasting is crucial for mitigating uncertainties of the future [20]. In that, individual building level demand forecasting is crucial as well as forecasting aggregate loads. In terms of demand response, building level forecasting helps carry out demand response locally since the smart grids incorporate distributed energy generation. The advent of smart meters has made the acquisition of energy consumption data at building and individual site level feasible. Thus data driven and statistical forecasting models are made possible [21].

Most of the research for prediction of electricity load forecasting concentrate on aggregate load at the system level .However variation at the individual level also play an important role in determining accuracy of load prediction.

Machine learning is closely related to computational statistics, which also focuses on prediction-making through the use of computers. It has strong ties to mathematical optimization, which delivers methods, theory and application domains to the field. Machine learning is sometimes conflated with data mining, where the latter subfield focuses more on exploratory data analysis and is known as unsupervised learning. Machine learning can also be unsupervised and be used to learn and establish baseline behavioral profiles for various entities and then used to find meaningful anomalies.

This paper presents the various methodologies for electricity load prediction using machine learning approaches and the comparison of these methodologies in terms of their advantages and disadvantages..

II. SURVEY OF LOAD PREDICTION METHODS

Considering the enormity of the field, we went through a number of papers, but the following are the few examples while the rest are mentioned in the references. A lot of research work has been done by the researchers in order to solve the problem of electricity load forecasting. This section presents some of the research papers reviewed.

A. Forecasting using RNN networks [1-2][5-8]

RNNs are designed specifically to operate over sequential data or time series [7]. RNN Networks act as a feedback network that permits the signals to traverse both forward and backward. Whereas feed-forward networks allows to travel from input layer to the output layer. They introduce loops in the network and allow internal connections among hidden units.

RNN [26] is best method to predict the future data on the basis of past data by using so much internal connections. In particular, RNNs make it possible to explore temporal relationships among the data that are far away from each other [8].

Some problem arises during the training procedure, when RNN architecture [18] can model any system which is dynamic because it has to model the system which must be learned from data in order to solve a target task. RNNs are more suitable for exploiting the information in the past data to forecast the future data [22-25].

B. Load Forecasting using GRU-RNN Mode [9]

The Gated Recurrent Unit (GRU) is another notorious gated architecture, originally proposed by Cho et al., which adaptively captures dependencies at different time scales. GRU can control how much hidden layers can be remembered or not by combining forget and input gates are into a single update gate, the internal state in GRU is always fully exposed in output, due to the lack of a control mechanism, like the output gate in LSTM [3-4].

GRU can modify the memory content of the cells because of highly nonlinear statistical dependencies therefore making this an efficiently model. The performance of GRU is amazing in many sequence learning problems, the additional complexity of the complicated gated mechanisms seems to be unnecessary in many time series predictions tasks.

The limitation of this approach is regarding training. The training is slower and complex due to unfolding and back propagation with time. However, while some precautions need to be taken in the design of these networks, satisfactory results can be obtained with minimal fine-tuning and by selecting default hyper parameters. This implies that a strong expertise on the data domain is not always necessary.

C. Load Forecasting using NARX Model[10]

Nonlinear Auto-Regressive exogenous inputs neural network (NARX) is one of the type of RNN.NARX networks are recurrent dynamic architectures with several hidden layers and they are inspired by discrete-time nonlinear models called Nonlinear Auto-Regressive with eXogenous inputs. Differently from other RNNs, the recurrence in the NARX network is given only by the feedback on the output, rather than from the whole internal state.NARX networks have also been adopted by Plett in a gray-box approach for nonlinear system identification. A NARX network can be implemented with a Multilayer Perceptron (MLP).

NARX network is not immune to the problem of vanishing and exploding gradients. NARX struggles in reaching performance comparable with the other architectures. The optimal configuration of NARX network is characterized by a quite large number of hidden nodes and layers, which denote a network of higher complexity with respect to the ones identified in the other tasks. On average, the NARX network achieved the lowest performance, especially on synthetic problems.

D. Traditional Methods like ARIMA Model[11]

ARIMA model is an auto-regressive statistical forecasting model. In particular, the ARIMA is one of the most popular and commonly used methods for time series forecasting. However, these methods work under the assumption that the observed time series and the future time series are linearly related, which makes them less effective for time series with significant nonlinear characteristics.

Cao et al. [11] adopted autoregressive integrated moving average (ARIMA) model and similar day method for intraday load forecasting. The mechanism of their similar day method is to group the targeted day with meteorologically similar days in the history and predict the load based on the average demand of those days. It was demonstrated that in ordinary days, ARIMA performs better. It works best when data exhibits a stable or consistent pattern over time with a minimum amount of outlier.

E. Power load forecasting using Support Vector Machine[12]

SVM is a classifier that tries to find a hyper plane which can divide data into the correct classes. The Support vector is a part of data that could help to determine the hyper plane.SVM is a classifier that tries to find a hyper plane which can divide data into the correct classes. The Support vector is a part of data that could help to determine the hyper plane. It offers great predictive capability for a limited sample size. Although SVM is a promising approach, the following challenges need to be addressed on making the accurate electricity price forecasting.

High computational complexity: According to Hu’s work [12], SVM is weak in processing uncertain information and has high computational complexity. In the electricity price forecasting, irrelevant and redundant features take great computation complexity to the training process of SVM and also decrease the forecasting accuracy.

Hard to tune parameters: There are three super parameters which are cost penalty, insensitive loss function parameter, and kernel parameter. It’s hard to tune these parameters for higher accuracy and more efficiency.

The common method used to adjust SVM super parameters is gradient descent (GD) algorithm or cross validation [13]. However, these two methods bring much computational complexity and may be unable to converge.

F. An Efficient K Nearest Neighbour Algorithm[14][15]

K-Nearest Neighbor is a simple algorithm that memorizes the training data and performs classification only if the attributes of an object exactly match one of the training examples.

It finds the group of k objects in the training set that are closest to the test object and assigns a label to the test set based on the predominance of a particular class in this neighborhood. Among clustering formulations that are based on minimizing a formal objective function, perhaps the most widely used and studied is k-means clustering.

K-nearest neighbor (KNN) algorithm had also seen some successful examples on load forecasting [11-12], whose dominant advantage is its efficiency. It is simple and powerful. No need for tuning complex parameters to build a model. No training involved (“lazy”).

However it doesn't handle 'soft' boundaries - i.e. areas where some cases appear on either side of a boundary. The main disadvantage of the KNN algorithms is that it is a lazy learner, i.e. it does not learn anything from the training data and simply uses the training data itself for classification. However, all of the above methods focus on learning and forecasting load at the system or substation level. In order to support future smart grid applications, effective load forecasting techniques for electricity users are gaining increasing interest.

III. MERITS AND SHORTCOMINGS OF VARIOUS APPROACHES

After the detailed literature review of various approaches of discovery we can lists the advantages and shortcomings of each approach. The merits and limitations of all the existing approaches are given below in Table 1.

TABLE 1. Merits and Shortcomings of various Electricity Load prediction approaches

	Merits	Shortcomings
RNN networks based approach	1. More suitable for exploiting the information in the past data to forecast the future data. 2. Suitable for dynamical system.	1. Difficult training when dataset is large. 2. Results in less reduced error 3. Complex architecture. 4. Difficult to implement
GRU-RNN based approach	1. It can model highly nonlinear statistical dependencies 2. Strong expertise on the data domain is not always necessary 3. Less time to reach convergence	1. The training is slower due to unfolding. 2. More complex in nature due to backpropogation through time procedure. 3. Minimum fine tuning is required for satisfactory results.
NARX Model based approach	1. Number of delays are finite. 2. It supports discrete-time nonlinear models 3. It can be implemented with Multi layer perceptron network 4. Non linear mapping function is there.	1. NARX network is not immune to the problem of vanishing. 2. Large number of hidden nodes and layers. 3. High complex network architecture 4. Lowest performance
ARIMA Model based approach	1. Statistical forecasting model . 2. Effective and reliable Approach. 3. It works best when data exhibits a stable or consistent pattern	1. Difficult to implement. 2. Large number of parameters. 3. Unexplainable random error.
Support Vector Machine based approach	1. High predictive capability for a limited sample size 2. Data item as a point in n-dimensional space. 3. Better results with reduced dataset.	1. High computational complexity. 2. Irrelevant and redundant features 3. Decreased the forecasting accuracy.
K Nearest Neighbour based approach	1. It is simple and powerful. 2. Highly efficient 3. No training involved. 4. Easy to implement	1. It doesn't handle 'soft' boundaries - i.e. areas where some cases appear on either side of a boundary. 2. it is a lazy learner, i.e. it does not learn anything from the training data.

After the brief discussion on the approaches to predict the electricity load, it is observed that there are some merits and shortcomings of each. So the comparison of approaches on the basis of some parameters is mentioned in Table 2 given below.

TABLE 2. Comparison of various Electricity Load prediction approaches

Parameters / Approaches	Trend Analysis	End-use Analysis	Econometric modeling	Similar day approach	Regression	Time series modeling
RNN networks based approach	√	√	√	√	√	×
GRU-RNN based approach	√	√	√	√	√	×
NARX Model based approach	√	√	√	√	√	√
ARIMA Model based approach	√	√	√	√	√	√
Support Vector Machine based approach	√	√	√	√	√	×
K Nearest Neighbor based approach	×	×	×	×	×	×

IV. CONCLUSION

This paper describes the research work done by the researchers to predict the electricity load in order to solve the issues of power usage, theft, demand and supply etc. It concludes the merits and demerits of all the methodologies to predict the electricity load using various approaches.

In order to eliminate all the limitations of these methods, there is a need to develop a system for efficient load prediction using machine learning approach which will provide better results than existing approaches.

V. FUTURE WORK

A deep learning approach of machine learning can be implemented by using Recurrent Neural Networks which not only removes the existing issues but also improves the prediction results.

VI. REFERENCES

[1]. Jian Zheng, Cencen Xu ,Ziang Zhang and Xiaohua Li,"Electric Load Forecasting in Smart Grids Using Long Short Term Memory based Recurrent Neural Network.", IEEE Transactions on Power Systems, 2017.

[2]. Weicong Kong,Zhao Yang Dong, Youwei Jia,"Short -Term Residential Load Forecasting based on LSTM Recurrent Neural Network", IEEE Transactions on Smart Grid,2017.

[3]. S. Ryu, J. Noh, and H. Kim, "Deep neural network based demand side short term load forecasting," IEEE International Conference on Smart Grid Communications (SmartGridComm), pp. 308-313, 2016.

[4]. Y. Bengio, P. Simard, and P. Frasconi, "Learning long-term dependencies with gradient descent is difficult," IEEE Transactions on Neural Networks, vol. 5, pp. 157-166, 1994.

[5]. S. Hochreiter, Y. Bengio, P. Frasconi, and J. Schmidhuber, "Gradient flow in recurrent nets: the difficulty of learning long-term dependencies," A field guide to dynamical recurrent neural networks, 2001.

[6]. F. A. Gers, r. A. Schmidhuber, and F. A. Cummins, "Learning to Forget:Continual Prediction with LSTM," Neural Comput., vol. 12, pp. 2451-2471, 2000.

[7]. L. Medsker and L. Jain, "Recurrent neural networks," Design and Applications, 2001.

[8]. R. Pascanu, T. Mikolov, and Y. Bengio, "On the difficulty of training recurrent neural networks." ICML (3), vol. 28, pp. 1310–1318, 2013.

[9]. F. M. Bianchi, E. De Santis, A. Rizzi, A. Sadeghian, Short-term electric load forecasting using echo state networks and PCA decomposition, IEEE Access vol. 3, pp. 1931–1943, 2015.

[10]. R. Napoli, L. Piroddi, Nonlinear active noise control with NARX models, IEEE Transactions on Audio, Speech, and Language Processing Vol. SMC-15, No.1, pp.286–295 2010.

[11]. Z.Yun, Z. Quan, S. Caixin, L. Shaolan, L. Yuming, and S. Yang, "RBF Neural Network and ANFIS-Based Short-Term Load Forecasting Approach in Real-Time Price Environment," IEEE Transactions on Power Systems, vol. 23, pp. 853-858, 2008.

[12]. S. Ryu, J. Noh, and H. Kim, "Deep neural network based demand side short term load forecasting," IEEE International Conference on Smart Grid Communications (SmartGridComm), pp. 308-313, 2016.

[13]. A. Shiri, M. Afshar, A. Rahimi-Kian, and B. Maham, "Electricity price forecasting using Support Vector Machines by considering oil and natural gas price impacts," Proceedings of IEEE International Conference on Smart Energy Grid Engineering (SEGE), pp.1-5, 2015.

[14]. R. Zhang, Y. Xu, Z. Y. Dong, and W. Kong, "A Composite k-Nearest Neighbor Model for Day-Ahead Load Forecasting with Limited Temperature Forecasts," presented at the IEEE General Meeting Boston, 2016.

[15]. F. H. Al-Qahtani and S. F. Crone, "Multivariate k-nearest neighbour regression for time series data - A novel algorithm for forecasting UK electricity demand," in Neural Networks (IJCNN), presented at International Joint Conference, pp. 1-8,2013.

[16]. L.P. Lombard, J. Ortiz, C. Pout, "A review on buildings energy consumption information," in Energy and Buildings, Vol. 40, pp. 394–398, 2008.

[17]. K. Amarasinghe, D. Wijayasekara, H. Carey, M. Manic, D. He, W. Chen, "Artificial Neural Networks based Thermal Energy Storage Control for Buildings," in Proc. 41st Annual Conference of the IEEE Industrial Electronics Society, IEEE IECON 2015, Yokohama, Japan, Nov. 09-12, 2015.

[18]. D. Wijayasekara, M. Manic, "Data-Fusion for Increasing Temporal Resolution of Building Energy Management System Data," in Proc. 41st Annual Conference of the IEEE Industrial Electronics Society, IEEE IECON 2015, Yokohama, Japan, Nov. 09-12, 2015.

[19]. Sareh Naji, Afram Keivani, Shahaboddin Shamshirband, U. Johnson Alengaram, Mohd Zamin Jumaat, Zulkefli Mansor, Malrey Lee, "Estimating building energy consumption using extreme learning machine method", Energy, Volume 97, Pages 506-516, 15 February 2016.

[20]. Elena Mocanu, Phuong H. Nguyen, Madeleine Gibescu, Wil L. Kling, Deep learning for estimating building energy consumption, Sustainable Energy, Grids and Networks, Pages 91-99,6 June 2016.

[21]. Antans Sauhats,Renata Varfolomejeva ,Olegs Linkevics ,"Analysis and prediction of electricity consumption using smart meter data", presented at IEEE 5th International Conference on Power Engineering.Energy and Electrical Drives,2015.

[22]. F. Rosenblatt, Principles of neurodynamics: perceptrons and the theory of brain mechanisms, ser. Report (Cornell Aeronautical Laboratory). Spartan Books, 1962.

[23]. M. Ghofrani, M. Hassanzadeh, M. Etezadi-Amoli, and M. S. Fadali, "Smart meter based short-term load forecasting for residential customers," North American Power Symposium (NAPS), pp. 1-5, 2011.

[24]. M. Espinoza, J. A. Suykens, R. Belmans, and B. De Moor, "Electric load forecasting," IEEE Control Systems, vol. 27, no. 5, pp. 43–57, 2007.

[25]. I. Sutskever, O. Vinyals, and Q. V. Le, "Sequence to sequence learning with neural networks," Advances in neural information processing systems, 2014, pp. 3104–3112.

[26]. Z. C. Lipton, J. Berkowitz, and C. Elkan, "A critical review of recurrent neural networks for sequence learning," arXiv pre printarXiv:1506.00019, 2015.

[27]. M. D. Zeiler, "ADADELTA: an adaptive learning rate method," arXiv preprint arXiv:1212.5701, 2012.

[28]. I. Sutskever, O. Vinyals and Q. V. Le, "Sequence to Sequence Learning with Neural Networks," in NIPS 2014, 2014.

Ajay Kumar, Assistant Professor in Department of Computer Science & Engineering at JSS Academy of Technical Education, Noida. He is pursuing Ph.D. in YMCA UST Faridabad. He received his M.Tech degree from YMCA UST Faridabad. He did his B.Tech from U.I.E.T at M.D.U. Rohtak. His areas of interests are Machine Learning, Semantic Web & Web Services, and Information Retrieval System.

Rishabh Jain, Student (final year) B.Tech in the Department of Computer Science and Engineering at JSS Academy of Technical Education, Noida. He works at the junction of Software Development & Research which lie at the intersection of Machine Learning and Computer Vision.

Awinash Pandey, Student (final year) B.Tech in Department of Computer Science Engineering at JSS Academy of Technical Education,Noida. His areas of interest are Machine Learning and Database Management System.

Shivani Prakash Gupta, Student (final year) B.Tech in the Department of Computer Science and Engineering, JSS Academy of Technical Education, Noida. Her areas of interests are User experience, User Interface, and Graphic Design.

Shivangi Garg, Student (final year) B.Tech in the Department of Computer Science Engineering at JSS Academy of Technical Education, Noida. Her areas of interests are Backend Development, Machine Learning, and Database Management.