



Comparison of Short-Term Electricity Load Forecasting Using Different Deep Learning Methods

Ipek Atik^{1*}

¹*Faculty of Engineering and Natural Sciences, Department of Electrical and Electronics Engineering, Gaziantep Islam Science and Technology University, Gaziantep, Turkey (ORCID: 0000-0002-9761-1347), ipek.atik@gibu.edu.tr

(First received 1 November 2021 and in final form 25 December 2021)

(DOI: 10.31590/ejosat.1017137)

REFERENCE: Atik I. (2021). Comparison of Short-Term Electricity Load Forecasting Using Different Deep Learning Methods. *European Journal of Science and Technology*, (31), 616-623.

Abstract

Estimation of the amount of electricity generation plays an important role in the planning of transmission and distribution systems, generation economy, unit work schedules and maintenance repair timing. With accurate forecasting models, uninterrupted and reliable electrical energy production can be achieved. In our study, 1-hour, 2-hour and 3-hour ahead predictions were made with different deep learning algorithms using Turkey's hourly electricity generation data. With the MAE, RMSE and correlation coefficient values of the models, their performances were compared. Among the compared methods, the best estimation result was obtained with the LSTM model. R values for 1-h, 2-h and 3-h ahead were obtained as 0.996, 0.973 and 0.959, respectively. The study aimed to determine the model that makes the closest estimation to the real values. In this context, it is anticipated that the study will be useful for future prediction studies

Keywords: Deep learning, Short-term prediction, LSTM, SVR, ANN

Farklı Derin Öğrenme Metotları ile Kısa Dönem Elektrik Yükü Tahmin Karşılaştırması

Öz

Elektrik üretim miktarının tahmin edilmesi iletim ve dağıtım sistemlerinin planlaması, üretim ekonomisi, ünite çalışma programları ve bakım onarım zamanlaması için önemli role sahiptir. Doğru tahmin modellemeleri ile kesintisiz ve güvenilir elektrik enerjisi üretimi sağlanabilir. Çalışmamızda Türkiye için saatlik elektrik üretim verileri kullanılarak farklı derin öğrenme algoritmaları ile 1 saat, 2 saat ve 3 saat ilerisi için tahmin çalışmaları yapılmıştır. Modellere ait MAE, RMSE ve korelasyon katsayısı değerleri ile performansları karşılaştırılmıştır. Karşılaştırılan yöntemler arasında en iyi tahmin sonucu LSTM modeli ile sağlanmıştır. 1-h, 2-h ve 3-h ilerisi için R değerleri sırasıyla 0.996, 0.973 ve 0.959 olarak elde edilmiştir. Çalışmada gerçek değerlere en yakın tahmin yapan modelin belirlenmesi amaçlanmıştır. Çalışmanın bu bağlamda gelecek tahmin çalışmaları için faydalı olacağı öngörülmektedir.

Anahtar Kelimeler: Derin öğrenme, Kısa dönem tahmin, LSTM, SVR, ANN

*Corresponding Author: ipek.atik@gibtu.edu.tr

1. Introduction

Electricity consumption is constantly rising with advancing technology. Providing continuous and reliable electricity generation is one of the factors that determine the level of development of countries. Forecasting studies are important in terms of planning of electricity transmission and distribution networks, economic criteria, and supply-demand balance. (Sarhani & El Afia, 2015) Electric load forecasting generally fall into three categories. These are short-term forecasts, midterm forecasts and long-term forecasts. Short-term forecasts include a period from a few hours to a few weeks; medium-term forecasts include a period between one week and one year, and long-term forecasts include a period of more than one year (Hochreiter & Schmidhuber, 1997). Short-term forecasts have an important role in the operation of electricity generation, energy market, unit work schedules, production economy, effective free-market offers and unit maintenance (Khan et al., 2018).

Many techniques related to short-term load forecasting (STLF) have been developed in the literature. The techniques generally include artificial intelligence-based and traditional approaches. Statistical methods are used more often in traditional approaches. Kaynar et al. (O. Kaynar, H. Ozekicioglu, & Demirkoparan, 2017) provided an error margin less than five percent with the hybrid algorithm in which the support vector regression algorithm and chaotic methods were used together for forecasting. Turkyay and Demre (B. E. Türkay & D. Demren, 2011) made a demand forecasting using the 2006-2009 load data with the library for support vector machine (LibSVM) algorithm. Ellatar et al. (Elattar, Goulernas, & Wu, 2010) conducted a forecasting study using LWSVR, a hybrid model of support vector, and stated that it could be an alternative to other methods. Nazarko and Zalewski (Nazarko & Zalewski, 1999) conducted an estimation study in which they compared the fuzzy regression and standard regression methods. Mostafa and Nagasaka (Al Mamun & Nagasaka, 2006) conducted an estimation study comparing SVM and ANN methods. Sarhani and Al Afia (Sarhani & El Afia, 2015) used SVA and PSA in a hybrid model to make predictions. Matijas (Matijaš, 2013) compared the least squares support vector machine (LS-SVM) and Robust LS-SVM methods with other methods and observed that they gave better results. Bozic and Stojanovic (Božić & Stojanović, 2011) made an estimation using the SVM method and calculated MAPE as 4.25%. Ghanbari et al. (Ghanbari, Naghavi, Ghaderi, & Sabaghian, 2009) compared the results of linear regression, logarithmic linear regression and ANN in their study using artificial neural network algorithm. Khan et al. (Khan et al., 2018) conducted an estimation study using ANN and bagged regression tree (BRT), and obtained better results

with the BRT algorithm. Kong et al. (Kong et al., 2017) applied long short-term memory (LSTM) based recurrent RNN in short term forecasting. They compared and tested the results with ELM, k-NN, and backpropagation algorithm models and proved that the LSTM model exhibited the best performance. Zheng et al. (Zheng, Ristovski, Farahat, & Gupta, 2017) used the long short-term memory (LSTM) RNN method on the smart meter data taken from a region connected to the smart grid and compared the obtained results with the seasonal autoregressive integrated moving average (SARIMA), Nonlinear Autoregressive Neural Network (NARX), Support Vector Regression (SVR) methods.

In our study, artificial neural networks (ANN), long short term memory (LSTM), linear regression (LR), support vector machines regression (SVR), bagged tree (BT) and fine tree (FT) deep learning methods will be used for short-term electrical load prediction. The first part of the study emphasizes the importance of load prediction and previous studies on short-term forecasting are summarized. The second part includes information about deep learning methods, data set and performance criteria used in the study. The third part includes forecasting results from the data set transferred to the model, and the comparison of the model performances. The last part includes results obtained from the study.

2. Material and Methods

In the study, deep learning methods such as linear regression, SVR, Bagged Tree, Fine Tree, ANN and LSTM will be used. Method descriptions, data set and model performance criteria are given in this section.

2.1 Data Collection

Data were obtained from the open database for Turkey's hourly electricity generation load between 01.05.2013-12.09.2020 as megawatt-hour (Kaggle, 2021). Hourly total electricity generation data of Turkey was used in the data set. A total of 64602 data is available in the specified date range. 64602 rows of data were found in the data set. 60000 rows of the used data set were reserved for training and validation, while the remaining 4602 rows were reserved for testing. Values in the data set were normalized between 0 and 1. The mathematical equation of the normalization is shown in Eq.1 (Zhang, Patuwo, & Hu, 1998). The hourly change graph of Turkey's electricity load for the date range given for the input data is as seen in Figure 1.

$$x_{new} = \frac{x - x_{min}}{x_{max} - x_{min}} \quad (1)$$

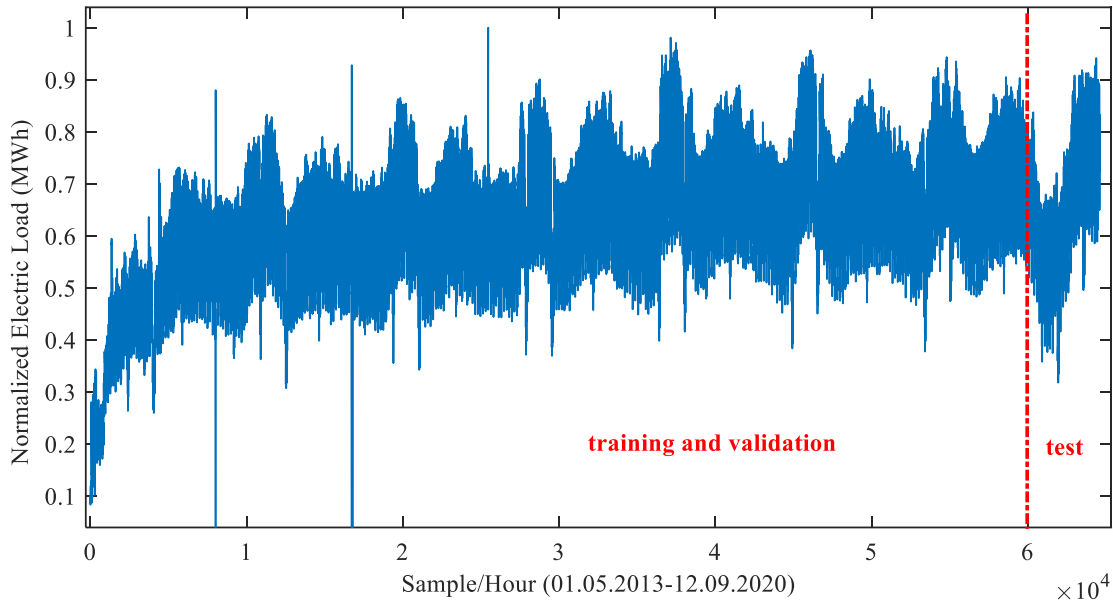


Figure 1. Hourly change of Turkey's normalized electricity load between 01.05.2013-12.09.2020

2.2 Deep learning algorithms

2.2.1 Regression Analysis

Regression analysis is a parametric method to determine the relationship between two or more variables. It is one of the methods used to determine the relationship between two or more variables that have a cause-effect relationship between them and to make future predictions about that subject based on this relationship. Regression model uses estimation, classification and analytical data tools to determine the importance of many explanatory variables. There are two types of regression models: simple regression with one independent variable and multiple regression analysis with more than one independent variables(Karaca & Karacan, 2016). In the study; linear regression, bagged tree, fine tree and SVR algorithms in Matlab R2020a program were applied to our dataset.

Linear regression algorithm is a way of modeling the relationship of a dependent variable to one or more independent variables, which is the basis of regression analysis. The main purpose here is to estimate the value of dependent variable based on the known or fixed values of independent variables(Karaca & Karacan, 2016).

Regression trees: Decision trees are in the form of a tree structure that can be built on both regression and classification models. Regression is used for numerical target data, while classification is used for categorical data (eg yes/no). Decision trees consist of decision nodes and leaf nodes according to feature and target.

Support vector machines regression: The SVR algorithm draws the parallel furthest line to separate the data formed in two groups(Omidi, Barakati, & Tavakoli, 2015).

2.2.2 Artificial Neural Networks

Artificial neural networks (ANN) consist of combining artificial nerve cells, which were inspired by the working structure of biological nerve cells. ANNs can learn and generalize, model nonlinear structures, be adapted for different problems, and have fault tolerance (Tosun, Ozturk, & Taspinar, F., 2019). Neural networks do not need prior knowledge between inputs and output variables. By giving input and output variables to the network, it is made to learn the relationship between them. This process is called supervised learning and the backpropagation algorithm, which is one of supervised learning methods, was used in the study. In artificial neural networks, the backpropagation learning algorithm consists of two stages. In forward phase, the input entering the network are processed with weight matrices and the output is calculated. Then, the network weights are rearranged with the backward propagation of the network, aiming to minimize error between the output value created by the network and the actual value. This process continues until the network produces the desired output(Sun, Qingdang, & Wang, 2020). Artificial neural networks are used in forecasting models based on time series and cause-effect relationships. In ANN models, the input is independent variable whereas the output is dependent variable. Nonlinear functional relationship for ANN is as shown in Eq. 2. (Zhang et al., 1998);

$$Y = f(x_1, x_1, \dots \dots x_n) \quad (2)$$

Here, $x_1, x_1, \dots \dots x_n$ n stands for independent variable and Y stands for dependent variable. For time series-based forecasts, ANN input variables represent the historical values of the data set, whereas the output represents the forecast value. The nonlinear relationship determined by the ANN is represented as in Eq.3 (Sun et al., 2020).

$$Y_{t+1} = f(y_t, y_{t-1}, \dots \dots y_{t-n}) \quad (3)$$

2.2.3 LSTM

Long Short Term Memory (LSTM) has three gates that control the flow of information, called input, forget, and output gates. Basically, these gates are simply generated by logistic functions of weighted sums; which can be calculated by backpropagation during training (Altan, 2019). Cell is managed through the input gate and forget gate. The output is repeatedly connected to the block input and to all gates.

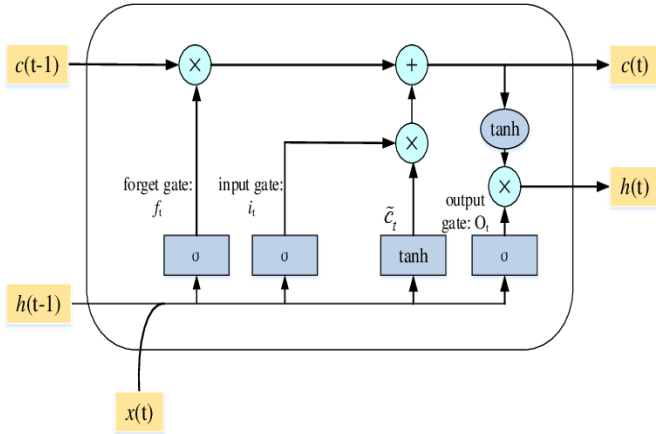


Figure 2. The diagram of LSTM memory block (Yuan, Li, & Wang, 2019)

The basic LSTM architecture is as shown in Figure 2. There are three gates in the LSTM unit; these are input, forget and output gates. For the basic LSTM structure, its external inputs are its previous cell state $c(t-1)$, the previous hidden state $h(t-1)$ and the current input vector $x(t)$. The three gates equations are given as;

$$i_t = \sigma(x_t W_{xi} + h_{t-1} W_{hi} + b_i) \tag{4}$$

$$f_t = \sigma(x_t W_{xf} + h_{t-1} W_{hf} + b_f) \tag{5}$$

$$o_t = \sigma(x_t W_{x0} + h_{t-1} W_{h0} + b_o) \tag{6}$$

where σ denotes the nonlinear activation function. In the LSTM, intermediate state $C(t)$ is formed as;

$$C_{(t)} = \tanh(x_t W_{xc} + h_{t-1} W_{hc} + b_c) \tag{7}$$

Then the memory cell and the hidden state are expressed as follows;

$$C_{(t)} = f_t \odot C_t + i_t \odot C_t \tag{8}$$

$$h_{(t)} = o_t \odot \tanh(C_t) \tag{9}$$

Where \odot and \tanh represents the multiplication operation and nonlinear activation function respectively. (Kong et al., 2017; Yuan et al., 2019).

2.3 Statistical Error Measures

Three evaluation measures were chosen to evaluate model performance; root mean square error (RMSE), correlation coefficient (R) and mean absolute error (MAE). While RMSE demonstrates the error between anticipated values and observed values, R expresses the relationship between inputs and outputs. MAE determines the average size of the errors in forecasts. Detailed equations of these expressions can be described as shown in Eq.10-12 (Kong et al., 2017);

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (Y_i - X_i)^2} \tag{10}$$

$$R = \frac{\sum_{i=1}^N (Y_i - \bar{Y}_i)(Y_i - \bar{X}_i)}{\sqrt{\sum_{i=1}^N (Y_i - \bar{Y}_i)^2 \sum_{i=1}^N (Y_i - \bar{X}_i)^2}} \tag{11}$$

$$MAE = \frac{1}{N} \sum_{i=1}^N |Y_i - X_i| \tag{12}$$

Where N refers the total number of testing samples, and Y_i denotes the forecast versus and X_i denotes actual values. Among these performance measures, R is the accuracy coefficient of the model. Higher values indicate a good forecasting relationship. Since RMSE and MAE are error measures, low values indicate high performance inversely proportional to performance (B. E. Türkay & D. Demren, 2011; Wang, Zhu, Zhang, & Lu, 2010). For example, RMSE equal to zero means high accuracy.

3. Results and Analysis

In the study, out of 64602 sample datasets for Turkey's hourly electricity generation, 60000 were used for training and 4602 were used for testing. The Intel(R) Core(TM) i7-10750H CPU @ 2.60GHz 2.59 GHz processor was used in the analysis. 1h, 2h and 3h ahead predictions were obtained by using the forecasting algorithms LSTM, linear regression (LR), SVR, Bagged Tree (BT), Fine Tree (FT) and ANN. Aforementioned methods, 1-h, 2-h and 3-h ahead results are as seen in Fig. 3, Fig.4 and Fig.5 respectively. The specifications of the algorithms and performance criteria metrics were obtained as summarized in Table.1-4.

Table 1. The specifications of the algorithms

Algorithm	Model Setup			
	Preset	Parameters		
Support Vector Regression	SVR	Fine Gaussian	Kernel Function	Gaussian
			Kernel Scale	0.64
Bagged Tree	BT	Bagged Trees	Min. Leaf Size	10
			Number of Learner	28
Decision Tree Regression	DTR	Fine Tree	Min. Leaf Size	8
			Decision Split	Off
Linear Regression	LR	Linear	Term	Interaction
			Robust Option	Off
			Maximum Epochs	100
			ValidationFrequency	25
Long Short Term Memory	LSTM	Linear	GradientThreshold	1
			Initial Learning Rate	0.001
			Learn Rate Drop Period	100
			Learning Rate Drop Factor	0.5

Figure 3. Prediction results for 1-h ahead of models

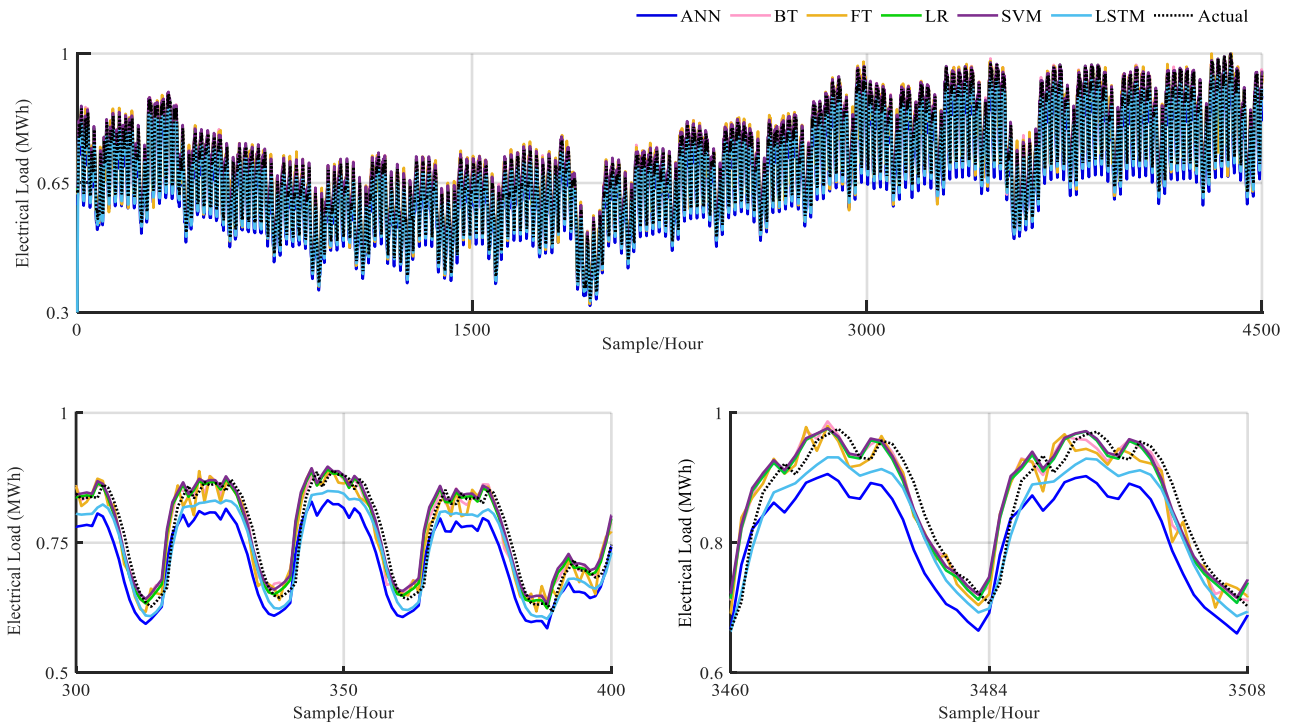


Table 2. Performance metrics of 1-h ahead prediction results

Metrics	LSTM	SVR	LR	ANN	BT	FT
R	0.996	0.984	0.982	0.980	0.977	0.971
RMSE	0.019	0.029	0.030	0.032	0.039	0.042
MAE	0.016	0.022	0.024	0.028	0.032	0.035

Comparison of the 1-hour ahead model forecasting performances shows that the highest R value was obtained in the LSTM model with 0.996, while the lowest R value belonged to FT with 0.971. As can be seen in Figure 3, while the actual values are close to the predicted values in LSTM, the error is higher in FT. Evaluation of model performances showed that the RMSE values of LSTM, SVR, LR, ANN, BT and FT models were obtained as 0.019, 0.029, 0.030, 0.032, 0.039 and 0.042, respectively. Comparison of MAE values revealed the lowest MAE value in LSTM with 0.016, while the model with the highest MAE value was FT with 0.035.

Figure 4. Prediction results for 2-h ahead of models

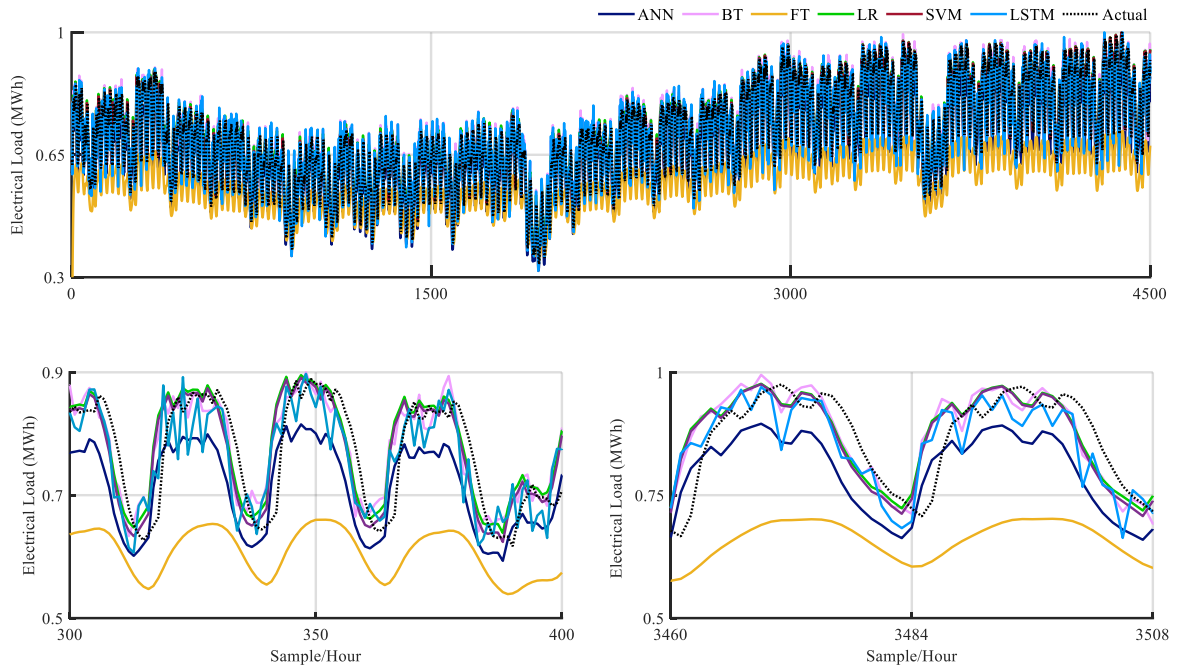


Table 3. Performance metrics of 2-h ahead prediction results

Metrics	LSTM	SVR	LR	ANN	BT	FT
R	0.973	0.945	0.938	0.934	0.924	0.916
RMSE	0.045	0.052	0.057	0.060	0.063	0.067
MAE	0.039	0.043	0.049	0.055	0.061	0.065

Prediction was performed for 2-hour ahead with the same methods, and the performance criteria were summarized in Table 2. Comparison of the 2-hour ahead model prediction performances shows that the highest R value was obtained in the LSTM model with 0.973, while the lowest R value belonged to FT with 0.916. Considering RMSE, the RMSE values of the LSTM, SVR, LR, ANN, BT and FT models for 2-hour ahead prediction were obtained as 0.045, 0.052, 0.057, 0.060, 0.063 and 0.067, respectively.

Figure 5. Prediction results for 3-h ahead of models

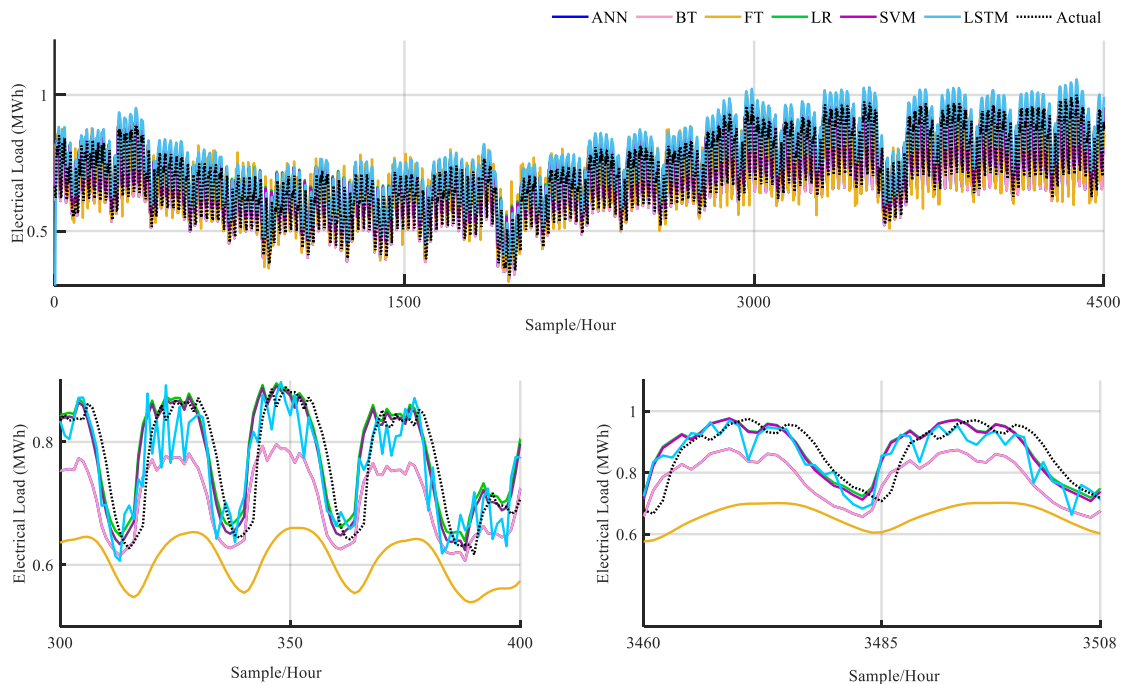


Table 4. Performance metrics of 3-h ahead prediction results

Metrics	LSTM	SVR	LR	ANN	BT	FT
R	0.959	0.882	0.878	0.872	0.852	0.817
RMSE	0.072	0.080	0.087	0.090	0.093	0.099
MAE	0.072	0.075	0.076	0.082	0.087	0.092

In comparison of R values for 3-hour ahead prediction, the best performance was observed in LSTM with 0.959 and FT with 0.817. Considering MAE, the MAE values of the LSTM, SVR, LR, ANN, BT and FT models for 3-hour ahead prediction were obtained as 0.072, 0.075, 0.076, 0.082, 0.087 and 0.092, respectively.

4. Conclusions

In this study, a short-term forecasting was performed using deep learning methods based on Turkey's electricity generation data. In the analysis, 60000 of the hourly total 64602 data between 01.05.2013-12.09.2020 were used for training and 4602 for testing. The results demonstrate the LSTM model showed the highest performance in 1-hour, 2-hour and 3-hour ahead predicts, whereas the Fine Tree method showed the worst performance. For 1-h ahead, R values of LSTM, SVR, LR, ANN, BT and FT models are respectively; 0.996, 0.984, 0.982, 0.980, 0.977 and 0.971 were obtained. For the 2-h ahead, the R values of these models are respectively; 0.973, 0.945, 0.938, 0.934, 0.924 and 0.916 were obtained. R values of the same models for 3-h ahead, respectively; 0.959, 0.882, 0.872, 0.852 and 0.817 were obtained. When the MAE and RMSE values are also examined, the model with the best performance was LSTM. In conclusion, SVR method exhibited higher success than linear regression, artificial neural networks, bagged tree and fine tree methods in estimating short-term electricity generation. It has been concluded that the study can be useful in making electricity load prediction and taking necessary precautions for network planning and production economy, unit work schedules and maintenance and repair timing. In future studies, performance analysis will be made with different algorithms in electric load estimation.

References

Al Mamun, M., & Nagasaka, K. (2006). Implementation of an Intelligent Method to Forecast Long-term Electric Demand. *Iranian Journal of Electrical and Computer Engineering (IJECE)*, 5(2), 75–82.

Altan, G. (2019). DeepGraphNet: Grafiklerin Sınıflandırılmasında Derin Öğrenme Modelleri. *Avrupa Bilim ve Teknoloji Dergisi*, 319–327. doi:10.31590/ejosat.638256

B. E. Türkay, & D. Demren. (2011). Electrical Load Forecasting Using Support Vector Machines (pp. 49–53). Presented at the International Conference on Electrical and Electronics Engineering, Nagpur.

Božić, M., & Stojanović, M. (2011). Application of SVM Methods for Mid-Term Load Forecasting. *Serbian Journal Of Electrical Engineering*, 8(1), 73–83.

Elattar, E. E., Goulermas, J., & Wu, Q. H. (2010). Electric Load Forecasting Based on Locally Weighted Support Vector Regression. *IEEE Transactions on Systems, Man, And Cybernetics—Part C: Applications And Reviews*, 40(4), 438–447.

Ghanbari, A., Naghavi, A., Ghaderi, S. F., & Sabaghian, M. (2009). Artificial Neural Networks and Regression Approaches Comparison for Forecasting Iran's Annual Electricity Load (pp. 675–679). Presented at the International Conference on Power Engineering, Energy and Electrical Drives. doi:10.1109/POWERENG.2009.4915245

Hochreiter, S., & Schmidhuber, J. (1997). *Long short-term memory* (Vols. 1-8, Vol. 9). Neural Computation.

Kaggle. (2021, December 6). Kaggle. *Kaggle data set*. dataset. Retrieved from <https://www.kaggle.com/datasets>

Karaca, C., & Karacan, H. (2016). Investigation Of Factors Affecting Demand For Electricity Consumption With Multiple Regression Method. *SUJEST*, 4(3), 182–195.

Khan, A. R., Razaq, S., Alquthami, T., Moghal, M., Amin, A., & Mahmood, A. (2018). Day Ahead Load Forecasting for IESCO Using Artificial Neural Network and Bagged Regression Tree. Presented at the 1st International Conference on Power, Energy and Smart Grid, Mirpur.

Kong, W., Dong, Z. Y., Jia, Y., Hill, D., Xu, Y., & Zhang, Y. (2017). Short-Term Residential Load Forecasting based on LSTM Recurrent Neural Network. *IEEE Transactions on Smart Grid, PP*, 1–1. doi:10.1109/TSG.2017.2753802

Matijaš, M. (2013). *Electric Load Forecasting Using Multivariate Meta-Learning* (Ph.D Thesis). University of Zagreb, Zagreb.

Nazarko, J., & Zalewski, W. (1999). The Fuzzy Regression Approach to Peak Load Estimation. *IEEE Transactions on Power Systems*, 14(3), 809–814.

O. Kaynar, H. Ozekicioglu, & Demirkoparan, F. (2017). Forecasting of Turkey's Electricity Consumption with Support Vector Regression and Chaotic Particle Swarn Algorithm. *Journal of Administrative Sciences*, 15(29), 2011–224.

Omidi, A., Barakati, S., & Tavakoli, S. (2015). Application of nuSupport Vector Regression in Short-Term Load Forecasting (pp. 32–36). Presented at the The 20th Iranian Electrical Power Distribution Conference, Zahedan.

Sarhani, M., & El Afia, A. (2015). Electric Load Forecasting Using Hybrid Machine Learning Approach Incorporating Feature Selection. Presented at the Proceedings of the International Conference on Big Data Cloud and Applications, Morocco.

Sun, Z., Qingdang, L., & Wang, L. (2020). Deep Learning Based Visual Object Tracker With Template Update. *University Politehnica of Bucharest Scientific Bulletin Series C-Electrical Engineering And Computer Science*, 82(2), 65–76.

Tosun, S., Ozturk, A., & Taspinar, F. (2019). Short Term Load Forecasting for Turkey Energy Distribution System with Artificial Neural Networks. *Tehnicki Vjesnik-Technical Gazette*, 26(6), 1545–1553.

Wang, J., Zhu, S., Zhang, W., & Lu, H. (2010). Combined modeling for electric load forecasting with adaptive particle swarm optimization. *Energy*, 35(4), 1671–1678.

Yuan, X., Li, L., & Wang, Y. (2019). Nonlinear Dynamic Soft Sensor Modeling With Supervised Long Short-Term Memory Network. *IEEE Transactions on Industrial Informatics*, 1–1. doi:10.1109/TII2019.2902129

Zhang, G., Patuwo, B. E., & Hu, M. Y. (1998). Forecasting with Artificial Neural Networks: The State of the Art. *International Journal of Forecasting*, 14(1), 35–62.

Zheng, S., Ristovski, K., Farahat, A., & Gupta, C. (2017). *Long Short-Term Memory Network for Remaining Useful Life estimation*.