# Building Deep Neural Network Model for Short Term Electricity Consumption Forecasting

Widyaning Chandramitasari, Bobby Kurniawan, Shigeru Fujimura
Graduate School of Information, Production, and System
Waseda University
Kitakyushu, Japan
akasari@akane.waseda.jp, b.kurniawan@fuji.waseda.jp, fujimura@waseda.jp

***Abstract*—Electricity consumption forecasting has a main role in the energy supply management system of a power supply company. A power supply company needs to keep the balancing of the electricity demand and supply for their customers. The target is to forecast the electricity consumption in manufacturing company for each 30-minutes in the next day to prevent the lack of electricity supply from a power supply company. Due to this problem, it is the challenge for short term electricity time series consumption forecasting. In this work, we proposed the model of deep learning neural network with approach the combination of Long Short-Term Memory (LSTM) and Feed Forward Neural Network (FFNN) to perform the electricity forecasting. This proposed method (LSTM-FFNN) was implemented in the time-series data of electricity consumption on a manufacturing company. In our experiment, we used LSTM to perform the time-series forecasting by using historical data of electricity consumption, and we performed FFNN along with additional information which represented by one-hot encoding shape to increase the forecasting performance. Experimental results showed that LSTM-FFNN gave the better result as we compared with our baseline which is the original LSTM and Moving Average (MA) based on the Root Mean Squared Error (RMSE) score.**

***Keywords—deep neural network; long short-term memory (LSTM); electricity forecasting; short term electricity forecasting; feed-forward neural network***

## I. INTRODUCTION

Electricity forecasting has an important role for the energy management system of power supply company. It is very necessary in the power plant company for having the good decision and planning in case of electricity supply. Some customers of power supply company can be local industries, home, public facilities, manufacturing company, etc. The major problem of power supply company is to keep the balancing of demand and supply for the customer. Therefore, it has been difficult to store the electricity when they have been produced a lot, but the lack of electricity supply also caused the disadvantage for customer and supplier. In the case of manufacturing company, lack of electricity supply can be inhibited the process of production. In this research, our concern is to help the small power supply company which provides the electricity supply for manufacturing company. This power supply company needs to forecast the electricity consumption of manufacturing company for each 30minutes. In this case, this problem belongs to short term electricity forecasting.

Some researchers have been proposed some methods for solving the short-term electricity forecasting. Taylor and Mcsharry were applied approach of Auto Regressive Integrated Moving Avegarge (ARIMA) [1]. Hipert *et al.* investigated the application of Artificial Neural Network (ANN) [2]. Sanjit and Sarat proposed a novel Recurrent Neural Network (RNN) [3]. Chang et. al. proposed a novel approach of Long Short-term Memory (LSTM) based on RNN [4]. Moreover, combination of some different methods also has been proposed for solving this problem, for example Wan He proposed the combination of RNN and Convolutional Neural Network (CNN) for solving load forecasting problem [5]. These all method has been using the historical load data to perform the load forecasting.

Neural Network (NN) technology is a statistical machine learning method which widely used in various kind of cases. One of them is electricity consumption forecasting. The advantage of NN technology is they can learn from the example data (learning process) and after it has finished, NN able to know the non-linear dependencies. In this research, we tried some hyperparameter settings by using test case, in order to check the performance of the forecasting model. In the test case, we have tried an experiment using single and multiple hidden layer. Dealing with time series historical electricity consumption data, one of NN type which is RNN has their advantages to be used. RNN is the type of NN that can be processed the sequential data. However, RNN has the problem with long-term dependencies and LSTM came to overcome the problem of RNN. The brief introduction of LSTM will be explained in the Section III. The core of LSTM is introducing the idea of self-loops for overcoming the long duration of dependencies so that become the solution of gradient vanishing/exploding which appeared in RNN [6]. Deep Neural Network (DNN) is NN with more than one hidden layer (multiple hidden layers). Therefore, in DNN at least there is more than three layers and it means the process of learning is "deeper". There is the complex connection between input-output and complex pattern of dataset and it can be learned by those multiple layers [7].

978-1-5386-5280-0/18/$31.00 ©2018 IEEE

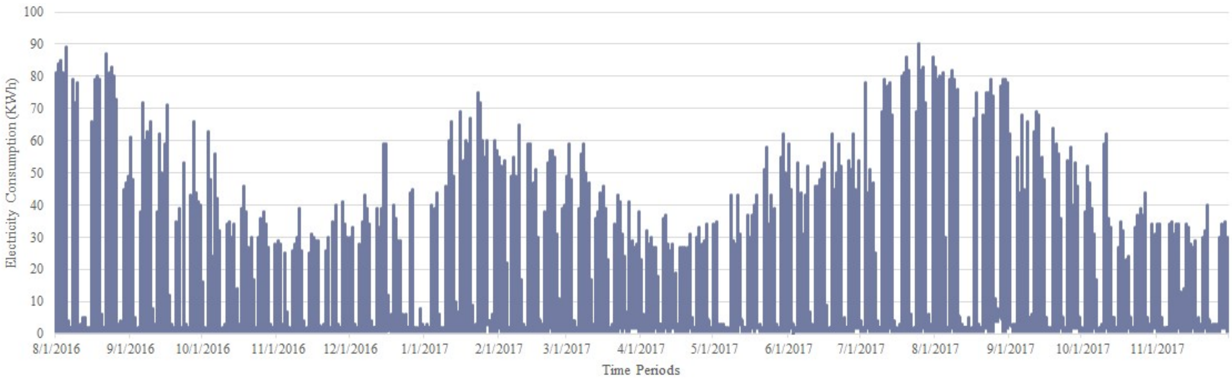


Fig. 1. The overview of electricity consumption dataset (August 2016 – November 2017) with time-interval 30-minutes.

While analyzing the electricity time series data, we can obtain the result contain the trend of the data and seasonal variation for the electricity consumption. Fig. 1 shows the historical data of electricity consumption in manufacturing company with interval 30 minutes. These data provided by one of the small power supply company in Japan which supply the electricity for manufacturing companies. The target of this research is to help the power supply company to prevent the high penalty caused by the high loss of forecasting. Therefore, by using this historical data, the analysis and process of data extraction to the information are needed to produce the good forecasting results.

This paper proposed model by the combination of LSTM and Feed-Forward Neural Network (FFNN) to perform the short-term electricity forecasting. The main features in this proposed model are LSTM perform the time series forecasting and the FFNN will perform the forecasting with the additional information to minimize the loss of forecasting. The result of forecasting derived from this proposed method was compared with the original LSTM.

## II. Feed Forward Neural Network

This section will give the brief introduction about FFNN.

FFNN was the standard type of artificial neural network [8]. The network structure in FFNN as shown in Fig. 2 usually has one or more hidden layer and using the sigmoid activation function. The number of neurons in the input and output layer is given based on the number of input and output variables in the related cases [9].

In this presented research, the numbers of input neuron is four, which each items representing the electricity consumption together with one-hot encoding of days, time-scale, and seasons information. The number of output neuron is one which is the predicted value of electricity consumption. Inside hidden layer, for each neuron there is a main process of FFNN. The process in hidden layer will be shown in the (1) and (2).

$$h_{[i]} = \sigma\left(W_h x_{[i]} + b_1\right) \quad (1)$$

$$out_{[i]} = \sigma\left(W_o h_{[i]} + b_2\right) \quad (2)$$

as shown in equation (1) and (2), the process in FFNN in each step is added the bias with the weighted input and activate it with the activation function.

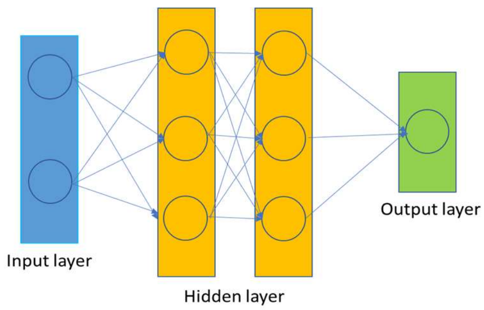


Fig. 2. The model of standard FFNN.

The training process in the neural network is the process of adjusting the weights and biases of a network for minimizing the loss function on the network outputs (predicted and ground-truth), which is in this research is the case of supervised learning (given the set of input and the ground-truth output).

## III. Long Short Term Memory

This section will give the brief introduction about LSTM. LSTM was originally proposed by Hochreiter and Schmidhuber in 1997. LSTM is a special kind of RNN which able to handle the learning of long-term dependencies [6]. This method has designed to overcome the problem of vanishing/exploding gradient which happened in RNN. Both RNN and LSTM are usually using the Backpropagation Through Time (BPTT) [10]. Compared with hidden layer in RNN, LSTM has more complex structure in hidden layer process.

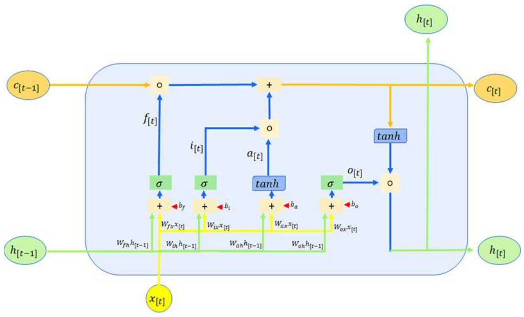


Fig. 3. LSTM cell

LSTM network model able to store the information for long period of time in their memory cell and this model has combination of 4 gates which is: input gate ($i_{[t]}$), input activation gate ($a_{[t]}$), forget gate ($f_{[t]}$) and output gate ($o_{[t]}$). Each gate has their own role inside the process of LSTM network as shown in Fig. 3. LSTM network always store the information on the memory cell state during the process of self-loop.

As shown in Fig. 3 for the LSTM cell, the equation (3) to (6) express the operation inside the single LSTM cell.

$$f_{[t]} = \sigma\left(W_{fx}x_{[t]} + W_{fh}h_{[t-1]} + b_f\right) \tag{3}$$

$$i_{[t]} = \sigma\left(W_{ix}x_{[t]} + W_{ih}h_{[t-1]} + b_i\right) \tag{4}$$

$$a_{[t]} = tanh\left(W_{ax}x_{[t]} + W_{ah}h_{[t-1]} + b_a\right) \tag{5}$$

$$o_{[t]} = \sigma\left(W_{ox}x_{[t]} + W_{oh}h_{[t-1]} + b_o\right) \tag{6}$$

Produce:

$$c_t = c_{[t-1]} \circ f_{[t]} + i_{[t]} \circ a_{[t]} \tag{7}$$

$$h_{[t]} = o_{[t]} \circ tanh(c_t) \tag{8}$$

The role of forget gate (3) in LSTM cell is to help decide what kind of information from the previous state to be remember and forgot in the memory cell. For the input gate has the role to decide what information to be stored in memory cell (4) and input activation gate (5) has the role to decide the update of the information in memory cell. Last, the output gate (6) decides the output for the new hidden state from information of the memory cell. Usually the activation function for forget gate, input gate and output gate is using sigmoid activation function and for input activation gate is using hyperbolic tangent activation function. Equations (7) and (8) show the internal memory cell state and hidden state.

## IV. Methodology

This section will introduce the methodology and the shape of dataset used for training and testing the proposed model in this research.

The objective of methodology in this research is to forecast the electricity consumption in the manufacturing company. Given the historical time series data of electricity consumption, we forecasted the electricity consumption for some time steps in the future by using our proposed method. In this case, we need to forecast the electricity consumption in the next day for each 30minutes (48 time-steps in a day). For the training data, we divided the data based on our case, which is 1-day time horizon and 30minutes for the time resolution. As shown in Fig. 4, the dataset has been divided for each day contains the electricity consumption data for each 30minutes. N represents the number of samples data. While X and Y represent set of input and output.

In this research, for the proposed model we combined the two type of NN components to perform the electricity consumption forecasting. The experiments have been performed by using those divided data. The historical time-series data of electricity consumption is the main input for this model.

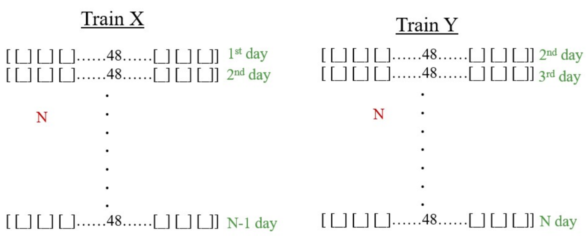


Fig. 4. Data pre-processing for training process

The whole proposed architecture model has shown in Fig. 5. In performing the electricity consumption forecasting, firstly we set LSTM to do the time series forecasting with the historical data. Then, the output from LSTM will be input for FFNN together with additional inputs. The additional inputs are the day, time-scale, and season information which represented by using one-hot encoding.

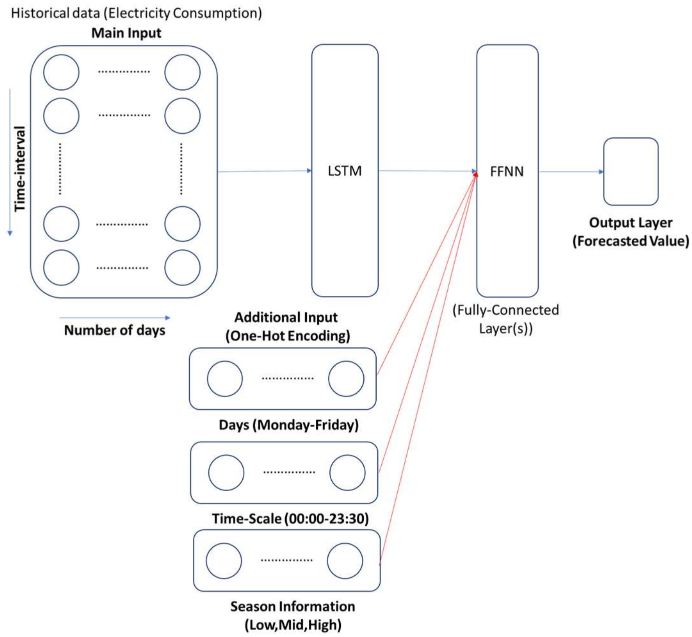


Fig. 5. LSTM-FFNN architecture for electricity consumption forecasting

In this presented work, we proposed three kinds of additional time information. That information is represented by one-hot encoding, they are as follows:

1. **Day information**
   This information represents the day information in each electricity consumption data, from Monday – Sunday.
2. **Time-scale information**
   This information represents the time-scale information in each electricity consumption data, from 00:00-23:30.
3. **Season information**
   In this season information we divided it into 3 seasons, which is **low season** (November and December), **middle season** (January – June, September and October), and **high season** (July and August).

Before performing the electricity consumption forecasting, some of hyperparameter setting needs to be settled. The hyperparameters is including the number of LSTM and FFNN layers, number of neurons in every layer, etc. The model of LSTM and FFNN are shown in the Fig. 6 and Fig. 7.

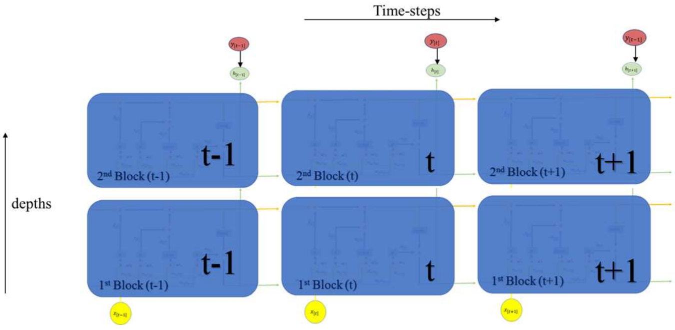


Fig. 6. LSTM model.

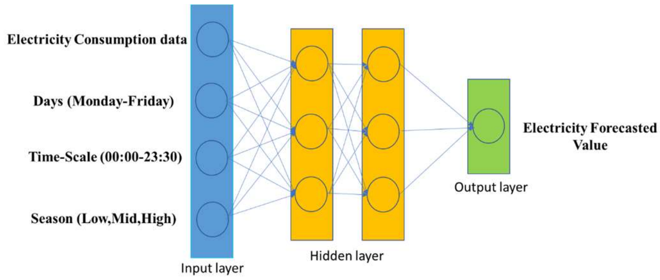


Fig. 7. FFNN model.

For the training process in LSTM we used the BPTT and standard backpropagation in FFNN. In performing the optimization, we used Adaptive Moment Estimation (ADAM) optimizer and Root Mean Square Propagation (RMSProp). During the training process, both weights and biases were updated to minimize the loss function.

## V. Dataset & Experimental Result

This section will introduce about the dataset used for this research experiments and the result of the experiments obtained from the proposed model.

### A. Dataset

The dataset was obtained from one of the small Power Company in Japan which contains the historical data of electricity consumption from one of manufacturing company. The proposed method was implemented on that dataset.

The dataset consists the electricity consumption data for 1 year and 4 months (August 2016 – November 2017) with 30-minutes resolution. Furthermore, for each day contains 48 data of electricity consumption. The detail of the dataset has shown in Table II.

As shown in Table II, we will use the “All Day” data (23376) and the “Weekdays” data (15360). The reason is show the advantage of our proposed model, we were compared the result of the experiment using different amount of data. Our goal is to forecast the electricity consumption for one day ahead each 30-minutes.

### B. Experimental Setting and Result

The proposed model in this paper is LSTM-FFNN. LSTM perform the time series forecasting and the output of LSTM will be going into FFNN model together with the additional information of day, time-scale, and season information. The hyperparameter of the model was tuned based on some test-case as shown in Table I. This test case has been decided to perform the forecasting process to check the forecasting model performance.

TABLE I. Test-Case

| Test Case | LSTM Layer | LSTM Units | FFNN Layer | FFNN Units |
|---|---|---|---|---|
| 1 | 1 | 10 | 2 | 10 |
| 2 | 1 | 50 | 2 | 50 |
| 3 | 2 | 10 | 2 | 10 |
| 4 | 2 | 50 | 2 | 50 |

The performance of the proposed model was trained and tested in comparison with the original LSTM. The reason is because the idea behind the proposed model based on the original LSTM combine with Neural Network model which is FFNN and became LSTM-FFNN. For the evaluation of the performance, we used Root Mean Squared Error (RMSE). The RMSE is defined as shown in (9).

$$RMSE = \sqrt{\frac{1}{n}\sum_{j=1}^{n}\left(y_j - y_j'\right)^2} \tag{9}$$

where $y_j$ represents the ground-truth value and $y_j'$ represents the forecasted value for time-step $j$.

As shown in Fig. 8, the proposed method reached the maximum value in the actual data. This LSTM-FFNN performed better based on the RMSE Score comparison with the Standard LSTM for training and testing process as shown in Table III.

In the forecasting result which has shown in Figs. 8 and Fig. 9, the result was using the test case 4. Fig. 8 shows the forecasting performance of the proposed model and the standard LSTM for one day in each 30-minutes. This experiment was using the “Weekdays” data. This result also has shown that the usage of additional information improved the performance of forecasting. Fig. 9 shows the forecasting performance of the proposed model and the standard LSTM for one day in each 30-minutes using “All Day” data.

Table III shows the comparison of RMSE Score for LSTM and LSTM-FFNN using the different amount of data and based on the test case for the testing process. The result of testing process has shown the LSTM-FFNN gave the better result than original LSTM. It has shown from RMSE Score of LSTM-FFNN is smaller than the RMSE Score of original LSTM. The additional information affected the result of the experiments. It has increased the performance in LSTM-FFNN and gave the better result rather than original LSTM.

TABLE II. DETAIL OF DATASET

| Month | Total Days | Total data (All Day) | Holiday & Weekend | Total data (Holiday & Weekend) | Weekdays | Total data (Weekdays) |
|---|---|---|---|---|---|---|
| August – 2016 | 31 | 1488 | 11 | 528 | 20 | 960 |
| September – 2016 | 30 | 1440 | 10 | 480 | 20 | 960 |
| October – 2016 | 31 | 1488 | 11 | 528 | 20 | 960 |
| November – 2016 | 30 | 1440 | 10 | 480 | 20 | 960 |
| December – 2016 | 31 | 1488 | 13 | 624 | 18 | 864 |
| January – 2017 | 31 | 1488 | 12 | 576 | 19 | 912 |
| February – 2017 | 28 | 1344 | 8 | 384 | 20 | 960 |
| March – 2017 | 31 | 1488 | 9 | 432 | 22 | 1056 |
| April – 2017 | 30 | 1440 | 10 | 480 | 20 | 960 |
| May – 2017 | 31 | 1488 | 11 | 528 | 20 | 960 |
| June – 2017 | 30 | 1440 | 8 | 384 | 22 | 1056 |
| July – 2017 | 31 | 1488 | 11 | 528 | 20 | 960 |
| August – 2017 | 31 | 1488 | 12 | 576 | 19 | 912 |
| September – 2017 | 30 | 1440 | 10 | 480 | 20 | 960 |
| October – 2017 | 31 | 1488 | 10 | 480 | 21 | 1008 |
| November – 2017 | 30 | 1440 | 11 | 528 | 19 | 912 |
| Total | **487** | **23376** | **167** | **8016** | **320** | **15360** |

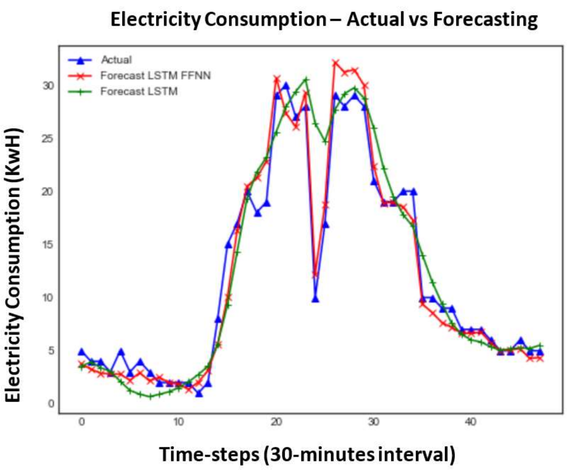


Fig. 8. Electricity consumption forecasting result (Weekdays)

TABLE III. RMSE SCORE FOR TESTING PROCESS

| Test Case | All Day | | Weekdays | |
|---|---|---|---|---|
| | LSTM | LSTM-FFNN | LSTM | LSTM-FFNN |
| 1 | 0.042266 | 0.024037 | 0.035474 | 0.023706 |
| 2 | 0.036668 | 0.032504 | 0.032294 | 0.022658 |
| 3 | 0.045702 | 0.039169 | 0.051316 | 0.027376 |
| 4 | 0.039173 | 0.028984 | 0.037987 | 0.019279 |

Furthermore, based on the presented RMSE Score in the Table III, the result of using “Weekdays” data is better than using “All Day” data in performing the forecasting on weekday.

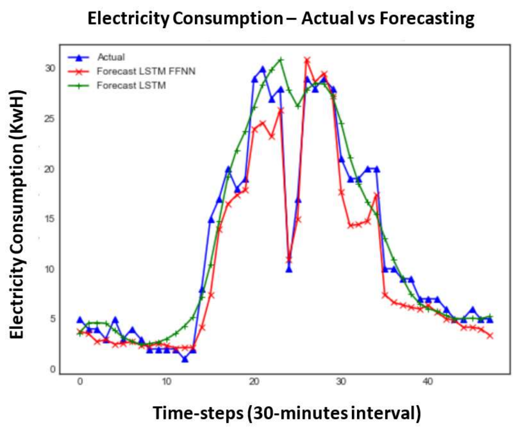


Fig. 9. Electricity consumption forecasting result (All Day)

After we have tried several experiments compared with our baseline which is original LSTM, we also tried to compare the experiment result with our other baseline which is Moving Average (MA). This MA has been used as the forecasting method in this power supply company to fulfill the electricity demand from manufacturing company. In this experiment, we tried the 5-days moving average to forecast the electricity demand for the next day. In this part, we have tried using 1-week testing data for each 30-minutes.

TABLE IV. COMPARISON OF RMSE SCORE (1-WEEK TESTING DATA)

| Dataset | Forecasting Method | | |
|---|---|---|---|
| | LSTM | LSTM-FFNN | MA |
| All Day | 5.459319 | 3.436282 | 8.906759 |
| Weekdays | 3.337392 | 2.519170 | 2.677063 |

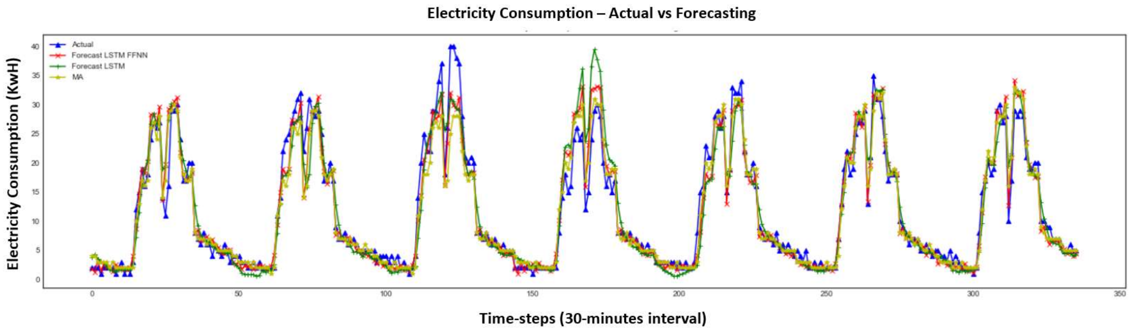


Fig. 10. Electricity consumption forecasting result 1-week Testing Data (Weekdays Data)

Based on the RMSE Score on the Table IV, our proposed method still reached the better result of the forecasting process. MA forecasted the next value based on the electricity consumption previous 5-days, different with our proposed method that not only just learn the pattern of the electricity consumption but also the detail of the additional information (days, time-scale, and seasons). As shown in Fig. 10, it has shown the comparation of forecasting results of LSTM, LSTM-FFNN, and MA.

Therefore, to forecast the weekday is better to separate the weekday and special days (holiday and weekend). The reason is electricity consumption in special day not as much as the usage in weekday. The separation of the data will help the network learn the pattern of electricity consumption in weekday.

## VI. Conclusion and Future Work

In this paper, the goal of the presented research was to forecast the electricity consumption of the next day for each 30-minutes. The proposed method based on combination of LSTM and FFNN have been presented in this paper to solve that problem. We obtained the better result on the proposed method rather than our baseline (original LSTM and MA). Based on the RMSE Score of our experiment result, we obtained the better result for the usage of "Weekdays" data compare with "All Day" data. In this work, we focus to forecast the electricity forecast for the weekday. Our experiment results (based on RMSE Scores) showed it is better to separate the way of electricity consumption forecasting on different type of day (weekday and holiday).

Furthermore, for the future work we will try to perform the model with using the different kind of dataset. Different data analysis is needed for different type of customer (ex: home, public facility, etc.). For the same case, we will focus to improve the performance of the model to give the better experimental result in the future.

## Acknowledgment

This research is supported by Scholarship from Indonesia Endowment Fund for Education (Lembaga Pengelola Dana Pendidikan (LPDP) Indonesia). LPDP has been providing the funds for the experimental equipment.

## References

[1] J. W. Taylor and P. E. Mcsharry, "Short-Term Load Forecasting Methods: An Evaluation Based on European Data", IEEE Transactions on Power Systems, Vol.22, No.4, pp. 2213-2219, 2007.

[2] H. S. Hippert, C. E. Pedreira, and R. C. Souza. "Neural Networks for Short-Term Load Forecasting: A Review and Evaluation", IEEE Transactions on Power Systems Vol.16, No.1, pp. 44-55, 2001.

[3] S. Mishra and S. K. Patra, "Short term load forecasting using a novel recurrent neural network", TENCON IEEE Region 10 Conference, Hyderabad, pp. 1-6, 2008.

[4] C. Liu, Z. Jin, J. Gu, and C. Qiu, "Short-term load forecasting using a long short-term memory network", IEEE PES Innovative Smart Grid Technologies Conference Europe (ISGT-Europe), 2017.

[5] Wan He, "Load Forecasting via Deep Neural Networks", Procedia Computer Science Vol.12, pp. 308-314, 2017.

[6] S. Hochreiter and J. Schmidhuber, "Long short-term memory", Neural Computation Vol.9, No.8, pp. 1735–1780, 1997.

[7] K. Amarasinghe, D. L. Marino, and M. Manic, "Building Energy Load Forecasting using Deep Neural Networks", 42nd Annual Conference of the IEEE Industrial Electronics Society (IECON), 2016.

[8] J. Schmidhuber, "Deep learning in neural networks: An overview", Neural Networks Vol.61, pp. 85–117, 2014.

[9] Limin F. "Neural networks in computer intelligence", Har/Ds ked. McGraw-Hill, International Series in Computer Science, 1994.

[10] P. J. Werbos, "Backpropagation through time: what it does and how to do it," Proceedings of the IEEE, Vol.78, No.10, pp.1550 – 1560, 1990.

[11] K. Amarasinghe, D. L. Marino and M. Manic, "Deep neural networks for energy load forecasting", IEEE 26th International Symposium on Industrial Electronics (ISIE), Edinburgh, pp. 1483-1488, 2017.

[12] D. P. Kingma and B. Jimmy, "Adam: A Method for Stochastic Optimization", 3rd International Conference for Learning Representations, San Diego, 2015.

[13] D. Grzegorz, "Neural networks for pattern-based short-term load forecasting: A comparative study", Neurocomputing Vol.205, pp. 64-74, 2016.

[14] R. K. Agrawal, F. Muchahary, M. M. Tripathi, "Long term load forecasting with hourly predictions based on long-short-term-memory networks", IEEE Texas Power and Energy Conference (TPEC), 2018.

[15] Ryan G. Hefron, Brett J. Borghetti, James C. Christensen, and Christine M. Schubert Kabban,"Deep long short-term memory structures model temporal dependencies improving cognitive workload estimation", Pattern Recognition Letters Vol.94, pp. 96-104, 2017.

[16] M. Elena, N. Phuong, and G. Madeleine, L. K. Wil, "Deep Learning for Estimating Building Energy Consumption", Sustainable Energy, Grids and Networks. Vol.6, pp. 91-99, 2016.

[17] T. Dave and L. Kornel, "Neural Network for Time Series Prediction", 2006.