International Journal of Scientific & Engineering Research Volume 11, Issue 12, December-2020
ISSN 2229-5518

Influence of Increasing Number of Epochs on Electric Power Consumption Dataset

Md. Abdur Rahman
Department of Mathematics, Jahangirnagar university, Savar, Dhaka, Bangladesh
E-mail: marahmanju@juniv.edu

Abstract— Electricity energy consumption prediction plays an important role in short-term energy allocation with the increasing demand for energy in recent years. Predicting this accurately helps to make better decisions by minimizing the cost and making energy efficient. Nowadays, machine learning and deep learning approaches are used to predict energy consumptions for the individual home, building, or campus. Different models developed by these approaches are required different times for training which largely differs on their architecture as well as the dataset. We studied the hidden patterns of time series data using heat-map, autocorrelation, and partial correlation. This proposed work focused on how epochs affect the prediction of power consumption for time series hourly data through internal analysis of long short term memory networks (LSTMs). The proposed network used the PJME dataset of the USA to predict the consumption patterns by avoiding over-fitting. The prediction accuracy is recorded from 94.23% to 95.87%. Finally, this work offers important information about how the variation of the number of epochs affects the training loss and validation loss as well as accuracy based on the size of the model and the pattern of data of the dataset.

Index Terms— Energy Consumption, Autocorrelation, Deep Learning, Recurrent Neural Networks, Long Short Term Memory Networks, Heatmap, Error Metrics.

1 INTRODUCTION

ELECTRICAL energy demand is thriving nowadays for buildings, population growth, and economic enhancement [1]. Researchers around the world invented several approaches from diverse data sets to predict building energy consumption [2-5]. In a developed country like the United States of America, 40% of the entire energy demand arises from the buildings [6]. Consequently, for the utmost prediction, it is so much essential to detect the best norm and approaches. Short-term energy consumption forecasting techniques are very workable for residential energy demand [7]. At present, in this fast-growing Artificial Intelligence (AI) Technology era, the deep learning methods are proficient in executing very short-term energy consumption forecasting outcomes with impressive high prediction accuracy [8-11].

For predicting electricity cost, several algorithms developed using several kinds of methods like linear regression [12], Support Vector Machines (SVM) [13], Autoregressive integrated moving average (ARIMA) [14], Artificial Neural Networks (ANN) [15], etc. For time series prediction, ARIMA and SVM applied in household and mercantile buildings' electric energy demand in the utmost familiar manner. [16-17]. Nevertheless, to acquire better outcomes, the non-linear models have been demonstrated for short-term forecasting [15]. When we compared these models with ANN, it appears as immensely improved models with higher performance than previous methods [18]. So, for predicting on time-series datasets, deep learning models have a higher competency of accuracy; for example, Boukoros et al. (2017) has shown in their work how it has more accuracy on time-series datasets [19]. Fayaz and Kim [4] used one year's weekly and monthly datasets for electronic energy consumption prediction of residential buildings and proposed a proficient model using a deep extreme learning machine. Krishnan et al. (2018) enhanced the optimization approach for electric energy demand prediction using RNN [3]. The utmost distinguished deep neural networks are RNN and LSTM. Consequently, deep learning models are extensively applied for electricity energy prediction; for example, W. Kong et al. (2017) has shown the utmost possible accuracy in residential households using LSTM recurrent neural networks [11].

The rest of this paper is ordered as the following ways. Data as well as patterns of observations are provided in Section 2. Methods used to analyze and predict energy consumption describes in the Section 3. Section 4 deduces the results and comparison of the experiential analysis, and it concludes with section 5.

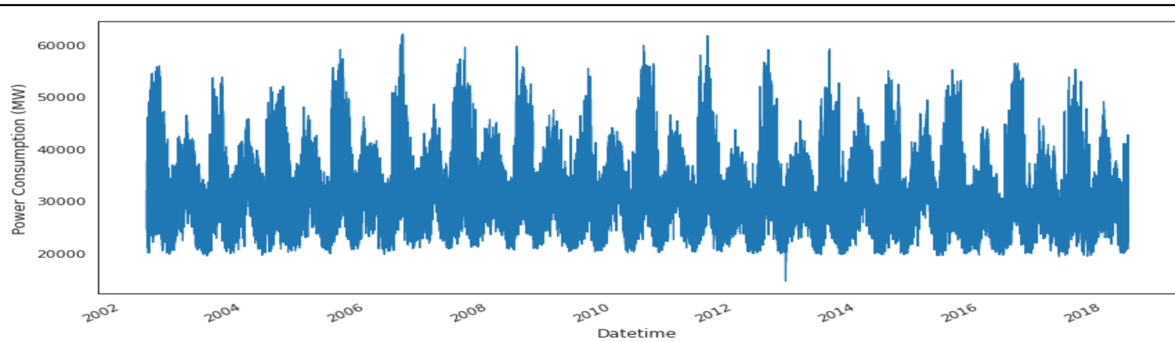


Fig. 1 Power consumption for a building electric power consumption of PJME dataset (2002-2018)

ISSN 2229-5518

Table 1. PJME Dataset Information

Epoch	Count	Mean	Std	Min	Max	25%	50%	75%
Total Consumption (MW)	145366	32080.22	6464.01	14544	62009	6464.01	14544	62009

2 Data

This proposed work used PJME dataset generated by PJM, USA which has 145366 records for a building. It contains hourly data from 2002 to 2018. This whole dataset is divided into two: 112134 for training, 28011 for testing. Fig. 1 shows the hourly power consumption for this building. Also, Table 1. provides statistical information of this dataset in which mean, standard deviation, minimum, maximum, and percentile information are displayed. There are higher power consumption in Summer (July-August) shown in Fig. 2. Fig. 3 tells the lower power consumption at the starting of week (Monday-Tuesday) and higher consumption (Friday-Saturday).

Autocorrelation is mostly defined that how much connection is there among a specific time point and a previous point. For example, today's power consumption is greatly connected with the previous week's consumption. Also, yesterday's consumption is greatly correlated with today's consumption. From the graph (Fig. 4(a)), it is observed that it looks repetition itself every six months. We can conclude that this is happened because of seasonality or weather patterns. There are two main seasons: winter and summer. Residents are supposed to consume the highest electric power because of more heat or cooling required in these sea-sons. These same scenarios are observed from the heatmap (Fig. 5) in which there are highest consumption in winter (December and January) and summer (July-August). Bold red denotes the highest electric consumption.

3 Methods

3.1 Artificial Neural Networks (ANN)

It closely looks like the human brain and uses set of algorithms. They process sensory data through a perception to recognize patterns. An ANN usually contains a collection of processors organized in tiers and working in parallel. The first tier is for receiving the raw input, and each the preceding tier. The output of the system is produced from the last tier. Fig. 6 (a) shows the basic diagram of neural network, and typical two-layer neural network is expressed by Fig. 6 (b).

3.2 Convolutional Neural Networks

Convolutional Neural Networks (CNNs) have been successfully applied to many problems in computer vision and medical image analysis. In our application, the convolutional layers were constructed using one-dimensional kernels that move through the sequence (unlike images where 2D convolutions are used). These kernels act as filters which are being learned during training. As in many CNN architectures, the deeper the layers get, the higher the number of filters become.

3.3 Long Short Term Memory networks (LSTMs)

Long Short Term Memory networks (LSTMs) are a type of Recurrent Neural Networks (RNNs). These are efficient of learning long-term dependencies. Hochreiter & Schmidhuber (1997) introduced these networks. LSTMs are devised in order to avoid the long-term dependency issue. They can remember required information for long periods of time. These equations (1-6) are used to express input gate (it), forget gate (ft), output gate (ot), cell status, etc.

$$i_t = \sigma(W_i.[h_{t-1}, x_t] + b_i) \quad (1)$$
$$\tilde{C}_t = tanh(W_C.[h_{t-1}, x_t] + b_C) \quad (2)$$
$$C_t = f_t.C_{t-1} + i_t.\tilde{C}_t \quad (3)$$
$$f_t = \sigma(W_f.[h_{t-1}, x_t] + b_f) \quad (4)$$
$$O_t = \sigma(W_o.[h_{t-1}, x_t] + b_o) \quad (5)$$
$$h_t = O_t * tanh(C_t) \quad (6)$$

4 Results and Discussion

We used PJME dataset generated by PJM, USA which has 145366 records for a building. It contains hourly data from 2002 to 2018. This whole dataset is divided into two: 112134

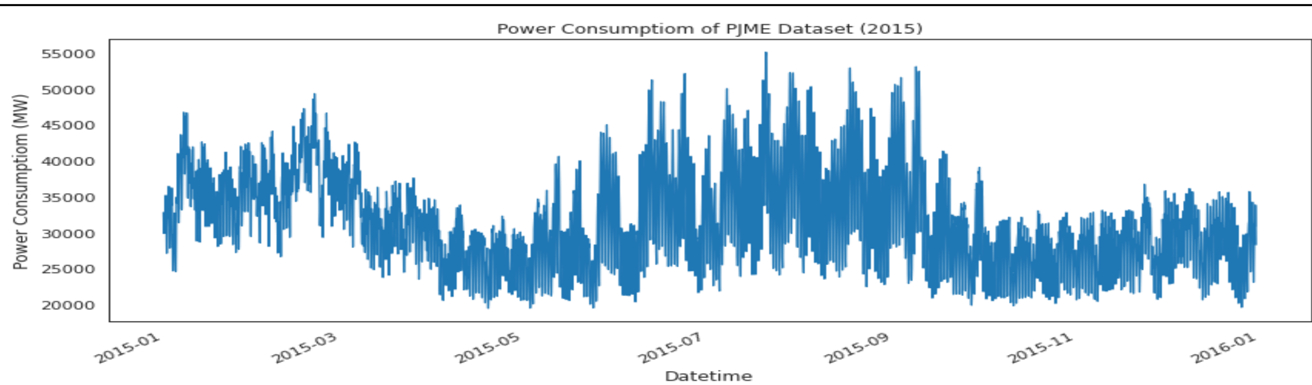


Fig. 2 Power consumption for a building electric power consumption of PJME dataset (2015)

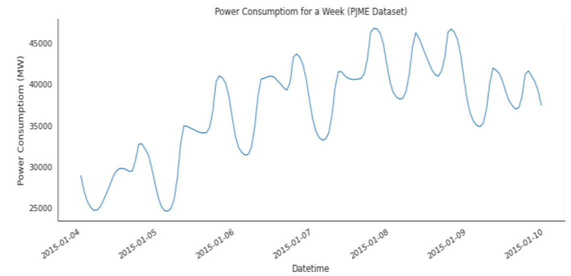


Fig. 3 Power consumption for a building electric power consumption for a week (Monday-Sunday)

for training, 28011 for testing. This section provides the results from experiments through using performance metrics for time series prediction. This model is trained with 5-30 epochs, batch size 70, and adam optimization. Also, initial learning rate is 0.001. We have chosen Python along with Keras library with google colab environment to run the simulation code.

We have used different number of epochs: 5, 10, 15, 20, 25, and 30. Different number of epochs that are required to train the model are responsible to determine prediction accuracy. Indeed, this accuracy also varies in terms of the size of the model and the variation of the dataset. We have studied how the training loss reduces with increasing the number of epochs. Once the curve is much flat or there is not much improvement, it is better to stop. In other words, it is important to set the number of epochs as low as possible until the curve is much flat, and then terminate training it.

4.1 Evaluation Metrics

In order to evaluate of the proposed approach with other approaches, we will use four useful metrics: Mean Square Error (MSE) Root Mean Square Error (RMSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE).

$$MSE = \frac{1}{n}\sum_{1}^{n}(Y_i - \overline{Y}_i)^2 \tag{7}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{1}^{n}(Y_i - \overline{Y}_i)^2} \tag{8}$$

$$MAE = \frac{1}{n}\sum_{1}^{n}|Y_i - \overline{Y}_i| \tag{9}$$

$$MAPE = \frac{1}{n}\sum_{1}^{n}\left|\frac{Y_i - \overline{Y}_i}{Y_i}\right| \tag{10}$$

In average, the values of RMSE, MAE, and MAPE of the proposed method are higher for epochs 5, 10, and 15 while comparing epochs 20,25, and 30, as expressed in Table 2. When the number of epochs increase to 25, we observed that almost all values are lower than epochs. Also, these values are very close when the epochs are increased to 30. Therefore, we chosen the epochs 25 to keep the values from the proposed approach. Secondly, the accuracy is 94.49% for epoch 10, however it increases to about 95.87% while the number of epochs goes up to 25 shown in Table 2.

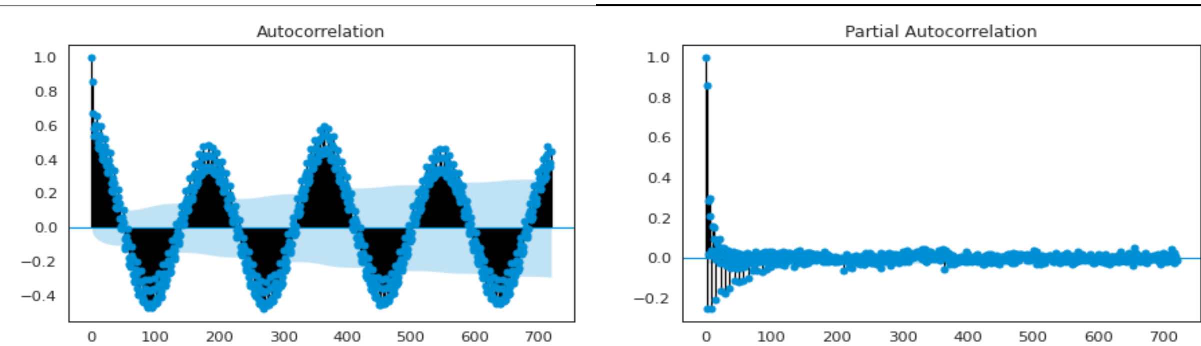
4.2 Discussion

In general, training loss is lower than validation loss. However, we observed higher training loss while starting to train the model in Fig. 8. This loss is decreasing in terms of the number of epochs. Most cases, training loss intersects the validation loss. However, validation losses are fluctuated after each epochs, and another losses are usually decreasing from one epoch to another epoch.

When we have chosen 5, 10, or 15 epochs, both losses are not close at the end of training (Fig. 8(a-c)). Interestingly, these losses are overlapped at the end of training (Fig. 8(d-f)). During training the proposed deep neural network, applying regularization helps it to have higher validation accuracy. In most cases, Regularization is responsible to increase validation/testing accuracy through decrementing training accuracy. As a result, validation loss is observed lower than the training loss in some cases. In order to avoid it, we have used dropout(0.2) in the simulation code. The curves of training and validation loss look more similar while factoring to validation loss in regularization. From the plotting of training and validation loss, overfitting can be visualized.

5 Conclusions

This approach has used the LSTM method which extract correlations spontaneously and other information from time series dataset. We also studied the hidden patterns of time series data using heat-map, autocorrelation, and partial correlation. This work also focused how epochs affect prediction of power



(a) (b)

Fig. 4(a) Autocorrelation and (b) partial autocorrelation of the observations for a electric power consumption.

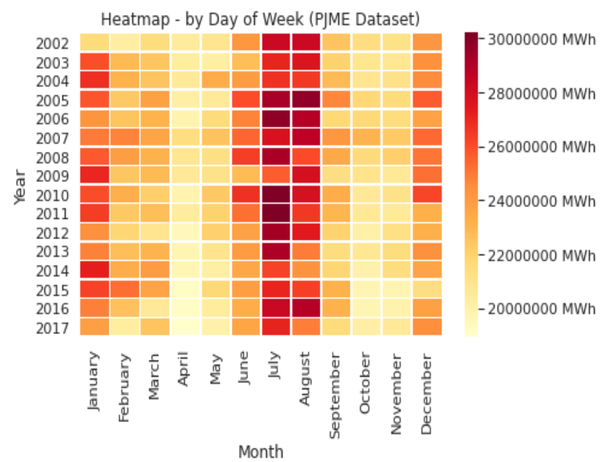


Fig. 5 Heatmap for a building electric power consumption

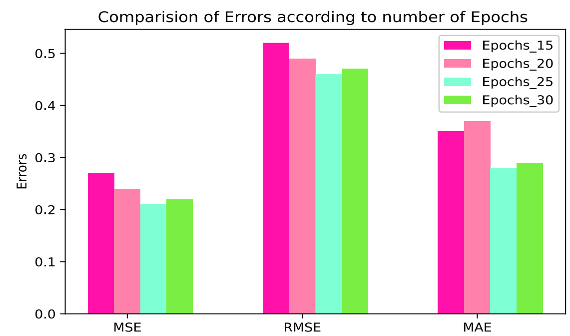
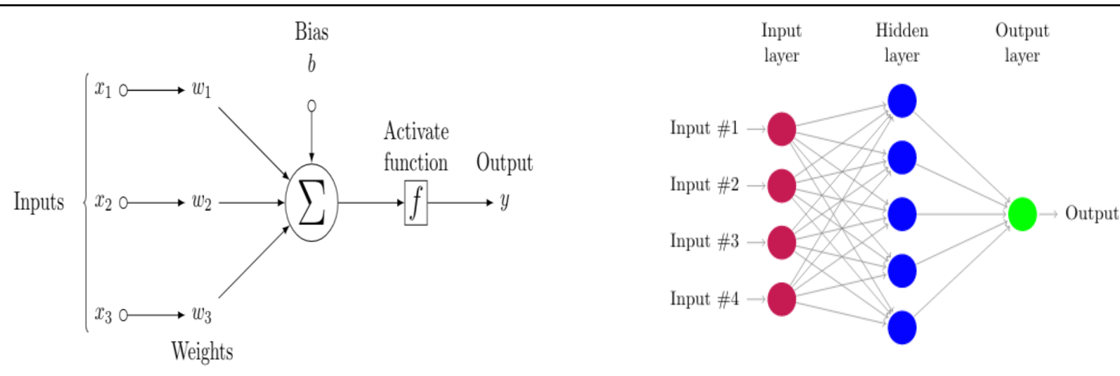


Fig. 7 Performance of the approaches according to the different number of epochs

consumption from time series data through internal analysis of LSTM. The proposed network predicts the consumption patterns through avoiding over-fitting. Also, this work offers important information about how the variation of number of epochs affects the training loss and validation loss as well as accuracy based on the size of the model and the pattern of data of the dataset. There is still over-fitting and delay in predicting patterns of dataset.

Table 2. Errors based on number of epochs

Epoch	MSE	RMSE	MAE	MAPE	ACC.
5	0.37	0.61	0.125	28.49	0.9423
10	0.32	0.56	0.119	32.33	0.9449
15	0.27	0.52	0.35	32.51	0.9506
20	0.24	0.49	0.37	33.14	0.9573
25	**0.21**	**0.46**	**0.28**	**32.87**	**0.9587**
30	0.22	0.47	0.29	33.17	0.9585



(a) (b)

Fig. 6 (a) Basic diagram of neural network **(b)** A typical two-layer neural network

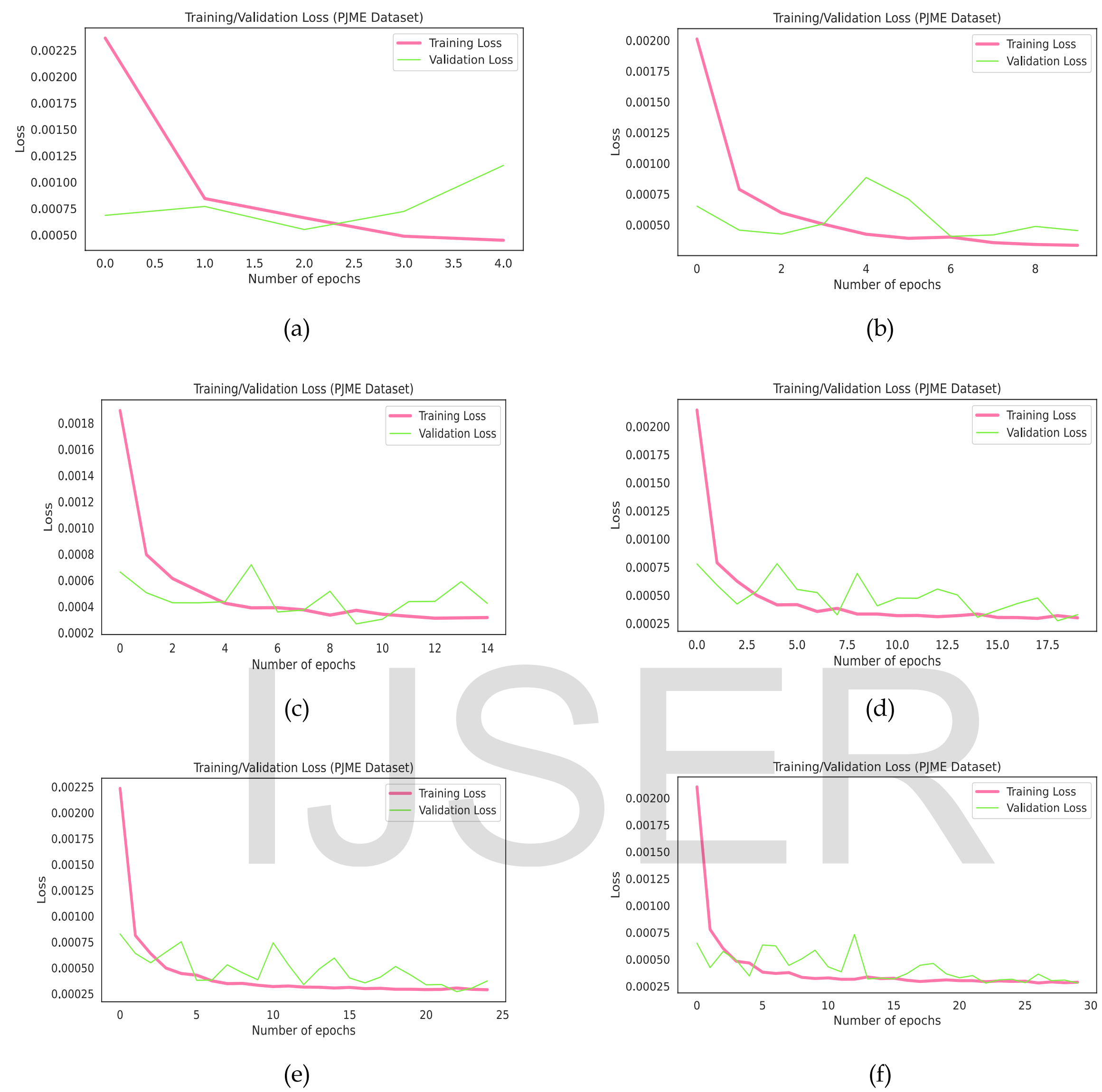


Fig. 8 shows the training losses/validation losses in terms of different number of epochs: 5, 10, 15, 20, 25, and 30

References

[1] Li, C.; Ding, Z.; Zhao, D.; Yi, J.; Zhang, G. Building energy consumption prediction: An extreme deep learning approach. Energies 2017, 10, 1525

[2] Oliveira, E.M.; Oliveira, F.L. Forecasting mid-long term electric energy consumption through bagging ARIMA and exponential smoothing methods. Energy 2018, 144, 776–788.

[3] Krishnan, M.; Sakthivel, R.; Vishnuvarthan, R. Neural network based optimization approach for energy demand prediction in smart grid. Neurocomputing 2018, 273, 199–208.

[4] Fayaz, M.; Kim, D. A Prediction Methodology of Energy Consumption Based on Deep Extreme Learning Machine and Comparative Analysis in Residential Buildings. Electronics 2018, 7, 222.

[5] Divina, F.; Torres, M.G.; Goméz Vela, F.A.; Noguera, J.L.V. A Comparative Consumption Prediction in Smart Buildings. Energies 2019, 12, 1934.

[6] Amarasinghe, K.;Wijayasekara, D.; Carey, H.; Manic, M.; He, D.; Chen,W.-P. Artificial neural networks based thermal energy storage control for buildings. In Proceedings of the IECON 2015 41st Annual Conference of the IEEE Industrial Electronics Society, Yokohama, Japan, 9–12 November 2015.

[7] N. Amjady, "Short-term hourly load forecasting using time-series modeling with peak load estimation capability," IEEE Trans. Power Syst., vol. 16, no. 3, pp. 498–505, Aug. 2001.

[8] N. Mohan, K. Soman, and S. S. Kumar, "A data-driven strategy for shortterm electric load forecasting using dynamic mode decomposition model," Appl. Energy, vol. 232, pp. 229–244, Dec. 2018.

[9] H. Choi, S. Ryu, and H. Kim, "Short-term load forecasting based on ResNet and LSTM," in Proc. IEEE Int. Conf. Commun., Control, Comput. Technol.

Smart Grids, Oct. 2018, pp. 1–6.

[10] W. Shen, V. Babushkin, Z. Aung, and W. L. Woon, ''An ensemble model for day-ahead electricity demand time series forecasting,'' in Proc. 4th Int. Conf. Future Energy Syst., May 2013, pp. 51–62.

[11] W. Kong, Z. Y. Dong, Y. Jia, D. J. Hill, Y. Xu, and Y. Zhang, ''Short-term residential load forecasting based on LSTM recurrent neural network,'' IEEE Trans. Smart Grid, vol. 10, no. 1, pp. 841–851, Sep. 2017.

[12] Bianco, V., Manca, O., Nardini, S.: Electricity consumption forecasting in Italy using linear regression models. Energy. 34(9), 1413-1421 (2009).

[13] Chen, Y.; Xu, P.; Chu, Y.; Li, W.; Wu, Y.; Ni, L.; Bao, Y.; Wang, K. Short-term electrical load forecasting using the Support Vector Regression (SVR) model to calculate the demand response baseline for office buildings. Appl. Energy 2017, 195, 659–670.

[14] Barak, S.,Sadegh, S.S.: Forecasting energy consumption using ensemble ARIMA– ANFIS hybrid algorithm. International Journal of Electrical Power & Energy Systems. 82, 92-104 (2016).

[15] Darbellay, G.A.,Slama, M.: Forecasting the short-term demand for electricity: Do neural networks stand a better chance? International Journal of Forecasting. 16(1), 71- 83 (2000)

[16] Chou, J.-S.,Tran, D.-S.: Forecasting energy consumption time series using machine learning techniques based on usage patterns of residential householders. Energy. 165, 709-726 (2018).

[17] Ahmad, A., Hassan, M., Abdullah, M., Rahman, H., Hussin, F., Abdullah, H., Saidur, R.: A review on applications of ANN and SVM for building electrical energy consumption forecasting. Renewable and Sustainable Energy Reviews. 33, 102-109 (2014).

[18] Pombeiro, H.; Santos, R.; Carreira, P.; Silva, C.; Sousa, J.M. Comparative assessment of low-complexity models to predict electricity consumption in an institutional building: Linear regression vs. fuzzy modeling vs. neural networks. Energy Build. 2017, 146, 141–151.

[19] Boukoros, S., Nugaliyadde, A., Marnerides, A., Vassilakis, C., Koutsakis, P., Wong, K.W.: Modeling server workloads for campus email traffic using recurrent neural networks. In: International Conference on Neural Information Processing, pp. 57-66. Springer, (2017).