Çukurova Üniversitesi Mühendislik Fakültesi Dergisi, 39(1), ss. 167-174, Mart 2024
Cukurova University Journal of the Faculty of Engineering, 39(1), pp. 167-174, March 2024

Load Balance Forecasting Based on Hybrid Deep Neural Network

Hajir KHALAF[1] ORCID *0000-0003-2887-225X*
Nezihe YILDIRAN[*2] ORCID *0000-0002-5902-1397*

[1]Bahçeşehir University, Faculty of Engineering and Natural Sciences, Electric-Electronic Engineering Department, İstanbul, Türkiye
[2]Bahçeşehir University, Faculty of Engineering and Natural Sciences, Energy Systems Engineering Department, İstanbul, Türkiye

Geliş tarihi: 12.08.2023   Kabul tarihi: 28.03.2024

Atıf şekli/ How to cite: KHALAF, H., YILDIRAN, N., (2024). Load Balance Forecasting Based on Hybrid Deep Neural Network. Cukurova University, Journal of the Faculty of Engineering, 39(1), 167-174.

Abstract

Load forecasting is the foundation of utility design, and it is a fundamental business problem in the utility industry. Load forecasting, mainly referring to forecasting electricity demand and energy, is being used throughout all segments of the electric power industry, including generation, transmission, distribution, and retail. In this paper, a long short-term memory network with a hybrid approach is improved with a dense algorithm and proposed for electricity load forecasting. A long short-term memory network is designed to effectively exhibit the dynamic behavior of load time series. The proposed model is tested for Panama study including historical data and weather variables. The prediction accuracy is validated by performance metrics, and the best of the metrics are attained when mean absolute error is 5.262, mean absolute percentage error 0.0000376, and root mean square error 18.243. The experimental results show a high prediction rate for load balance forecasting of electric power consumption.

Keywords: Dense layer, Load forecasting, Long short-term memory, Machine learning

Hibrit Derin Sinir Ağına Dayalı Yük Dengesi Tahmini

Öz

Yük tahmini, hizmet tasarımının temelidir ve hizmet sektöründe temel bir iş sorunudur. Ağırlıklı olarak elektrik talebini ve enerjiyi tahmin etmeye atıfta bulunan yük tahmini, üretim, iletim, dağıtım ve perakende dahil olmak üzere elektrik enerjisi endüstrisinin tüm segmentlerinde kullanılmaktadır. Bu bildiride, hibrit bir yaklaşıma sahip uzun bir kısa süreli bellek ağı, yoğun bir algoritma ile geliştirilmiş ve elektrik yükü tahmini için önerilmiştir. Uzun bir kısa süreli bellek ağı, yükleme süresi serilerinin dinamik davranışını etkili bir şekilde sergilemek için tasarlanmıştır. Önerilen model, tarihsel verileri ve hava durumu değişkenlerini içeren Panama çalışması için test edilmiştir. Tahmin doğruluğu, performans ölçümleriyle doğrulanır ve ölçümlerin en iyisi, ortalama mutlak hata 5,262, ortalama mutlak yüzde hatası 0,0000376 ve kök ortalama kare hatası 18,243 olduğunda elde edilir. Deneysel sonuçlar, elektrik gücü tüketiminin yük dengesi tahmini için yüksek bir tahmin oranı göstermektedir.

Anahtar Kelimeler: Makine öğrenimi, Uzun kısa süreli bellek, Yük tahmini, Yoğun katman

[*]Sorumlu yazar (Corresponding Author): Nezihe YILDIRAN, *nezihe.kucukyildiran@eng.bau.edu.tr*

1. INTRODUCTION

Recently, forecasting systems behavior have been considered as an important operation in succeeding and getting the desired results in multiple fields, as in predicting the financial affairs of exchanging goods or indication of markets, in business to schedule employees, manage the stockpiled, and forecast the request, in medicine to observe the prevalence of illness, and in meteorology for forecasting weather. Equally, forecasting plays a crucial function in con- trolling the usage of electric, and electric power purchase in interconnected power system [1]. The term of load forecasting is a technique used by electric utilities or by energy management system (grid operator) to predict the electricity needed to balance the generation and load demand [2]. Lately, deep learning, machine learning, and artificial intelligence techniques own numerous contests, and have been successfully employed in electric consumption load predictions, financial market predictions, and reliability predictions, be- cause of its ability in handling with a huge data and gives the right prediction results [3].

The desired forecasts attain the veritable patterns, and intended relation- ships, that owns by the historical data, but do not duplicate the previous situations, that will not appear again. In time-series forecasting, there are resemblance between the variant types of application in the field of preprocessing, and the adapted techniques. Forecasting algorithms can be implemented and used only for a specific type of forecasting. In general, there are three main types of forecasting that depends on the type of application and field: short-term, medium-term, and long-term forecasting [4].

Multiple number of researchers develop criteria to perform the operation of load balanced forecasting in a vast number of fields, depending on the utilization of the machine learning algorithms and deep learning. The load forecast models are produced within the use of four supervised machine learning techniques and algorithms including support vector machine (SVM), regression tree (RT), feed forward neural network (FFNN), and multiple linear regression (MLR) [5]. Recurrent extreme learning machine (RELM) is presented as a proposed method for electricity load forecasting in an accurate manner [6]. Deep learning (DL) along with data mining techniques are obtained for forecasting electricity load and price [7].

Long short term memory (LSTM) with multiple configurations to build the models of forecasting for short to medium-term overall load forecasting is applied to give a solution of the accuracy in load forecasting, in which multiple linear and non-linear machine learning algorithms are used and trained, and the superior baseline is picked, selecting superior features within the use of casing, and embedded feature selection approaches, and at the end the genetic algorithm (GA) is utilized to allocate the best time tardiness, and the count of the LSTM model layers' predictive performance enhancement [8]. A novel model is introduced for measuring load-forecasting based on improved version of LSTM that clearly denoted the recurrent characteristic in the electric load by utilizing many concatenations of entries of time tardiness. An autoregressive model is designed within an Auto-Correlation Function (ACF) to retract consumption and clarify the most pertinent time tardiness to feed the multi-sequence LSTM [9]. A deep recurrent neural network with LSTM units (DRNN-LSTM) model is proposed to forecast gathered residential power load, electric vehicles, energy storage system, and photovoltaic power produce in community micro-grid. To boost the balance of the supply-request, the failures of both PV power outcome, and residential power load are consumed in the model by combining the results of forecasting [10].

In this paper, hybrid deep neural network model is designed and validated. The proposed model is based on LSTM model, and it is improved by using dense layers to attain better performance in feature representation of prediction in electricity forecasting, and considering the time dependencies and consuming, and overcome non-linearity problems as well as the most important achieve a higher forecasting accuracy. The organization of this research article is given as: Section 2 presents the theoretical background. Section 3 describes the proposed model including algorithm and criteria. The performance of proposed method by using simulation results is demonstrated in Section 4. Finally, the conclusion of proposed research work and future suggestions are given in Section 5.

2. THEORETICAL BACKGROUND

In general, there are numbers of deep learning algorithms within variant architectures that have been developed and introduced to be applied in many kinds of domains, and the most popular utilized architectures of DL neural networks are conventional neural network (CNN), auto encoder, restricted Boltzmann machine (RBM), and long short-term memory (LSTM) [11]. LSTM was firstly introduced by Hochreiteri and Schmidhuber in 1997 to beat the problems of vanishing the gradient, mainly happening when learning using dependencies of long-term, even when the minimum periods of time are too long. It is composed of a set of recurrently linked subsets, noted as blocks of memory. Each block is composed of several cells belonging to the memory that can keep and save the information over an intended period.

Also, networks have gates that are used for controlling the flow of the information through the memory cells of neurons, and basically named as input, output, and forget gate. Every gate of LSTM acquires the same input as the input neuron. Moreover, each of them possesses an activation function [12]. LSTM does not resemble the traditional recurrent unit, because it overwrites its tenor in every time step. The unit of LSTM have the ability to make decision whether to keep the current memory by determined the gates such as, if the input gate notifies a maximum activation, the input will have kept in the cell of memory. If the output gate detects this high activation, it will emit the saved information to the afterward neurons, in another manner if the forget gate observes the intended high activation, the memorycell will be erased [13].

3. PROPOSED METHODOLOGY

The proposed model design is considered an important operation, in which it must be illustrated the model in an applicable manner, because of its significant effect when implementing it in real-world applications. The proposed system aims to perform an electricity load forecasting, within the use of the deep learning concept, in which a proposed model that depends on the LSTM algorithm is used to provide a data regression on the attained data set, that contain a description about some effects on the electricity load, and what are the number of loads in the week, month, and year. The system will go through two phases after splitting the data into two groups including data normalization and data regression. Data normalization is performed using max-min normalization technique, in which all the data set elements are first normalized before getting into the regression phase. The data regression phase accomplished using a proposed deep hybrid-based model is implemented, as illustrated in Figure 1.

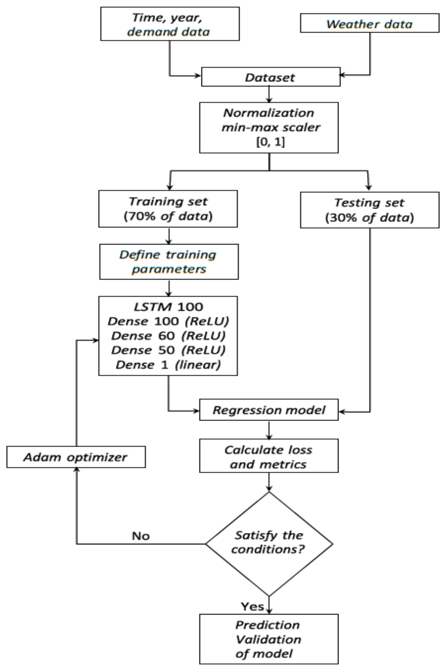

Figure 1. The flow chart of proposed deep hybrid model for load forecasting

3.1. Dataset Description

The attained dataset that the proposed load forecasting depending on is illustrated in this section, that known as short-term electricity load forecasting (Panama case study) dataset and can be found as Kaggle dataset.

The main sources of the data obtain the post-transmission of the electricity load, that depends on both of daily, and weekly pre-transmission of generated electricity of load forecast data on a weekly basis, both with hourly subdivisions. The data of time and periods of holidays and school, also weather data is illustrated as an important element in this dataset. The dataset of the electricity forecasting is hourly records, and composed of historical electricity load, historical weekly forecasts, calendar information related to school periods, calendar information related to holidays, weather parameters, including temperature, relative humidity, liquid precipitation, and wind speed of three major cities in Panama (David, Santiago, and Tocumen).

3.2. Normalization

Min-max scaler standardization is one of the most popular utilized techniques for data normalization, in which it implements a linear conversion on the basic data, be performing a rescaling operation on the features or on the outcomes in any domain into a new one. Mainly, the features are scaled in a range fall in between [0, 1], or [1, 1]. When min-max method is implemented, every feature stays the same, while obtaining place in the new domain, by considering that all the relational properties in the data are still without changing [14]. When the feature has a constant value in a determined data, it should be ejected, due to that it does not give any information, and min-max normalization has the benefit of maintaining the exact relationships exist in the data, and can be computed with Equation (1) where X_{min} denotes the minimum value in X feature, and X_{max} is the maximum value in X feature [15].

$$X' = \frac{X - X_{min}}{X_{max} - X_{min}} \tag{1}$$

3.3. Regression Model

The neural network methods are the most prevalent in the models of Artificial Intelligence (AI) that attained to give a solution for the problems of time series that are nonlinear. Furthermore, if there are variety of hidden layers, the neural networks do not operate in an acceptable manner, due to the technique of the back propagation. It may consume too long time, and in many cases, it gets local minimum with a poor count, and slow concourse, because of the random initialization. To solve the complexity problems facing when dealing with machine learning, deep learning architectures considered to be as the best choice for it.

In neural networks, Long Short-Term Memory considered to be one of the main categories that can be utilized in deep architecture for data regression purpose. A proposed model that depends on the LSTM as well as dense network is introduced in this research to perform data regression and provide an accurate prediction result for the electricity forecasting. The designed model of network is a back propagation type, in which the weights related to a determined network communications are repeatedly regulated to decrease the difference between the original output vector of the determined net, and the requested output vector. In general, there are some important variables and functions of the deep LSTM, such as activation function, optimizer, and dense layer.

Activation function basically exists in the core of deep networks and give the network the ability to learn the complicated mappings operation. Mainly, it is described as mathematical equations utilized to decide the outcome of the intended neural network, connect to every neuron in the network, locate whether it must be activated or not. There are various activation functions, and two types of activation functions (linear and rectified linear unit) are utilized in the proposed system. The simplest activation function is simply a linear transform on the data. The linear activation used for linear regression tasks and has a limited form solution. Rectified linear unit (ReLU) is non-linear activation function, employed in the neural networks having a variety of layers. The major concern is that ReLU function does not give the ability to activate all the neurons at the same intended amount of time. However, ReLU function can accelerate the training speed of deep neural networks with a comparison within the traditional kinds. Additionally, ReLU function does not urge the problem of vanishing gradient when the count of layers grows more, due to this function does not have a close upper and lower bounds in their ranges. The important purposes to prefer ReLU are gradient stability and low computational complexity.

Optimizer is an algorithm, or an approach utilized to change the parameter related to the neural network including weights along with learning rate to minimize the value of losses. In general, the optimizers are implemented to solve optimization problems by reducing the amount of functionality. One of the methods is adaptive moment estimation (ADAM) optimizer and it operates by calculating the rates related to the adaptive learning for every parameter to keep and save the exponentially average deterioration of the previously squared gradients value along with the exponentially deterioration average of the prior gradients value.

Dense layer is also known as the fully connected layer communicating within the dense circuits to attain a feature map from all prior layers to all posterior layers. As a result, the model can have the ability in defining the relationships exist between the values describing the data in a simple manner when the intended model is operating.

4. EXPERIMENTAL RESULTS

4.1. Performance Measurements

The evaluation of a specific model in terms of its performance is being done based on certain measurements, and the attained once to evaluate the proposed model includes mean absolute error (MAE), mean squared error (MSE), root mean squared error (RMSE), and mean absolute percentage error (MAPE).

4.2. Implementation Environment Description

The utilized programming languages to accomplish the proposed system function is Python version 3.6.5. The computer machine, that attained to perform, and accomplish the proposed system procedures have the type of Lenovo laptop with CORE i7 the seventh generation a 16 GB RAM size and the screen cart is NVIDIA GTX 6G as well as the hard of the computer is SSD and the operating system is Windows 10 64 bit.

4.3. Implementation Results Using Proposed Hybrid Deep Neural Network

The proposed deep model is implemented on the electricity load fore- casting dataset to provide load prediction in an accurate manner. The deep model has two phases in general: training and testing. Total number of parameters are 71661 and the proposed model layers with respect to number of parameters are given in Table 1.

The data covering nearly five years period in David, Santiago, and Tocumen cities of Panama including historical electricity load, historical weekly forecasts, calendar information related to school periods, calendar information related to holidays, weather parameters, including temperature, relative humidity, liquid precipitation, and wind speed is utilized. The total 582289 data from 03/01/2015 to 28/05/2018 is used for training phase, while the data from 28/05/2018 to 06/10/2020 is separated for testing phase.

4.4. Training Phase Results

The system has been trained each time with a different epoch and look back value to provide the best training result within the best time. Epoch refers to cycle through the full training dataset. Look back is the term that show the period illustrate the number of the preceding time steps are attained to predict the posterior time step. Table 2 shows the attained results when implementing the training phase of the proposed system using different epoch and look back values. Whenever epoch is increased the training phase behavior is enhanced and the time required for training the system become less. As well as look back influence the attained results and the accuracy of the training. In general, the best attained result in the training phase is appeared when the epoch equal to 200, and the lock back is 40.

Table 1. Proposed deep hybrid model layers

Layer type	Parameter	Output shape
LSTM	52400	$(non, 100)$
Dense	10100	$(non, 100)$
Dense	6060	$(non, 60)$
Dense	3050	$(non, 50)$
Dense	51	$(non, 1)$

4.5. Testing Phase Results

Testing phase of the proposed system considered to be the most important, in which it provides the electricity load forecasting in its accurate manner, and in an obvious view. The testing phase results are presented by examining the evaluation parameters in Table 3 for epoch in range of 100 to 500 with look back values from 10 to 50.

It is noticed from the prediction results when the epoch is equal to 100, the best result is attained for a look back of 10, and whenever the look back increased the results become worse and the testing prediction result reaches its final value when the look back is equal to 50.

As given in Figure 2, result values for epoch of 200 show that the best result is attained when the look back is equal to 10 and the worst case is shown when the look back is equal to 30 but it is still better than the worst case in epoch of 100 and look back of 50.

Results when epoch is equal to 300 show an enhancement in the attained results for all look back values, and the best result attained when the look back equal to 10, while the worst one is when the look back is 30.

As illustrated in Table 3 for epoch of 400, the best result is obtained when the look back is equal to 10, and it is also the best attained result as shown in Figure 3. The prediction line is nearly identical to the testing line. However, the worst result is attained for epoch of 400 when the look back is equal to 50.

Epoch was not increased more, because there is a slight difference some- times increasing sometimes decreasing in results with epoch of 400, when epoch is equal to 500.

Table 2. Training results with various epoch and look back values

Epoch	Look back	MAE	RMSE	Time
100	10	5.090	16.405	88589.95
	20	4.321	10.924	6508.74
	30	12.290	6.019	5711.86
	40	8.732	12.901	4869.84
	50	12.086	15.299	7125.30
200	10	5.020	16.414	9055.91
	20	5.081	10.716	88809.90
	30	9.221	13.268	7010.93
	40	3.407	9.626	12472.09
	50	5.594	10.445	60526.38
300	10	4.615	16.349	5882.70
	20	5.861	11.437	11855.75
	30	6.673	11.553	7530.17
	40	5.500	10.767	18404.54
	50	5.959	10.935	6702.38
400	10	4.494	16.341	4792.02
	20	8.252	13.253	16743.95
	30	5.769	11.498	6967.99
	40	4.659	10.496	7479.41
	50	8.307	11.850	9996.48
500	10	7.382	16.909	5886.55
	20	5.414	10.895	35192.07
	30	3.795	11.151	4733.94
	40	7.760	11.923	6393.03
	50	4.358	10.753	6193.44

Table 3. Testing results with various epoch and look back values

Epoch	Look back	MAE	RMSE	MAPE
100	10	5.957	18.462	0.0000399
	20	6.990	18.390	0.0000483
	30	8.610	20.106	0.0000747
	40	11.363	20.101	0.0000518
	50	14.502	22.090	0.0000657
200	10	5.842	18.292	0.0000443
	20	7.830	18.136	0.0000444
	30	11.671	19.754	0.0000791
	40	6.355	17.374	0.0000580
	50	9.111	19.386	0.0000712
300	10	5.454	18.245	0.0000400
	20	8.029	18.322	0.0000461
	30	9.536	19.578	0.0000791
	40	8.239	19.079	0.0000525
	50	8.721	19.466	0.0000600
400	10	5.262	18.243	0.0000376
	20	9.880	20.249	0.0000387
	30	8.803	19.268	0.0000557
	40	7.851	21.696	0.0000627
	50	11.258	20.259	0.0000575
500	10	8.311	18.814	0.0000376
	20	8.204	18.957	0.0000522
	30	6.328	19.737	0.0000587
	40	10.156	21.891	0.0000498
	50	7.332	19.728	0.0000595

4.6. Comparison with Linear Machine Learning Model

To further validation the performance of the proposed model, same system with same constraints is examined by using linear machine learning the simplest and popular algorithm. The proposed model has superior performance in terms of evaluation metrics as shown in Table 4 giving the comparison results of the two models. confirmed.

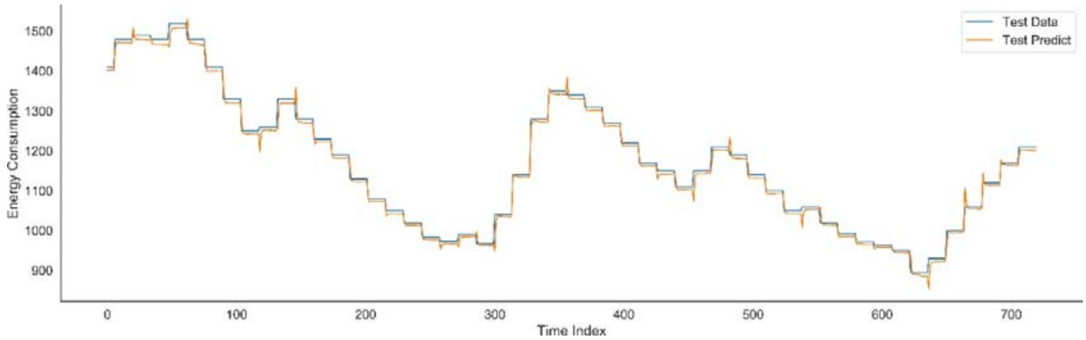

Figure 2. The forecasting result for epoch of 200 and look back of 30

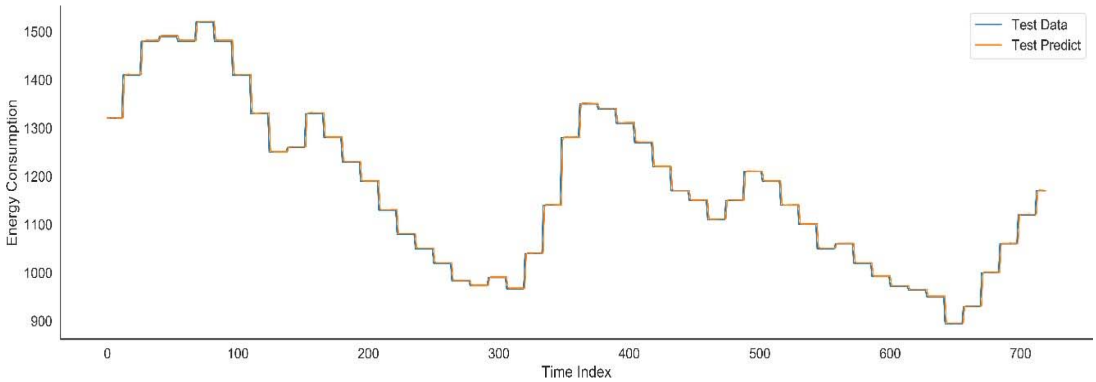

Figure 3. The forecasting result for epoch of 400 and look back of 10

Table 4. Testing results with various epoch and look back values

Model	MAE	RMSE	MAPE
Linear Regression	47.352	65.606	0.0002623
Proposed	5.262	18.243	0.0000376

5. CONCLUSION

Accurate and reliable forecasting for electricity consumption consideredan essential operation of energy management. The proposed system is depending on the famous deep learning *LSTM* network along with dense to provide the best result from the attained dataset. The experimental results show that the proposed deep learning model depending on the *LSTM* have a great performance and provide a high prediction results compared with the machine learning algorithm. The attained value of *MAE* for the proposed deep model is always less than 15 and the worst result of *RMSE* is equal to *22.090*, on the other hand, the best result obtained from the linear regression algorithm is equal to 47.352 for *MAE* and 65.606 for *RMSE*. The quality of the proposed deep model for electricity load forecasting reaching high prediction results is shown. When examining the system with epoch of 400, look back period of 10, the attained *MAE* is 5.262, *RMSE* is 18.243, and *MAPE* is 0.0000376. The results indicate nearly perfect prediction and the test prediction line is very close to the real data line.

6. REFERENCES

1. Kuster, C., Rezgui, Y., Mourshed, M., 2017. Electrical Load Forecasting Models: A Critical Systematic Review. Sustainable Cities and Society, 35, 257-270.
2. Fiot, J., Dinuzzo, F., 2016. Electricity Demand Forecasting by Multi-task Learning. IEEE Transactions on Smart Grid, 9(2), 544-551.
3. Dedinec, A., Filiposka, S., Dedinec, A., Kocarev, L., 2016. Deep Belief Network Based Electricity Load Forecasting: An Analysis of Macedonian Case. Energy, 115, 1688-1700.
4. Armstrong, J., 2001. Selecting Forecasting Methods. In Principles of Forecasting, Springer, 365-386.
5. Idowu, S., Saguna, S., Ahlund, C., Schelen, O., 2016. Applied Machine Learning: Forecasting Heat Load in District Heating System. Energy and Buildings, 133, 478-488.
6. Ertugrul, Ö., 2016. Forecasting Electricity Load by a Novel Recurrent Extreme Learning Machines Approach. International Journal of Electrical Power and Energy Systems, 78, 429-435.
7. Zahid, M., Ahmed, F., Javaid, N., Abbasi, R., Kazmi, H., Javaid, A., Bilal, M., Akbar, M., Ilahi, M., 2019. Electricity Price and Load Forecasting Using Enhanced Convolutional Neural Network and Enhanced Support Vector Regression in Smart Grids. Electronics, 8(2), 122.
8. Bouktif, S., Fiaz, A., Ouni, S., Serhani, M., 2018. Optimal Deep Learning LSTM Model for Electric Load Forecasting Using Feature Selection and Genetic Algorithm: Comparison with Machine Learning Approaches. Energies, 11(7), 1636.
9. Bouktif, S., Fiaz, A., Ouni, S., Serhani, M., 2019. Single and Multi-sequence Deep Learning Models for Short and Medium Term Electric Load Forecasting. Energies, 12(1), 149.
10. Wen, L., Zhou, K., Yang, S., Lu, X., 2019. Optimal Load Dispatch of Community Microgrid with Deep Learning Based Solar Power and Load Forecasting. Energy, 171, 1053-1065.
11. Shrestha A., Mahmood, A., 2019. Review of Deep Learning Algorithms and Architectures. IEEE Access, 7, 53040-53065.
12. Ryu, S., Noh, J., Kim, H., 2016. Deep Neural Network Based Demand Side Short Term Load Forecasting. Energies, 10(1), 3.
13. Hochreiter H., Schmidhuber, J., 1997. Long Short-term Memory. Neural Computation, 9(8), 1735-1780.
14. Dalton, B., 2019. Data mining: A Preprocessing Engine. Solid State Technology, 62(4), 9-16.
15. Liu, Z., 2011. A Method of SVM with Normalization in Intrusion Detection. Procedia Environmental Sciences, 11, 256-262.