



# Article Improving Electric Energy Consumption Prediction Using CNN and Bi-LSTM

Tuong Le <sup>1</sup>, Minh Thanh Vo <sup>2</sup>, Bay Vo <sup>3</sup>, Eenjun Hwang <sup>4</sup>, Seungmin Rho <sup>5</sup> and Sung Wook Baik <sup>1,\*</sup>

- <sup>1</sup> Digital Contents Research Institute, Sejong University, Seoul 05006, Korea; tuonglc@sju.ac.kr
- <sup>2</sup> Institute of Research and Development, Duy Tan University, Da Nang 550000, Vietnam; vmthanhit@gmail.com
- <sup>3</sup> Faculty of Information Technology, Ho Chi Minh City University of Technology (HUTECH), Ho Chi Minh City 700000, Vietnam; vd.bay@hutech.edu.vn
- <sup>4</sup> School of Electrical Engineering, Korea University, Seoul 02841, Korea; ehwang04@korea.ac.kr
- <sup>5</sup> Department of Software, Sejong University, Seoul 05006, Korea; smrho@sejong.edu
- \* Correspondence: sbaik@sejong.ac.kr

Received: 12 September 2019; Accepted: 3 October 2019; Published: 10 October 2019



Abstract: The electric energy consumption prediction (EECP) is an essential and complex task in intelligent power management system. EECP plays a significant role in drawing up a national energy development policy. Therefore, this study proposes an Electric Energy Consumption Prediction model utilizing the combination of Convolutional Neural Network (CNN) and Bi-directional Long Short-Term Memory (Bi-LSTM) that is named EECP-CBL model to predict electric energy consumption. In this framework, two CNNs in the first module extract the important information from several variables in the individual household electric power consumption (IHEPC) dataset. Then, Bi-LSTM module with two Bi-LSTM layers uses the above information as well as the trends of time series in two directions including the forward and backward states to make predictions. The obtained values in the Bi-LSTM module will be passed to the last module that consists of two fully connected layers for finally predicting the electric energy consumption in the future. The experiments were conducted to compare the prediction performances of the proposed model and the state-of-the-art models for the IHEPC dataset with several variants. The experimental results indicate that EECP-CBL framework outperforms the state-of-the-art approaches in terms of several performance metrics for electric energy consumption prediction on several variations of IHEPC dataset in real-time, short-term, medium-term and long-term timespans.

Keywords: electric energy consumption prediction; energy management system; CNN; Bi-LSTM

# 1. Introduction

With the development of data, internet as well as computing power of computers, machine learning and deep learning [1,2] have been used in many areas such as construction [3–5], cybernetic [6,7], economic [8–11] and medical [12,13] to help professionals save time and effort. Utilizing machine learning in economic, Hoang et al. [14] introduced a full-fledged geo-demographic segmentation model for identifying and gaining insights of the most probable cause of churn for a bank dataset. Meanwhile, Le et al. [8–11] developed several machine learning models for dealing with imbalance data problem to forecast the bankruptcy in South Korea. These studies can help the investors, managers and government devise appropriate strategies to optimize the profits in their business. In addition, medical diagnosis is another practical application of machine learning and deep learning with many studies in this decade. Recently, Hemanth et al. [12] proposed a deep CNN model to classify brain

images to diagnose abnormal brain tumors. In addition, Le and Baik [13] proposes FSX framework that utilizes an oversampling technique and extreme gradient boosting (XGB) classifier to improve the prediction ability in self-care problem identification for children with disability. Through the above studies, machine learning and deep learning are growing dramatically and are being applied to most important areas.

Long Short-Term Memory (LSTM) is a popular specialized model of artificial recurrent neural network (RNN) which is capable of modeling sequential or temporal aspects of data. Unlike original RNN models which often get problem of vanishing or exploding gradients when training sequential data, LSTM introduces three new gates in each cell including the input gate, output gate and forget gate which are capable of capturing the temporal changes in extreme long sequential data. Therefore, it has been used widely for text, videos, and time-series analysis. Park et al. [15] proposed a lightweight as well as real-time system utilizing LSTM, a model of recurrent neural networks, for fault detection in smart factories. Next, Huang and Kuo [16] combined Convolutional Neural Network (CNN) and LSTM to particulate matter forecasting in smart cities. Monitoring particulate matter concentration will reduce several diseases such as asthma, lung cancer, or cardiovascular diseases. Meanwhile, Ran et al. [17] propose a method based on LSTM with attention mechanism for travel time prediction to improve the effectiveness of intelligent transportation systems. In addition, Lin et al. [18] developed a neural-encoded mention-hypergraph (NEMH) approach to recognize the nested structure mention entities which has ability to extract features for dealing with both overlapping and nested structures mentions automatically.

Electric energy consumption prediction (EECP), a multivariate time series forecasting issue, is an interesting issue that needs to be addressed for stable power supply. In recent years, there are many approaches proposed to predict the electric energy consumption [19–29] from various datasets. In 2012, Hebrail and Berard released the IHEPC dataset [30] on UCI Machine Learning Repository collected from an individual house in France. In 2019, Kim and Cho [31] proposed an effective model that combine the convolutional neural network and long short-term memory. The proposed model extracts spatial and temporal features to effectively predict energy consumption for IHEPC dataset. Although this approach has been successful in predicting electric consumption, improving the predictive performance of the IHEPC dataset is necessary. Therefore, this study proposed an improve model denoted by EECP-CBL for predicting the electric consumption that utilizes the combination of CNN and Bi-LSTM. The experiments were conducted to compare the proposed model and the-state-of-the-art models for IHEPC dataset using various metrics which are Mean Square Error (MAE), Root Mean Square Error (RMSE), Mean Absolute Error (MAE) and Mean Absolute Percentage Error (MAPE). The experimental results confirm that the proposed model outperforms the state-of-the-art approaches for predicting the housing energy consumption on IHEPC dataset.

The remaining of the study is structured as follows. In the next section, the related works regarding to the electric energy consumption prediction problem will be introduced. Section 3 introduces the materials and methodology for electric energy consumption prediction utilizing the combination of CNN and Bi-LSTM. In Section 4, the experimental results in terms of four common performance metrics in time series prediction including MSE, RMSE, MAE and MAPE as well as the processing time on the experimental dataset are discussed. Then, the conclusions as well as future works are discussed in the last section.

## 2. Related Works

In 2018, Oliveira et al. [19] developed the efficient techniques that combine the decomposition and Bootstrap aggregating (Bagging) to improve univariate predictions for mid-long term electric energy demand (monthly energy demand) in both kind of countries including developed and developing ones. This method was successful in electric energy consumption prediction for several experimental datasets which were collected from Canada, France, Italy, Japan, Brazil, Mexico and Turkey. Then, Wu et al. [20] proposed Adaptive Rate of Change (ARC) approach for estimating the electrical maximum demand in short-term timespan. This issue can recognize the peak demand pattern for both commercial and industrial customers. This approach would provide to the customers several direct and indirect benefits. In addition, the utility providers also benefit from demand reduction, cost control, and system stability. This approach was successfully applied to three different manufactories to estimate and manage the electrical maximum demand in short-term timespan. Next, Krishnan, Sakthivel and Vishnuvarthan [21] developed a novel neural network-based optimization approach for electronic energy consumption prediction. The authors implemented two approaches including the Neural Network based Genetic Algorithm denoted by NNGA and the Neural Network based Particle Swarm Optimization namely NNPSO for optimizing the weight of neural network to improve the performance. The proposed approach yields better results compared with the results of the CNN model in predicting the future energy consumption. The proposed approach was verified by the real time dataset collected from pecan street. Meanwhile, Fayaz and Kim [22] proposed an efficient model for electronic energy consumption prediction in residential buildings using deep extreme learning machine. The dataset used on this study for weekly and monthly energy prediction was collected from four residential buildings from January to December in 2010. In addition, Bouktif et al. [23] utilized feature selection and genetic algorithm to optimize a LSTM model for electric load prediction. The proposed approach obtained higher accuracy than other machine learning models for France metropolitan's electricity consumption data. Unfortunately, most of the data from above studies are not available. In 2018, Tanveer et al. [24] introduced a comprehensive survey regarding to the approaches based on data driven and large scale for predicting the energy demand in a building to provide better overview of this issue. Next, Moon et al. [25] proposed a hybrid electric energy consumption forecasting model for week-ahead total daily electric energy consumption forecasting. They classified electric energy consumption data by pattern similarity using the decision tree technique. Then, they used Random Forest and Multilayer Perceptron to select models with a better prediction performance in similar time series. They compared their hybrid model with other machine learning techniques using three university building clusters. They verified their hybrid model yielded better prediction performance than other machine learning techniques.

In 2019, Johannesen et al. [26] utilized several regression approaches for predicting the urban area electrical energy demand in an urban area located in Sydney from 2006 to 2010. The experimental results show that Random Forest Regressor model archived the best performance for short-term electrical energy demand prediction (30 min) while kNN model obtained the relatively better results for predicting the daily timespan of electrical energy demand. Meanwhile, Divina et al. [27] introduced a comparative empirical evaluation of different time series forecasting strategies in both statistical and machine learning based approaches for Short Term Electric Energy Consumption Prediction on a dataset regarding the electricity consumption registered by thirteen buildings located at the Pablo de Olavide (UPO) University campus in Seville, Spain, collected over five and a half years. The results found that Machine Learning approaches including bagging and boosting ensemble schemes could achieve better results. Bouazza and Deabes [28] developed a framework using Petri Nets for smart temperature control to reduce energy consumption for smart building. The proposed system reduces the electric energy consumption by automatically controller electrical devices. The experiments show that this system reduces a significant amount of electric energy consumption at ratio 46.79%. Kim et al. [29] proposed a recurrent inception convolution neural network (RICNN) to predict the electric energy consumption (48-time steps with an interval of 30 min). They combined RNN and 1-D convolution inception module to help calibrate the prediction time and the hidden state vector values calculated from the nearby time steps. They used three industrial distribution complexes in South Korea. They verified their proposed RICNN outperformed better prediction performance than Multilayer Perceptron, Convolutional Neural Networks, and Recurrent Neural Networks. Recently, Kim and Cho [31] proposed an efficient model namely CNN-LSTM using the combination of CNN and LSTM for stably predicting electric energy consumption on IHEPC dataset. The CNN-LSTM method

in this study achieves almost perfect prediction performance for IHEPC dataset compared with Linear Regression, LSTM approaches.

## 3. Materials and Methods

## 3.1. Acronym

EECP	The electric energy consumption prediction
CNN	Convolutional Neural Network
Bi-LSTM	Bi-directional Long Short-Term Memory
IHEPC	The individual household electric power consumption dataset
LSTM	Long Short-Term Memory
RNN	Recurrent neural network
MSE	Mean Square Error
RMSE	Root Mean Square Error
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error
MLP	The Multilayer perceptron
ReLU	The Rectified Linear Unit
EECP-CBL	The Electric Energy Consumption Prediction model utilizing the combination of CNN and Bi-LSTM

## 3.2. Datasets

The IHEPC dataset provided by Hebrail and Berard [30] on UCI Machine Learning Repository was used to validate several electric energy consumption prediction models [31] recently. This dataset contains 2,075,259 measurements collected from a house in Sceaux, France in five years between December 2006 and November 2010. A total of 25,979 missing values on 28 April 2007 were removed in preprocessing step. Table 1 presents all the variables in the dataset. There are nine variables (day, month, year, hour, minute, global active power, global reactive power, voltage and global intensity) and three variables collected from energy consumption sensors (sub metering 1, sub metering 2, and sub metering 3) attached with the meaning of each variable.

#	Variable Description					
1	Day	A value from 1 to 31				
2	Month	A value from 1 to 12				
3	Year	A value from 2006 to 2010				
4	Hour	A value from 0 to 23				
5	Minute	A value from 1 to 60				
6	Global active power	The household global minute-averaged active power (in kilowatt)				
7	Global reactive power	The household global minute-averaged reactive power (in kilowatt)				
8	Voltage	The minute-averaged voltage (in volt)				
9	Global intensity	The household global minute-averaged current intensity (in ampere)				
10	Sub metering 1	This variable corresponds to the kitchen, containing mainly a dishwasher, an oven and a microwave, hot plates being not electric, but gas powered (in watt-hour of active energy)				
11	Sub metering 2	This variable corresponds to the laundry room, containing a washing machine, a tumble-drier, a refrigerator and a light (in watt-hour of active energy)				
12	Sub metering 3	This variable corresponds to an electric water heater and an air conditioner (in watt-hour of active energy)				

#### Table 1. The features of the IHEPC dataset.

From the IHEPC dataset, four datasets including minutely, hourly, daily, and weekly datasets, which indicate for real-time, short-term, medium-term and long-term predictions respectively, will be created. Figure 1 shows the samples of minutely, hourly, daily, and weekly datasets which originated from the IHEPC dataset. For each global active power in the time series shown in Figure 1, there are six variables including global reactive power, voltage, global intensity, sub metering 1, sub metering 2 and sub metering 3.



Figure 1. Samples of four electrical energy consumption datasets.

# 3.3. The EECP-CBL Model

Figure 2 introduces the overall architecture of the proposed model namely EECP-CBL for predicting electric energy consumption utilizing the combination of CNN and Bi-LSTM. The input variables from IHEPC dataset are extracted by two CNN layers in the first module and passed to the two Bi-LSTM layers in the second module. The Bi-LSTM layers are used for information analysis and time series prediction. Finally, the proposed method can generate predicted electrical energy consumption by the last module consisting of two fully connected layers. The results values predicted by the proposed model are evaluated by several performance metrics for time series forecasting such as MSE, RMSE, MAE and MAPE.



Figure 2. The overview architecture of EECP-CBL model.

In the first module, two one-dimensional CNN layers followed by two max pooling layers to reduce the computational complexity were used to analysis the input variables for feature extraction. At the early stage, MLP was often preferred as feature extractor where it is considered as a feedforward artificial neural network However, the model becomes inefficient and redundancy due to the face that

each perceptron is connected with every other perceptron. CNN, in the other hand, is a special type of MLP where every neuron does not need connected to other neurons. Those neurons only connected to a region of data (matrices or vector inputs) and are panned around entire data according to certain size and stride. The panning of filters in CNN essentially allows parameter sharing, weight sharing so that the filter looks for a specific pattern and is location invariant. Specifically, a number of neurons in a CNN layer is combined by weights and biases. These values have to be learned by the training process. In these models, several input variables are provided to each neuron. Next, the dot product operator is executed and followed by an optional nonlinearity function. In our application, convolution 1D layers, pooling 1D layers and fully connected layer are used. In details, CNN takes the time series data in one dimensional form wherein the data are arranged in the order of sequential time instants. Let the one-dimensional input vector be  $\mathbf{x} = \{x_1, x_2, \dots, x_n, ec\}$  where  $x_n \in \mathbb{R}^{fd}$  where *f* denotes the features and  $ec \in \mathbb{R}$  denotes the energy consumption value. The convolution 1D constructs a feature map *fm* by applying the convolution operator on the input data with a filter  $w \in \mathbb{R}^{fd}$  where *f* denotes the features inherent in the input data producing at its output, new set of features which is fed to input of the next block in line. A new feature map *fm* is obtained from a set of features *f* as the following equation.

$$hl_{i}^{fm} = tanh(w^{fm}x_{i:i+f-1} + b)$$
(1)

The filter hl in Equation (1) is utilized to each set of features f in the input data defined by  $\{x_{1:f}, x_{2:f+1}, \ldots, x_{n-f+1}\}$  and generate a feature map denoted by  $hl = [hl_1, hl_2, \ldots, hl_{n-f+1}]$ . In addition,  $b \in R$  is a bias term and  $hl \in R_{n-f+1}$ .

The output of convolutional layers is summed of weighted inputs which are comprised from multi linear transformations. Normally the linear transformation cannot capture the complex structures in the data, so that a non-linear activation layer is preferred to apply after convolutional layers to learning the data better in training step. In this study, we choose ReLU activation function that apply max(0, x)to each of the inputs. Then, the output of the convolutional layer is passed to the pooling layer which performs a down-sampling operator. In our model, the max-pooling layer is applied on each feature map  $\vec{hl} = max\{hl\}$ . This process selects the most significant features with highest values. The output of the max-pooling layer is denoted as follows.

$$x'_i = CNN(\mathbf{x}_i) \tag{2}$$

where  $x_i$  is the input vector to the CNN network with the energy consumption and  $x'_i$  is the output of the CNN network to be fed to the next Bi-LSTM network. The proposed framework first passes the input vector to CNN and obtains the new vector  $x'_i$  based on the Equation (2). Then, to understand Bi-LSTM, we introduce the LSTM with forget gate structure proposed in [32]. The formulation is denoted by,

$$i_t = \sigma(W_i([x_t, y_{t-1}])) \tag{3}$$

$$f_t = \sigma \Big( W_f([x_t, y_{t-1}]) \Big) \tag{4}$$

$$o_t = \sigma(W_o([x_t, y_{t-1}])) \tag{5}$$

$$g_t = tanh \Big( W_g([x_t, y_{t-1}]) \Big) \tag{6}$$

$$c_t = f \odot c_{t-1} + i \odot g \tag{7}$$

$$y_t = o \odot tanh(c_t) \tag{8}$$

where *i*, *f*, *o*, *g* and *c* are input, forget, output, and input modulation gate respectively. Note that these are in n-dimensional real vectors. In Equations (3)–(6), the notation  $\sigma$  is a sigmoid function and  $W_i$ ,  $W_f$ ,  $W_o$ , and  $W_g$  are fully connected neural networks for the input, forget, output, and input modulation gates respectively. In Equations (7) and (8), the notation  $\odot$  is an element-wise product operator. LSTM model only considers one directional information on a sequence which leads to reduce

the effectiveness of LSTM model. Moreover, multiple directional information on the sequence contains valuable information. Therefore, bidirectional long short-term memory denoted by Bi-LSTM was developed which combines the forward and backward directions in the sequence [33]. The key idea of Bi-LSTM model is that it looks a particular sequence both from the front-to-back as well from the back to front. In which, a LSTM layer is for forward processing while the remaining layer is for backward processing. In this way, the network could capture the change of energy power both its past as well as its future. To understand this model, let's considering an input sequence *x* with *n* elements. The order of the forward LSTM is  $\{x_1, x_2, \ldots, x_n\}$  while that of the backward LSTM is  $\{x_n, x_{n-1}, \ldots, x_1\}$ . After the training process of the forward and backward LSTMs separately, these LSTMs are integrated by fusing their outputs in the previous step denoted as,

$$y(t) = y_F(t) \oplus y_B(n-t+1) \tag{9}$$

where  $y_F$  and  $y_B$  are the outputs of the forward and backward LSTMs, respectively while the notation  $\oplus$  expresses any integration operator including a simple adder, a neural network etc. In other words, Bi-LSTM combines the forward and backward directions based on Equation (9). The outputs of Bi-LSTM are fed to two fully connected layers to generate energy consumption values. The structure and configuration details of proposed approach are presented in Table 2.

#No	Layer Type	Neurons	Parameters	
1	Convolution1D	(None, None, 6, 64)	192	
2	MaxPooling1D	(None, None, 3, 64)	0	
3	Convolution1D	(None, None, 2, 64)	8256	
4	MaxPooling1D	(None, None, 1, 64)	0	
5	Flatten	(None, None, 64)	0	
6	Bi-LSTM	(None, None, 128)	66,048	
7	Bi-LSTM	(None, 128)	98,816	
8	Fully connected layer	(None, 128)	16,512	
9	Dropout	(None, 128)	0	
10	Fully connected layer	(None, 1)	129	

Table 2. The structure and configuration details of the proposed model.

## 4. Experiments

In this section, the experiments were conducted to compare four common performance metrics for time series prediction including MSE, RMSE, MAE and MAPE as well as the processing time of the experimental methods. Since the IHEPC is recorded in 5 five years so that we took the first 3 years as the training and the rest 2 years for the testing. The experimental methods were implemented in Keras library and performed in the server with 4 GPU cards of GTX 1080 Ti. Our model (https://github.com/vmthanh/electric\_enegery\_comsumpsion) is trained in 100 epochs, batch size 30 using Adam optimization with initial learning rate as 0.001.

The first metric is MSE that measures the average of the squares of the errors. In other words, it is the average squared difference between the predicted values and the actual values. The equation of MSE is as follows.

$$MSE = \frac{1}{n} \sum_{1}^{n} (y - \hat{y})^2$$
(10)

Meanwhile, RMSE is the standard deviation of prediction errors. Firstly, let's consider residuals which are a measure of how far from the regression line data points are. Therefore, RMSE is a measure

of how spread out these residuals are. This metric is commonly used in climatology, forecasting, and regression analysis to verify the experimental models and are determined as follows.

$$RMSE = \sqrt{\frac{1}{n} \sum_{1}^{n} (y - \hat{y})^2}$$
(11)

In addition, MAE measures the average magnitude of the prediction errors and ignores their directions. In more detail, it's the average the absolute differences between prediction and actual values for all instances in the testing set. Note that this measurement considers all individual differences are the same weight. MAE is determined as the following equation.

$$MAE = \frac{1}{n} \sum_{1}^{n} \left| y - \hat{y} \right| \tag{12}$$

The last metric namely MAPE is a measure of prediction accuracy of a forecasting method such as time series prediction. This metric expresses accuracy in percentage the following equation:

$$MAPE = \frac{100\%}{n} \sum_{1}^{n} \left| \frac{y - \hat{y}}{y} \right|$$
(13)

To show the effectiveness of EECP-CBL model, this section compares four above performance metrics such as MSE, RMSE, MAE and MAPE computed by Equations (10)–(13) respectively of Linear Regression, LSTM and CNN-LSTM [31] and the proposed model for minutely, hourly, daily and weekly datasets derived from the IHEPC dataset.

Table 3 presents the results of experimental methods for minutely dataset. The proposed model reaches the best values of MSE, RMSE, MAE and MAPE at 0.051, 0.225, 0.098 and 11.66 respectively while the second approach, CNN-LSTM, obtains 0.374, 0.611, 0.349 and 34.84 of MSE, RMSE, MAE and MAPE respectively. Meanwhile, LSTM and Linear Regression did not achieve good results for this dataset. The gaps between the results in terms of the predictive performance of our approach and others are very large. Therefore, the proposed model explicitly outperforms other approaches including Linear Regression, LSTM and CNN-LSTM for minutely dataset. In addition, this study also reports the training time and predicting time of the proposed method (3950 s) is double that of CNN-LSTM (2070 s) because EECP-CBL analyzes both directions including forward and backward directions in the sequence. In the other side, EECP-CBL reduced nearly 30% of the predicting time compared with CNN-LSTM. Note that in machine learning, the predicting time is more important than the training time. Trained model needs to be trained once and stored it for many predicting later.

 Table 3. Performances of the experimental methods for minutely dataset.

#No	Model	MSE	RMSE	MAE	MAPE	Training Time (s)	Predicting Time (s)
1	Linear Regression	0.405	0.636	0.418	74.52	1028	37.48
2	ĽSTM	0.748	0.865	0.628	51.45	6880	114.26
3	CNN-LSTM	0.374	0.611	0.349	34.84	2070	62.99
4	EECP-CBL	0.051	0.225	0.098	11.66	3950	43.83

Next, the performance results for hourly dataset are introduced in Table 4. For MSE and RMSE, our approach is slightly better than Linear Regression, LSTM and CNN-LSTM. For details, our approach got 0.298 and 0.546 of MSE and RMSE while CNN-LSTM, LSTM and Linear Regression obtain (0.355, 0.596), (0.515, 0.717) and (0.425, 0.652) of MSE and RMSE respectively. For MAE, our approach just

stands in second place with 0.392 while the first place is CNN-LSTM with 0.332. The gap of the first and second approach in terms of MAE is quite small. In addition, our approach stands the third place after CNN-LSTM (the first place) and LSTM (the second place) in terms of MAPE. Therefore, our approach is best approach in terms of MSE and RMSE while CNN-LSTM archived the best values in terms of MAE and MAPE for hourly dataset. In terms of processing time, our method increases 50% in training time and decrease 20% in predicting time for hourly dataset compared with CNN-LSTM model.

#No	Model	MSE	RMSE	MAE	MAPE	Training Time (s)	Predicting Time (s)
1	Linear	0.425	0.652	0.502	83.74	692.12	2.88
2	LSTM	0.515	0.717	0.526	44.37	2281.50	5.95
3	CNN-LSTM	0.355	0.596	0.332	32.83	820.70s	2.31
4	EECP-CBL	0.298	0.546	0.392	50.09	1296.34	1.87

Table 4. Performances of the experimental methods for hourly dataset.

The proposed method also achieves the best results for daily dataset which are shown in Table 5. For MSE, our approach improves nearly 50% (0.065) compared with CNN-LSTM (0.104). Meanwhile, LSTM and Linear Regression obtained the high values in MSE at 0.241 and 0.253 respectively for daily dataset. In addition, EECP-CBL also reaches the best values of RMSE, MAE and MAPE with 0.255, 0.191 and 19.15 respectively. These values are significantly better than other methods for daily dataset. Therefore, EECP-CBL is the best method for electric energy consumption prediction in medium-term timespans. Moreover, the training time of EECP-CBL for daily dataset is 61.36 s while that of CNN-LSTM is 42.35 s. Similar to the previous dataset, the proposed approach only requires 0.71 s for predicting time that is only 37% of the predicting time required by CNN-LSTM model.

#No	Model	MSE	RMSE	MAE	MAPE	Training Time (s)	Predicting Time (s)
1	Linear Regression	0.253	0.503	0.392	52.69	27.83	1.32
2	LSTM	0.241	0.491	0.413	38.72	106.06	2.97
3	CNN-LSTM	0.104	0.322	0.257	31.83	42.35	1.91
4	EECP-CBL	0.065	0.255	0.191	19.15	61.36	0.71

Table 5. Performances of the experimental methods for daily dataset.

Finally, Table 6 shows the performances of the experimental methods for weekly dataset. Similar to daily dataset, the proposed method improves nearly 50% for MSE and 30% for RMSE, MAE and MAPE compared with the second place (CNN-LSTM). In the meantime, the results of LSTM and Linear Regression are not good in terms of all the performance metrics for weekly dataset. Therefore, our approach is the best method for electric energy consumption prediction in long-term timespans. The training time and predicting time in the last experiment are the same with the previous experiments. The training times of the proposed approach and CNN-LSTM are 20.7 s and 14.12 s respectively. However, the predicting time of our approach is only 20% of predicting time required by CNN-LSTM model.

To sum up, this study creates the Figure 3 which shows the average percentages of experimental methods over four datasets. We first determine the average for MSE, RMSE, MAE and MAPE over the minutely, hourly, daily and weekly datasets. Then, we scale these above values to percentage to show the graph better which are shown in Figure 3. Obviously, the proposed approach namely EECP-CBL is the best method for most above datasets including minutely, hourly, daily and weekly datasets derived from the IHEPC dataset in terms of four common performance metrics including MSE, RMSE, MAE and MAPE. Moreover, although our method has a larger training time than CNN-LSTM,

the prediction times of our method are a lot better than those of CNN-LSTM. Based on the above analysis, the proposed approach can be recommended to use in intelligent power management system to predict the electric power consumption in the future.

#No	Model	MSE	RMSE	MAE	MAPE	Training Time (s)	Predicting Time (s)
1	Linear Regression	0.148	0.385	0.320	41.33	11.23	1.48
2	LSTM	0.105	0.324	0.244	35.78	24.42	3.66
3	CNN-LSTM	0.095	0.309	0.238	31.84	14.12	2.06
4	EECP-CBL	0.049	0.220	0.177	21.28	20.7	0.4

Table 6. Performances of the experimental methods for weekly dataset.



Figure 3. The average percentages of experimental methods over four datasets.

# 5. Conclusions

This study proposed EECP-CBL model for electric energy consumption prediction that utilizes the combination of CNN and Bi-LSTM on the IHEPC dataset. The proposed model consists of three modules including CNN, Bi-LSTM and FC modules for predicting the energy consumption. Firstly, two CNN layers in the first module were used to extract information from several variables in IHEPC dataset. Next, the results of the previous module as well as the trends of time series in two directions will be passed to the next module namely Bi-LSTM module that consists of the two Bi-LSTM layers. In the last module, EECP-CBL used two fully connected layers to finally predict the values of energy consumption. The experiments were conducted to indicate that EECP-CBL outperforms CNN-LSTM, LSTM and Linear Regression, which are the state-of-the-art models for electric energy consumption prediction, in terms of several performance metrics such as MSE, RMSE, MAE and MAPE as well as processing time in different time frame settings.

In the future, we will improve the performance of electric energy consumption prediction model by applying several techniques such as evolutionary algorithms, optimized EECP-CBL model. In addition, we will try to collect more datasets regarding to electric energy consumption prediction to verify our proposed model.

Author Contributions: S.W.B. proposed the topic and obtained funding; T.L. and M.T.V. proposed and implemented the framework. T.L. wrote the paper. M.T.V., B.V., E.H. and S.R. improved the quality of the manuscript.

**Funding:** This research received no external funding.

Acknowledgments: This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. 2019M3F2A1073179).

Conflicts of Interest: The authors declare no conflict of interest.

# References

- 1. Lin, J.C.W.; Shao, Y.; Fournier-Viger, P.; Hamido, F. BILU-NEMH: A BILU neural-encoded mention hypergraph for mention extraction. *Inf. Sci.* **2019**, *496*, 53–64. [CrossRef]
- 2. Djenouri, Y.; Belhadi, A.; Lin, J.C.W.; Cano, A. Adapted K-Nearest Neighbors for Detecting Anomalies on Spatio-Temporal Traffic Flow. *IEEE Access* 2019, *7*, 10015–10027. [CrossRef]
- 3. Nguyen, T.N.; Lee, S.; Nguyen-Xuan, H.; Lee, J. A novel analysis-prediction approach for geometrically nonlinear problems using group method of data handling. *Comput. Methods Appl. Mech. Eng.* **2019**, *354*, 506–526. [CrossRef]
- Nguyen, T.N.; Thai, H.C.; Luu, A.T.; Nguyen-Xuan, H.; Lee, J. NURBS-based postbuckling analysis of functionally graded carbon nanotube-reinforced composite shells. *Comput. Methods Appl. Mech. Eng.* 2019, 347, 983–1003. [CrossRef]
- 5. Nguyen, N.T.; Thai, C.H.; Nguyen-Xuan, H.; Lee, J. Geometrically nonlinear analysis of functionally graded material plates using an improved moving Kriging meshfree method based on a refined plate theory. *Compos. Struct.* **2018**, *193*, 268–280. [CrossRef]
- 6. Nguyen, N.P.; Hong, S.K. Sliding Mode Thau Observer for Actuator Fault Diagnosis of Quadcopter UAVs. *Appl. Sci.* **2018**, *8*, 1893. [CrossRef]
- 7. Nguyen, N.P.; Hong, S.K. Fault-tolerant control of quadcopter UAVs using robust adaptive sliding mode approach. *Energies* **2019**, *12*, 95. [CrossRef]
- 8. Le, T.; Vo, M.T.; Vo, B.; Lee, M.Y.; Baik, S.W. A hybrid approach using oversampling technique and cost-sensitive learning for bankruptcy prediction. *Complexity* **2019**. [CrossRef]
- Le, T.; Vo, B.; Fujita, H.; Nguyen, N.T.; Baik, S.W. A fast and accurate approach for bankruptcy forecasting using squared logistics loss with GPU-based extreme gradient boosting. *Inf. Sci.* 2019, 494, 294–310. [CrossRef]
- 10. Le, T.; Le, H.S.; Vo, M.T.; Lee, M.Y.; Baik, S.W. A Cluster-Based Boosting Algorithm for Bankruptcy Prediction in a Highly Imbalanced Dataset. *Symmetry* **2018**, *10*, 250. [CrossRef]
- 11. Le, T.; Lee, M.Y.; Park, J.R.; Baik, S.W. Oversampling techniques for bankruptcy prediction: Novel features from a transaction dataset. *Symmetry* **2018**, *10*, 79. [CrossRef]
- 12. Hemanth, D.J.; Anitha, J.; Náaji, A.; Geman, O.; Popescu, D.E.; Le, H.S. A Modified Deep Convolutional Neural Network for Abnormal Brain Image Classification. *IEEE Access* **2019**, *7*, 4275–4283. [CrossRef]
- 13. Le, T.; Baik, S.W. A robust framework for self-care problem identification for children with disability. *Symmetry* **2019**, *11*, 89. [CrossRef]
- 14. Hoang, V.L.; Le, H.S.; Khari, M.; Arora, K.; Chopra, S.; Kumar, R.; Le, T.; Baik, S.W. A New Approach for construction of Geo-Demographic Segmentation Model and Prediction Analysis. *Comput. Intell. Neurosci.* **2019**. [CrossRef]
- 15. Park, D.; Kim, S.; An, Y.; Jung, J.Y. LiReD: A Light-Weight Real-Time Fault Detection System for Edge Computing Using LSTM Recurrent Neural Networks. *Sensors* **2018**, *18*, 2110. [CrossRef] [PubMed]
- 16. Huang, C.J.; Kuo, P.H. A Deep CNN-LSTM Model for Particulate Matter (PM2.5) Forecasting in Smart Cities. *Sensors* **2018**, *18*, 2220. [CrossRef]
- 17. Ran, X.; Shan, Z.; Fang, Y.; Lin, C. An LSTM-Based Method with Attention Mechanism for Travel Time Prediction. *Sensors* **2019**, *19*, 861. [CrossRef]
- 18. Lin, J.C.W.; Shao, Y.; Zhou, Y.; Pirouz, M.; Chen, H.C. A Bi-LSTM mention hypergraph model with encoding schema for mention extraction. *Eng. Appl. Artif. Intell.* **2019**, *85*, 175–181. [CrossRef]
- 19. 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. [CrossRef]
- 20. Wu, D.C.; Amini, A.; Razban, A.; Chen, J. ARC algorithm: A novel approach to forecast and manage daily electrical maximum demand. *Energy* **2018**, *154*, 383–389. [CrossRef]
- 21. Krishnan, M.; Sakthivel, R.; Vishnuvarthan, R. Neural network based optimization approach for energy demand prediction in smart grid. *Neurocomputing* **2018**, *273*, 199–208.

- 22. 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. [CrossRef]
- Bouktif, S.; Fiaz, A.; Ouni, A.; Serhani, M.A. Optimal Deep Learning LSTM Model for Electric Load Forecasting using Feature Selection and Genetic Algorithm: Comparison with Machine Learning Approaches. *Energies* 2018, 11, 1636. [CrossRef]
- 24. Tanveer, A.; Huanxin, C.; Yabin, G.; Jiangyu, W. A comprehensive overview on the data driven and large scale based approaches for forecasting of building energy demand: A review. *Energy Build.* **2018**, *165*, 301–320.
- 25. Moon, J.; Kim, Y.; Son, M.; Hwang, E. Hybrid Short-Term Load Forecasting Scheme Using Random Forest and Multilayer Perceptron. *Energies* **2018**, *11*, 3283. [CrossRef]
- 26. Johannesen, N.J.; Kolhe, M.; Goodwin, M. Relative evaluation of regression tools for urban area electrical energy demand forecasting. *J. Clean. Prod.* **2019**, *218*, 555–564. [CrossRef]
- Divina, F.; Torres, M.G.; Goméz Vela, F.A.; Noguera, J.L.V. A Comparative Study of Time Series Forecasting Methods for Short Term Electric Energy Consumption Prediction in Smart Buildings. *Energies* 2019, 12, 1934. [CrossRef]
- 28. Bouazza, K.E.; Deabes, W.A. Smart Petri Nets Temperature Control Framework for Reducing Building Energy Consumption. *Sensors* 2019, *19*, 2441. [CrossRef]
- 29. Kim, J.; Moon, J.; Hwang, E.; Kang, P. Recurrent inception convolution neural network for multi short-term load forecasting. *Energy Build.* **2019**, *194*, 328–341. [CrossRef]
- 30. Hebrail, G.; Berard, A. Individual Household Electric Power Consumption Data Set. UCI Machine Learning Repository. Available online: https://archive.ics.uci.edu/ml/datasets/individual+household+electric+power+ consumption (accessed on 1 September 2019).
- Kim, T.Y.; Cho, S.B. Predicting residential energy consumption using CNN-LSTM neural networks. *Energy* 2019, 182, 72–81. [CrossRef]
- Gers, F.A.; Schmidhuber, J.; Cummins, F. Learning to Forget: Continual Prediction with LSTM. *Neural Comput.* 2000, 12, 2451–2471. [CrossRef] [PubMed]
- 33. Graves, A.; Schmidhuber, J. Framewise phoneme classification with bidirectional LSTM and other neural network architectures. *Neural Netw.* **2005**, *18*, 602–610. [CrossRef] [PubMed]



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).