Electricity Load Forecasting in Thailand Using Deep Learning Models

Pyae P. Phyo[1], Chawalit Jeenanunta[1], and Kiyota Hashimoto[2]
[1] Sirindhorn Int. Inst. of Technology, Thammasat University, Pathum Thani, Thailand
[2] Prince of Songkla University, ESSAND, Phuket, Thailand
Email: {pyaepyaephyo1979; kiyota.hashimoto}@gmail.com; chawalit@siit.tu.ac.th

***Abstract*—The objective of this research is to improve the short-term load forecasting accuracy using deep learning models such as long short-term memory (LSTM) and deep belief network (DBN). The required historical data is provided by Electricity Generating Authority of Thailand (EGAT). Long short-term memory model which can learn to store time series data in memory and solve long dependencies problems and deep belief network model are investigated to overcome back propagation problems in the network. The proposed models are trained and tested using the cleaned data during the period of January 2016 to January 2017 by smoothing the raw data. Mean absolute percentage error (MAPE) and root mean square error (RMSE) are used to measure the forecasting accuracy. In this research, the results generated by the LSTM model are compared with those of the DBN model. The results show that the LSTM model execute higher accuracy performance than the DBN model.**

***Index Terms*—short-term electricity load forecasting, long short-term memory, deep belief network, mean absolute percentage error, root mean square error.**

I. Introduction

Electricity manufacturing is a well developing industry in the current era with major consideration on the equilibrium of energy supply and load consumption. Intense research studies in this industry ensure great development in the country. Load forecasting plays a vital role in electric power industry. It can be divided in to three main sectors as short, medium and long depending on the duration of load forecasting with respect to a day, month and a year. Load reading is taken every hour or in a time period of 30 minutes to record daily load forecasting. This short-term load forecasting is highly accurate and will improve the availability of electricity if it is taken into consideration. This not only helps in reducing the generating and operating costs of the industry but accesses the security of the power system and performs short-term scheduling functions.

Forecasting techniques that are applied on load forecasting research studies can be classified into two major categories as traditional statistical models and artificial intelligent models. Traditional statistical models such as moving average [1], stochastic time series models [2], exponential smoothing [3], and regression analysis [4] have produced reliable results whereas artificial intelligence techniques including support vector machines [5], fuzzy time series [6], artificial neural networks [7] and hybrid model [8] produce effective results. Neural network outstands the other artificial models for nonlinear time series problems with its features [9].

However, back propagation method reduces the efficiency of neural network method due to the presence of multiple hidden layers [10]. In addition, the runtime of the algorithm is high with possibilities for poor local minimum and slow convergence because of using random initialization of the parameters [11]. The inputs and outputs of the models are independent in a simple neural network. If a data contains continuous information, the network fails to memorize along the training process [12]. Thus, it is important to include features which allow carrying respective memory of data, which is introduced via addition of long short-term memory [13].

The outline of the paper is as follows. In Section 2, the details of description about the long short-term memory model and deep belief network model are described. In Section 3, the design of experiments of this research is discussed. The key results of short-term electricity forecasting in Thailand obtained by using all proposed models with related discussion are presented in Section 4. The paper is concluded in Section 5 with remarks.

II. Methodology

A. Long Short-Term Memory

Long short-term memory (LSTM) was first proposed by Hochreiter & Schmidhuber and has been modified by many other researchers [14]. The goal of this study is to develop a forecasting system of electricity balance between supply and demand using the LSTM based RNN technique and evaluate the method by comparing it to other techniques. The overview of the process in the system is shown in Fig. 1. Firstly, the system is loaded with a dataset consisting eighteen input variables. The data is normalized by using the min-max scaling method. The data is then divided into the training and testing sets. Next, a LSTM network is constructed and trained. After that, the trained LSTM network is used for prediction.

Manuscript received September 25, 2018; revised November 8, 2018; accepted January 15, 2019.
Corresponding author: Pyae P. Phyo (email: pyaepyaephyo1979@gmail.com)

doi: 10.18178/ijeetc.8.4.221-225

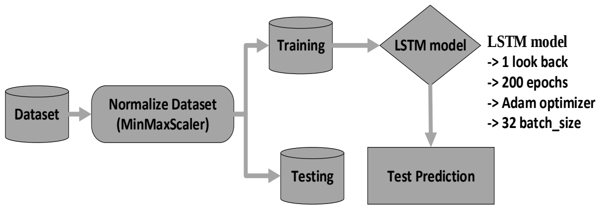

Fig. 1. A process of electricity load forecasting for LSTM model.

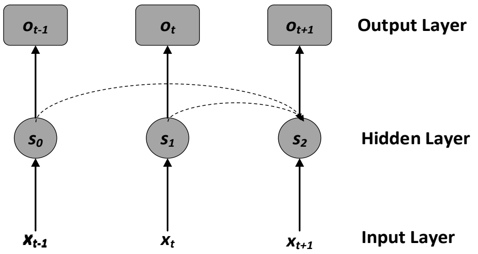

Fig. 2. A working process of LSTM networks.

A powerful type of neural network designed to handle sequence dependence is the recurrent neural network. In this research, the LSTM network is a recurrent neural network that is trained by using backpropagation through time to overcome the vanishing gradient problem. First of all, the eighteen input variables are imported into the LSTM network. In the LSTM network, the input units are fully connected to a hidden layer. The cell outputs are fully connected to the cell inputs, to all gates, and to the output units. During the model fitting, mean squared error is used as loss function for optimizing the parameters of the LSTM model, ReLU (rectified linear units) activation function, and 200 epochs, and Adam is used as the optimization algorithm for the loss function with the batch size of 32.

The electric load data is considered as time-series data because it is recorded every 30 minutes. In time-series data, the data collected at a time t naturally relates to the data at the previous adjacent time (i.e., at the time t–1). A LSTM network deals with this characteristic by using hidden-state information obtained during the network training with the data at the time $t-1$ to train the network with the data at the time t. Fig. 2 illustrates how the LSTM network works, where x_{t-1}, x_t, and x_{t+1} are the input at times $t-1$, t, and t+1, respectively, s_0, s_1, and s_2 are hidden-state information of hidden layers, and o_{t-1}, o_t, and o_{t+1} are the output at times $t-1$, t, and t+1, respectively.

B. Deep Belief Network

Deep belief network (DBN) is a popular and rewarding technique in application since the last decade. It reconstructs the inputs with consideration to probability and includes various learning modules which are less complex [15]. This research study uses DBN which includes a pre-trained restricted Boltzmann machine (RBM) model in each layer of both training and testing data. This unique model has the ability to learn probability distribution over a range of set inputs [16]. In addition, it has a single layer of hidden units with no interconnections and thus composed of symmetrical connections to visible layers of units with no directionality. However, restricted Boltzmann machine requires the formation of bipartite graph with their neurons. The primary goal of RBM application is that with conditionally independent states of hidden units, there are no connections among hidden units.

In this research, the DBN model is applied to compare the results with the LSTM model. Fig. 3 indicates the workflow of the whole process. Firstly, the system is loaded with a dataset consisting of eighteen input variables and the input data is normalized by using the min-max scaling method. After normalizing the data, it is separated into training and testing sets. Finally, a supervised DBN regression model is constructed and trained to predict the testing sets. The DBN model uses ten hidden layers, ten epochs, hundred iterations, ReLU activation functions and a batch size of fifty in the training process.

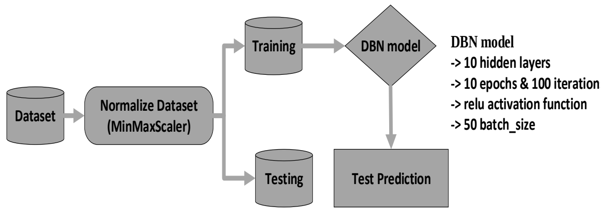

Fig. 3. A process of electricity load forecasting for DBN model.

III. Design of Experiments

The Electricity Generating Authority of Thailand (EGAT) collects the load data from five different regions such as Central, Bangkok and Metropolitan, South, North and North-East regions in Thailand. The collected load data has been recorded every 30 minutes from 2007 to 2017. However, this research uses the net peak load for the whole country from 2016 to 2017 to predict daily load demand.

A. Data Cleaning

The historical data need to be smoothed because there are many missing values, and outliers in the original raw data. If these outliers are included in the training data, the accuracy performance of load predictions would be lower [17]. In order to filter and smooth the raw data, we use local regression filtering technique that uses the method of regression analysis. It can be classified into four types: the lowess (locally weighted scatterplot smoothing) local regression which uses the method of linear regression analysis, the loess (locally estimated scatterplot smoothing) local regression which uses the method of polynomial square regression analysis, rlowess (robust locally weighted scatterplot smoothing) local regression and rloess (robust locally estimated scatterplot smoothing) local regression which are the robust functions can be used to get rid of outlier values. If outliers are present in the dataset, robust lowess/loess (rlowess, rloess) procedure is used to overcome the problem of distorted values. Among them, the loess local regression filtering technique is applied in this research.

The filtering technique fits a local regression function to the data within a chosen neighborhood of data points. A chosen neighborhood which is also known as a smoothing parameter (0<smooth<=1) is specified by percentage of data points. The larger the smoothing parameter, the smoother the graphed function. For calculating smoothed values, this filtering technique will specify weight for every data point in the selected window by using the regression weight function:

$$w_i = \left(1 - \left|\frac{x - x_i}{d(x)}\right|^3\right)^3 \tag{1}$$

Once the regression function values are calculated with flexible weights and polynomial degree, rloess fit is complete.

Equation (1) indicates that w_i is the regression weight of points i, x is the predictor value associated with the response value to be smooth, x_i are the nearest neighbours of x defined by the selected window, and $d(x)$ is the distance along the abscissa from x to the most distant predictor value within the selected window. The relationship between the original data and the smoothed data is shown in Fig. 4.

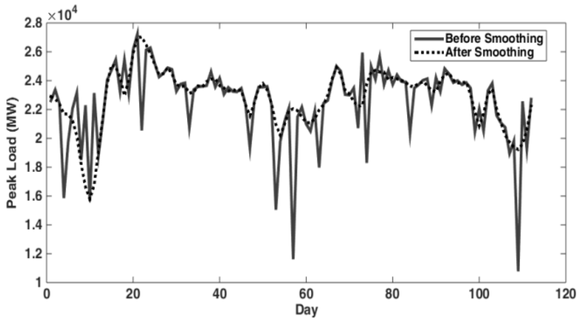

Fig. 4. Relationship between original data and smoothed data.

B. Data Segmentation

When modelling the networks for forecasting purposes, one of the most important steps in order to obtain good results is the selection of the input variables [18], [19]. There are two approaches when forecasting time series variables and solving regression problems. Time series forecasting is either deduced with the use of only the previous x value of the variable or with unique input variables [20]. However with selection of unique input variables, points from the historical values in the training set and other correlated variables to the testing dataset are considered [21]. The load data is organized depending on the days. Forecasting for a specific day considers testing and training data of the load from the same day which generates seven groups. The testing and training process uses walk-forward testing routine. A dataset is tested using 52 training datasets and thus the testing data is sliding forward throughout as shown in Fig. 5.

In this study, there are eighteen input variables to train and test LSTM model and DBN model. They are regarded as load demand at time t on day (d−7) as $L_t d$-7), load demand at (t−2) on day (d−1) as $L_{t-2}(d\text{-}1)$, load demand at (t−1) on day (d−1) as L_{t-1}(d-1), load demand at t on day (d-1) as $L_t(d\text{-}1)$, temperature at t on day day (d-1) as $T_t(d\text{-}1)$, monthly seasonal index (SI) which is monthly load divided by yearly load, and January to December as 12 inputs which have values 0 and 1 (e.g, if the load demand in month of January is considered for forecasting, Jan is given 1 dummy variable while the other months are given 0) to forecast the forecasted load at time t on day d as $F_t(d)$. In this study, both models use 2016 as training datasets and 2017 as testing datasets. The example of data arrangement for both models is shown in Table I.

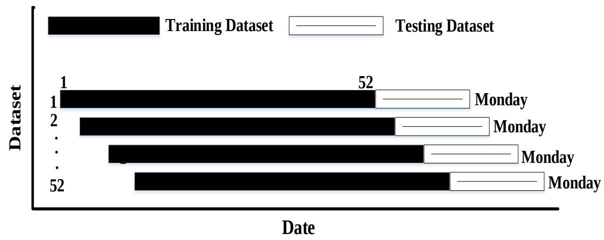

Fig. 5. Walk forward testing routine using sample dataset.

TABLE I. SAME DAY TRAINING DATA ARRANGEMENT FOR TESTING TARGET 1ST JAN 2017

		Inputs									Target
Training Dataset	No.	$L_t(d-7)$	$L_{t-2}(d-1)$	$L_{t-1}(d-1)$	$L_t(d-1)$	$T_t(d-1)$	SI	Jan	…	Dec	$F_t(d)$
	1	3/1/16 (Sun)	9/1/16 (Sat)	9/1/16 (Sat)	9/1/16 (Sat)	9/1/16 (Sat)	-	1	-	0	10/1/16 (Sun)
	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
	52	11/12/16 (Sun)	17/12/16 (Sat)	17/12/16 (Sat)	17/12/16 (Sat)	17/12/16 (Sun)	-	0	-	1	18/12/16 (Sun)
		Inputs									Output
Testing Dataset	No.	$L_t(d-7)$	$L_{t-2}(d-1)$	$L_{t-1}(d-1)$	$L_t(d-1)$	$T_t(d-1)$	SI	Jan	…	Dec	$F_t(d)$
	1	25/12/16 (Sun)	31/12/16 (Sat)	31/12/16 (Sat)	31/12/16 (Sat)	31/12/16 (Sat)	-	1	-	0	1/1/17 (Sun)

IV. RESULT AND DISCUSSION

In this research, the mean absolute percentage error at day d (MAPE$_d$):

$$\text{MAPE}_d = \frac{1}{t}\sum_{t=1}^{48}\frac{L_t(d) - F_t(d)}{L_t(d)} \times 100\% \tag{2}$$

is used as an accuracy measurement which indicates how many units of the forecasting demand is deviated from the actual demand. Moreover, the root mean square error (RMSE$_d$) given in (3) is also applied to measure accuracy of forecasting result:

$$\text{RMSE}_d = \sqrt{\frac{\sum_{t=1}^{48}(L_t(d) - F_t(d))^2}{48}} \tag{3}$$

where $L_t(d)$ is the cleaned load at period t for day d, $F_t(d)$ is the forecasted load at period t for day d, and t = 1, 2, 3, ⋯, 48 periods.

In the following, we compare all the $MAPE_d$ and $RMSE_d$ results executed from both models.

The average monthly MAPE and RMSE for LSTM model and DBN model are depicted in Table II. When we train both LSTM and DBN on load and temperature data, both models provide sufficient performance for all months. According to the results in Table II, LSTM outperforms DBN because the collected data shows similarity with time series data collected for every 30 minutes. Though there are similar values for some months from both the models, the significant difference during certain months ensure the selection of LSTM is much better than the DBN model. December shows the highest MAPE and RMSE since there are many holidays and the number of tourists are high. This affect January too because of New Year celebrations and increased number of organized public events.

TABLE II. AVERAGE MONTHLY MAPE AND RMSE FOR LSTM AND DBN IN 2017

Months	MAPE		RMSE	
	LSTM	DBN	LSTM	DBN
Jan	4.7296	5.8013	134.4252866	167.4041466
Feb	3.8164	5.3027	104.3676409	161.1015793
Mar	4.0269	4.5721	129.0100475	151.5026312
Apr	3.3238	5.8958	111.7291947	194.7090033
May	4.2867	7.4171	138.6456158	242.3300712
Jun	2.8422	3.4449	92.85113732	112.6369005
Jul	2.5922	6.0459	81.85598854	191.3437764
Aug	2.5601	4.1239	82.83007808	134.9175077
Sep	2.7629	4.4978	89.97302991	147.9916277
Oct	2.5263	4.8453	81.8529626	151.4246636
Nov	3.3853	4.1156	100.5808375	126.5700084
Dec	8.6518	8.1478	236.9885136	226.4363737
Yearly Average	3.7920	5.3509	115.4259	167.3640

March, April and May shows higher MAPE and RMSE due to the highest temperature in Thailand. The proposed LSTM model clearly shows that the errors are minimum and the values are well presented. However, the month of March and December, shows similarity with results from DBN model along with the presence of holidays.

TABLE III. AVERAGE MAPE AND RMSE FOR EACH DAY FOR LSTM AND DBN

Day	MAPE		RMSE	
	LSTM	DBN	LSTM	DBN
Monday	5.9692	5.3563	189.2100	166.4573
Tuesday	3.9260	5.4030	120.7098	171.0818
Wednesday	4.0965	4.9781	125.2023	159.6956
Thursday	3.1062	5.3870	94.3162	172.2459
Friday	3.1527	5.1281	96.3141	165.3722
Saturday	2.9949	6.0507	91.5656	192.6780
Sunday	3.3604	5.2243	93.0119	146.4636

The average daily MAPE and RMSE for LSTM model and DBN model are shown in Table III. Monday shows higher MAPE and RMSE due to the input of the previous day. In general, weekend load patterns show a low value comparing to the weekday load patterns and thus, Sunday being an input for the load pattern of Monday results in higher MAPE and RMSE. Weekday load fluctuation of LSTM model shows similar results while weekend load fluctuations of LSTM model shows specific pattern. DBN model specifies the weekday load fluctuation pattern and weekend load fluctuation pattern as two distinct groups with similar features too. LSTM model outperforms DBN model on the average daily MAPE and RMSE results except Monday.

According to the executed results, the day with the least MAPE and RMSE and the day with the highest MAPE and RMSE are chosen to show the variations in load pattern within the day using LSTM model. Fig. 6 and Fig. 7 indicate the load pattern variation between actual load and forecasted load, respectively. LSTM model can predict the least MAPE and RMSE on 27th June, 2017 which is 0.8852 and the highest MAPE and RMSE on of 5th January, 2017 which is 21.3393. The highest MAPE and RMSE day is assumed as a bridging holiday because there are three holidays before the specific day and a weekend after. However, the lowest MAPE and RMSE are generated on 27th June, 2017 with summer vacation and no public holidays in June.

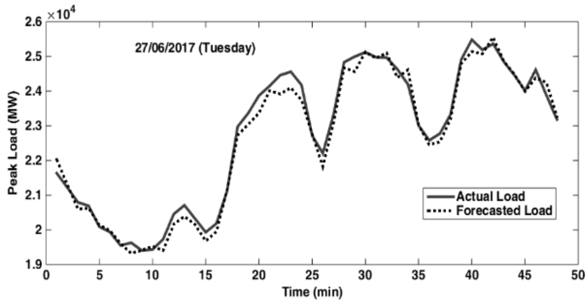

Fig. 6. Load pattern variation with minimum MAPE and RMSE using LSTM in 2017.

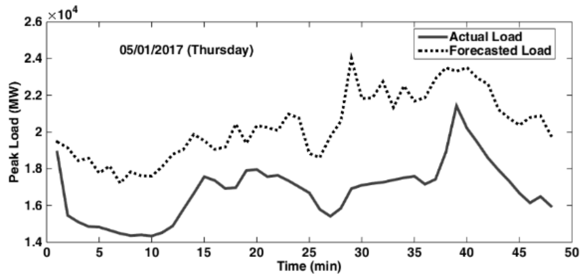

Fig. 7. Load pattern variation with maximum MAPE and RMSE using LSTM in 2017.

V. CONCLUSION

In this research, we proposed the long short-term memory to learn how to store time series data in memory and to solve long dependencies problems for short term load forecasting. In addition, deep belief network which is an unsupervised learning method is proposed to avoid random initialization of parameters in the network. The historical data is provided by the Electricity Generating Authority of Thailand (EGAT). LSTM model and DBN model are trained and tested by using cleaned data from 2016 to 2017 to forecast daily load demand in 2017. Eighteen input variables are selected to train the LSTM and DBN models. After executing the outcomes, the results between the proposed two models are compared.

The proposed LSTM model performs with higher accuracy than the DBN model.

ACKNOWLEDGMENT

This research is partially supported by the Logistics and Supply Chain Systems Engineering Research Unit (LogEn), Sirindhorn International Institute of Technology (SIIT), Thammasat University (TU), Thailand. Data used in this research is provided by EGAT. Therefore, we acknowledge their support for completing this research.

REFERENCES

[1] A. K. Singh, S. K. Ibraheem, M. Muazzam, and D. K. Chaturvedi, "An overview of electricity demand forecasting techniques," *Network and Complex Systems*, vol. 3, no. 3, pp. 38-48, 2013.

[2] K. Liu, S. Subbarayan, R. R. Shoults, M. T. Manry, C. Kwan, F. I. Lewis, and J. Naccarino, "Comparison of very short-term load forecasting techniques," *IEEE Trans. on Power Systems*, vol. 11, no. 2, pp. 877-882, 1996.

[3] W. R. Christiaanse, "Short-term load forecasting using general exponential smoothing," *IEEE Trans. on Power Apparatus and Systems*, vol. PAS-90, no. 2, pp. 900-911, 1971.

[4] D. Papalexopoulos and T. C. Hesterberg, "A regression-based approach to short-term system load forecasting," *IEEE Trans. on Power Systems*, vol. 5, no. 4, pp. 1535-1547, 1990.

[5] M. Mohandes, "Support vector machines for short‐term electrical load forecasting," *Int. Journal of Energy Research*, vol. 26, no. 4, pp. 335-345, 2002.

[6] A. Azadeh, M. Saberi, and A. Gitiforouz, "An integrated simulation-based fuzzy regression-time series algorithm for electricity consumption estimation with non-stationary data," *Journal of the Chinese Institute of Engineers*, vol. 34, no. 8, pp. 1047-1066, 2011.

[7] H. S. Hippert, C. E. Pedreira, and R. C. Souza, "Neural networks for short-term load forecasting: A review and evaluation," *IEEE Trans. on Power Systems*, vol. 16, no. 1, pp. 44-55, 2001.

[8] A. Kavousi-Fard, H. Samet, and F. Marzbani, "A new hybrid modified firefly algorithm and support vector regression model for accurate short term load forecasting," *Expert Systems with Applications*, vol. 41, no. 13, pp. 6047-6056, 2014.

[9] N. I. Sapankevych and R. Sankar, "Time series prediction using support vector machines: A survey," *IEEE Computational Intelligence Magazine*, vol. 4, no. 2, pp. 24-38, 2009.

[10] G. Zhang, B. E. Patuwo, and M. Y. Hu, "Forecasting with artificial neural networks: The state of the art," *Int. Journal of Forecasting*, vol. 14, no. 1, pp. 35-62, 1998.

[11] L. Bottou, "Large-scale machine learning with stochastic gradient descent," in *Proc. COMPSTAT'2010*, Physica-Verlag HD, 2010, pp. 177-186.

[12] P. P. Phyo and C. Jeenanunta, "Electricity load forecasting using deep neural network," *Engineering and Applied Science Research*, June 2018.

[13] A. Narayan and K. W. Hipel, "Long short term memory networks for short-term electric load forecasting," in *Proc. IEEE Int. Conf.* on *Systems, Man, and Cybernetics (SMC)*, October, 2017, pp. 2573-2578.

[14] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, vol. 9, no. 8, pp. 1735-1780, 1997.

[15] G. E. Hinton, S. Osindero, and Y. W. Teh, "A fast learning algorithm for deep belief nets," *Neural Computation*, vol. 18, no. 7, pp. 1527-1554, 2006.

[16] A. Dedinec, S. Filiposka, A. Dedinec, and L. Kocarev, "Deep belief network based electricity load forecasting: An analysis of Macedonian case," *Energy*, vol. 115, pp. 1688-1700, 2016.

[17] C. Jeenanunta, K. D. Abeyrathna, M. S. Dilhani, S. W. Hnin, and P. P. Phyo, "Time series outlier detection for short-term electricity load demand forecasting," *Int. Scientific Journal of Engineering and Technology*, vol. 2, no. 1, pp. 37-49, 2017.

[18] P. Wang, B. Liu, and T. Hong, "Electric load forecasting with recency effect: A big data approach," *Int. Journal of Forecasting*, vol. 32, no. 3, pp. 585-597, 2016.

[19] P. Lusis, K. R. Khalilpour, L. Andrew, and A. Liebman, "Short-term residential load forecasting: Impact of calendar effects and forecast granularity," *Applied Energy*, vol. 205, pp. 654-669, 2017.

[20] C. Bergmeir, R. J. Hyndman, and B. Koo, "A note on the validity of cross-validation for evaluating autoregressive time series prediction," *Computational Statistics & Data Analysis*, vol. 120, pp.70-83, 2018.

[21] M. Rana and I. Koprinska, "Forecasting electricity load with advanced wavelet neural networks," *Neurocomputing*, vol. 182, pp. 118-132, 2016.

Pyae Pyae Phyo is currently a doctoral student in Management Technology (MT) at Sirindhorn International Institute of Technology (SIIT). She graduated Bachelor of Technology (B.Tech) and Bachelor of Engineering (B.E) majoring in electrical power engineering from the Technological University Mandalay, Myanmar in 2013 and 2014 respectively. In 2015 and 2017, she received the Excellent Foreign Student (EFS) scholarship award from Sirindhorn International Institute of Technology (SIIT), Thammasat University (TU), Thailand. She finished her master of Science (M.Sc) majoring in Engineering and Management in 2017. Her current research interests include artificial intelligence, machine learning, deep learning, deep neural networks, recurrent neural networks, long short-term memory, deep belief networks, regression and forecasting.

Chawalit Jeenanunta is an associate professor of School of Management Technology (MT), Sirindhorn International Institute of Technology, Thammasat University, Thailand. He received a B.S. degree in Mathematics and Computer Science, and M.Sc. in Management Science from University of Maryland and he received his Ph.D. in Industrial and Systems Engineering from Virginia Polytechnic Institute and State University. His Research interests are in area of applications of operations research, simulation, large-scaled optimization and supply chain management.

Kiyota Hashimoto, D.Eng. works at ESSAND, Prince of Songkla University Phuket Campus, Thailand. He obtained his B.A. and M.A. from Kyoto University, Japan, and Doctor of Engineering from Nara Institute of Science and Technology, Japan. After he worked for 20 years at Japanese universities including Osaka Prefecture University as a full professor, he moved to Thailand in 2015. His research interest includes natural language processing, text mining, machine learning, and artificial intelligence in general.