INESEG

European Journal of Technique

journal homepage: https://dergipark.org.tr/en/pub/ejt

Vol.13, No.2, 2023

Research Article

Multilayer LSTM Model for Wind Power Estimation in the Scada System

Selahattin Barış Çelebi[1*], Ömer Ali Karaman[2]

[1*]Batman University, Department of Computer Technology, Batman, Turkey. (e-mail: sbariscelebi@gmail.com).
[2] Batman University, Department of Electronics and Automation, Batman, Turkey. (e-mail: omerali.karaman@batman.edu.tr).

ARTICLE INFO

Received: Oct., 30. 2023
Revised: Nov., 26 2023
Accepted: Nov, 27. 2023

Keywords:
Power forecasting
Wind turbine energy
Long short-term memory
Machine learning
Regression

Corresponding author: *Selahattin Barış Çelebi*

ISSN : *2536-5010* | e-ISSN: *2536-5134*

DOI: https://doi.org/10.36222/ejt.1382837

ABSTRACT

Wind energy is clean energy that does not pollute the environment. However, the complex and variable operating environment of a wind turbine often makes it difficult to predict the instantaneous active power generated. In this study, a wind turbine active power estimation system based on a long short-term memory network (LSTM) using time series analysis is proposed. The data obtained from the wind turbine SCADA system is used as input variables. In the proposed method, a multilayer LSTM architecture is designed to train the model. The first LSTM network consists of 64 units, and the second one consists of 32 units. This is followed by a dense layer consisting of 16 neurons. In the last layer, the architecture is finalized by using a linear activation function for the prediction process. The proposed deep learning (DL)-based LSTM model takes into account environmental factors such as wind speed and wind direction for active power forecasting. The results show that the LSTM-based time series analysis method is capable of effectively capturing time series features among the data. Thus, the proposed architecture can realize high-accuracy active power forecasting.

1. INTRODUCTION

Accurate forecasting of the active power from a wind turbine is critical for analyzing the energy demand [1], efficiency [2], and economic sustainability [3] of wind power plants. These forecasts are used to meet energy demand and reduce energy costs by influencing the structure of the energy grid [4]. Furthermore, the design and maintenance of wind turbines also rely on the predictions. If the power predictions are miscalculated, the energy production of the turbines may be lower or higher than expected. This can complicate the efficient use of resources and affect the stability of the energy grid [5]. Therefore, accurate active power forecasts are one of the key factors in the success of the wind energy industry [6].

Traditional statistical and machine learning-based prediction are the most widely used methods for turbine active power prediction [7]. Some common statistical techniques are time series analysis [8], Kalman filtering [9], and linear regression [10]. The fact that these methods are simple models is a great advantage. However, it is difficult to obtain satisfactory performance from statistical methods using big data from today's real-time applications [11]. When traditional machine learning-based methods are considered, there are some widely used algorithms such as support vector machines (SVM) [12], bagged trees (BT) [13], and extreme learning machines (ELM) [14]. In traditional machine-learning based methods, feature selection from the dataset is a difficult task [15, 16].

Machine learning technologies such as light gradient boosting machines (LightGBM), extreme gradient boosting (XGBoost), and recurrent neural networks (RNN) have been used for time series data analysis and prediction [17, 18]. XGBoost is a method that works well on datasets with large sizes. However, overfitting problems are encountered due to incorrect hyperparameters for this technique. Also, the method needs feature engineering, which requires technical skills and experience [17]. Another method, LightGBM is a method that stands out with its speed. However, it may require more memory compared to other traditional methods [18]. While RNN networks can be successful in short-time series, they face the problem of losing the information obtained in long-time interval dependencies. Due to this problem, called the vanishing gradient, the networks may experience various problems when analyzing time series consisting of large data [19]. For these reasons, LSTM networks are distinguished from other methods as an alternative machine learning method. They are especially used for long term time series analysis and have emerged as a solution method for these problems [20].

Recently, DL-based machine learning algorithms have achieved high accuracy in time series predictions [21]. Among these algorithms, the LSTM is frequently used in the literature as one of the most successful methods [22]. LSTM(s) can analyze complex connections between time-series data features. Nevertheless, it is crucial for the data to be continuous in order to accurately discern the relationship between features and attain a high level of prediction accuracy [23].

Time series analysis has been successfully applied in many fields, such as construction [24], transportation systems [25], and energy forecasting [26]. Output power forecasting is valuable information for the continuous support of power grids [27]. A highly accurate forecasting model with a suitable performance curve provided by the manufacturer can help renewable-based power grids operate efficiently and safely [28]. In this paper, an LSTM-based DL architecture is proposed to predict wind turbine active power using wind turbine data as input. R^2, MAE, MSE, and RMSE metrics are used to measure the prediction performance and accuracy of the proposed method. There are limited studies in the literature on energy forecasting using LSTM-based architecture, which is a relatively new technique. Therefore, encouraged by the above findings, we aim to design an LSTM-based architecture to estimate the active output power with high accuracy. Here is a synopsis outlining the main contributions of this study.

- In the proposed architecture, high-accuracy power estimation is achieved by performing time series analysis.
- The effectiveness of the LSTM-based method in power estimation is demonstrated with statistical performance indicators.
- The actual power of a wind turbine data set obtained from real-world applications is estimated by time series analysis.

As for the rest of the paper, Section II summarizes the study in the literature for turbine energy prediction. Section III presents the data acquisition process; the preprocessing steps used for the study and the method used in this study are described in detail. Section IV presents the results and discussion. This section provides information about the experimental settings. Then the results of the proposed method are described. Section V, the concluding section, discusses the results of the study and concludes the paper with future work.

2. RELATED WORKS

Forecasting methods using machine learning-based models can be broadly divided into two categories: shallow learning and DL-based models [29]. In some shallow studies, wind energy prediction has been performed with fuzzy logic [30], wavelet analysis [31], and least squares support vector machine (LSSVM) [32]. Another shallow learning model, artificial neural networks (ANN), has the ability to capture the high correlation between data [33–35]. Sun et al. developed an ANN-based model to predict wind turbine active power. They considered environmental factors in network training. In their study, they concluded that differently positioned wind turbines should use different yaw angle strategies [36]. DL is a machine learning approach using ANNs [37].

DL, a subset of machine learning, is a relatively new technique developed to overcome the shortcomings of shallow learning models [38, 39] DL-based methods have been successfully applied to classification [40] and prediction problems [41]. LSTM, a variant of RNN, can learn time-series information more accurately. It is capable of efficiently utilizing temporal information to predict new data points [42]. It has been successfully used in stock market forecasting [43], natural language processing [44], and medicine [45].

Studies using LSTM-based methods for energy estimation are available in the literature [46]. An LSTM method with physical constraints was proposed by Luo et al. When compared to conventional statistical and machine learning techniques, the physically constrained LSTM model greatly increased prediction accuracy [47]. Chen et al. selected strongly associated features using the Pearson correlation coefficient. Features related to temperature, humidity, and solar radiation intensity were chosen for the LSTM model's input. They contrasted the time series method, radial basis function (RBF) neural networks, and back-propagation (BP) neural networks with the one-layer LSTM model. When compared to previous methods, their suggested model made predictions with a higher accuracy [48]. Zherui et al. used the LSTM model as a deep network model to predict wind power output with appropriate reliability. To enhance the prediction outcomes, they suggested a double decomposition-based remedial method [49]. In addition, related works based on chaotic time series, hybrid back-propagation, decomposition, and wavelet transforms have been investigated in the literature [50].

Most of the methods proposed in the literature for predicting turbine output power are traditional machine learning-based techniques. These methods have problems, such as requiring feature selection engineering and overcoming the problems of dealing with big data. In addition, the studies lack visualization of time series that can help in understanding and analyzing the problems while evaluating the data set. Our research focuses on the visualization and forecasting of wind power generation. The proposed architecture helps to make sense of the problems that can be encountered in the energy forecasting process with the help of data preprocessing and visualization methods.

3. MATERIALS AND METHOD

3.1. Data Pre-processing

To forecast wind power, the features that machine learning algorithms will use must be properly chosen. The environmental factors surrounding the wind turbine should be taken into account in this situation. Additionally, a thorough assessment of its effect on the wind turbine's active power generation is necessary. In this study, the dataset is provided by Kaggle [51]. Environmental factors such as wind speed and wind direction are used as inputs in the model. The dateset is obtained from a N117/3600 model wind turbine manufactured by Nordex. The SCADA system contains time series data of the wind turbine for one year (01.01.2018–31.12.2018) recorded in 10-minute periods. The dataset consists of 50530 units and five attributes: Wind speed (m/s), wind direction (°), theoretical power (kW), active power (kW), and Date/Time (Table 1).

TABLE I
DATASET DESCRIPTION

Feature	Description
Date/Time	10 minutes period.
LV Active Power (kW)	Power produced at that precise instant by the turbine.
Wind Speed (m/s)	Wind speed used by the turbine to generate electricity.
Theoretical Power Curve (kW)	The power expected to be generated by the turbine manufacturer at this wind speed.
Wind Direction (°)	Wind direction measured from the turbine hub.

3.2. Impact Factors Analysis

It is of great importance to assess and quantify the effects of the characteristics in the dataset on active wind energy production. Considering the impact of several variables on energy production, understanding the relationships between these factors is a critical requirement. A correlation matrix could be used to investigate the correlations between different variables for this purpose. Pearson correlation coefficient analysis can select the appropriate influence factors of the input data for the model. Thus, it can investigate the degree to which different impacts are correlated factors of the data and active power. The Pearson correlation coefficient can be calculated using Equation 1 [52].

$$r_{jk} = \frac{\sum_{i=1}^{n} (x_{ij} - \bar{x}_j)(x_{ik} - \bar{x}_k)}{\sqrt{\sum_{i=1}^{n} (x_{ij} - \bar{x}_j)^2}\sqrt{\sum_{i=1}^{n} (x_{ik} - \bar{x}_k)^2}} \quad (1)$$

Where, the variables x_{ij} and x_{ik} represent the *i* value of data for class *j* and class *k*, respectively. Similarly, $\bar{x}_j$ and $\bar{x}_k$ denote the arithmetic mean of the data for class *j* and class *k*, respectively. The heat map in Figure 1 illustrates the results of Pearson correlation coefficient analysis applied to the dataset. The matrix, which numerically expresses the relationship between input variables and active power, presents the effect of one variable on the other between -1 and +1. Figure 1 shows that the correlation between actual power and wind speed is the highest, approximately 0.9. It can be seen that power and wind direction are negatively correlated. The correlation coefficient value of the wind direction is -0.063, which is less than 0.1. Therefore, the degree of correlation is weak.

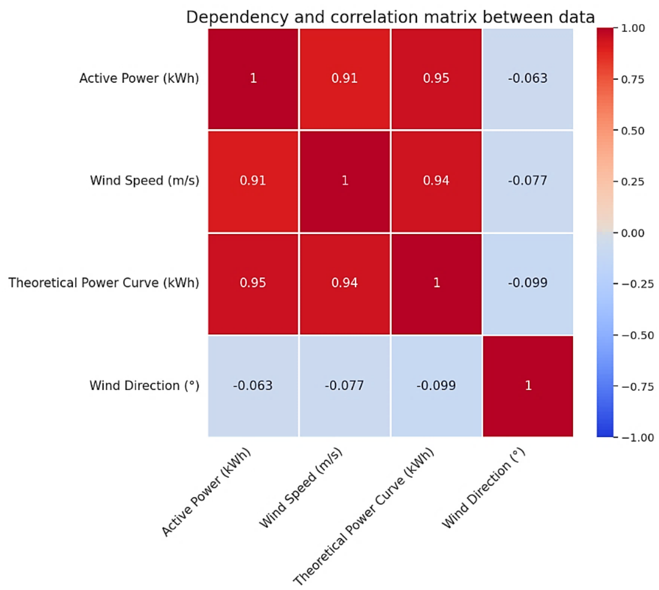

Figure 1. Pearson correlation matrix between active power and impact factors

3.3. Outlier Data Cleaning

One of the most important factors that negatively affects the performance of a model is outliers. Outliers can occur for various reasons. Outliers may occur in unexpected situations, such as wind outages and malfunctions. Due to these situations, it is difficult to obtain reliable wind power curves from raw wind data. For these reasons, it is necessary to extract these data [53]. A turbine only begins to produce electricity when the wind speed reaches the start-up value. The wind speed at which the machine generates its rated power is known as the "rated speed". In order to avoid failure and damage, electricity generation is halted when wind speeds reach high levels. Manufacturers can generate theoretical power curves under the assumption of perfect topographical and meteorological circumstances [54].

The study begins with the cleaning of outlier data. Then, the "LV ActivePower (kW)" feature is divided into sub-datasets in the range of 50 kW. This process is performed in increments of 50 between 20 and 3400 using a loop. At the end of this process, frames of 50 data points each are obtained. Since power generation starts when the wind speed reaches 3 m/s, this lower wind speed limit is taken as the starting value of power generation. 20 m/s is the upper wind limit of the turbine. After this speed, there will be no active power generation as the turbine will protect itself. After these filtering operations, outliers are removed from each sub-frame obtained. For this process, values other than 1.5 times the lower and upper quartiles of the data (q_{low} and q_{hi}) are considered outliers. Figure 2 shows the raw data set and wind speed graph. Figure 3 shows the plot of the cleaned data set obtained after the outliers are removed as a result of the data preprocessing described above.

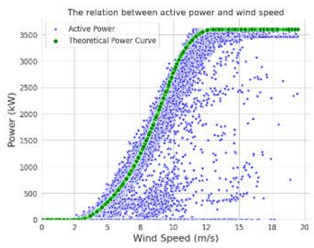

Figure 2. Relationship between actual power and wind speed in the raw data set

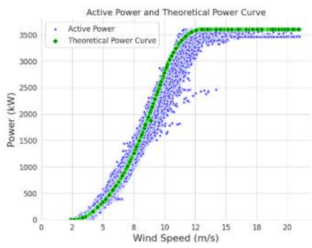

Figure 3. Relationship between actual power and wind speed in the preprocessed data set

At the end of the process, the sub-frames are merged to obtain a new data frame consisting of 37820 extracted data samples. Min-max normalization is applied to the input features to reduce the computational cost. At the end of normalization, the data range is compressed to [0, 1]. The normalization process is calculated using Equation 2 [55].

$$X_{scaled} = \frac{x_o - min(x)}{max(x) - min(x)} \quad (2)$$

Here, X_{scaled} is the normalized number, x_o is the original number, and *max(x)* and *min(x)* are the maximum and minimum numbers in the series, respectively.

3.4. LSTM Structure

The LSTM proposed by Hochreiter and Schmidhuber offers a solution to the problem of vanishing gradients in RNNs [56]. LSTM has a more complex structure than traditional RNNs, which includes cells and gates. An LSTM cell has the ability to preserve the temporal data from the earlier forecast and transmit this information to the network when needed [57]. The memory cell helps to preserve the temporal information of the previous prediction in the training of the LSTM and propagates it to the network when needed. Figure 4 shows the structure of a basic LSTM model.

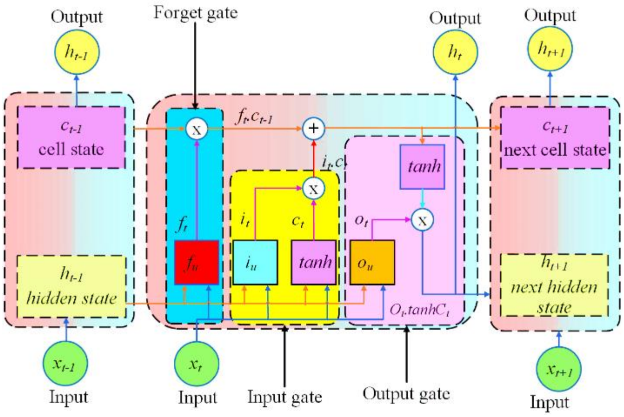

Figure 4. Basic structure of the LSTM model [57]

Compared with traditional RNN(s), the hidden layer of LSTM has more controllable units for information transfer to memory cells [58]. Three gates are added to the basic neural unit of the LSTM. These gates are input i_t, forget f_t, and output o_t. The gates take values in the interval [0, 1]. The primary role of the input gate is to update some attributes and determine the new attribute's content. The forget gate is designed to forget information that was previously useless. The output gate is used to determine what the output will be. All gates are connected at any time with the previous unit h_{t-1} and the current input x_t. Together, they decide the output. Below are the computational formulas for Equation (3) f_t, Equation (4) i_t, Equation (5) o_t, and the current neuron value, Equation (6) $\tilde{C}_t$ [59].

$$f_t = \sigma\left(W_{fx}x_t + W_{fh}h_{t-1} + b_f\right) \tag{3}$$

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

$$o_t = \sigma(W_{ox}x_t + W_{oh}h_{t-1} + b_o) \tag{5}$$

$$\tilde{C}_t = tanh(W_{Cx}x_t + W_{Ch}h_{t-1} + b_C) \tag{6}$$

Where W_{fx}, W_{fh}, W_{ix}, W_{ih}, W_{cx}, W_{ch}, W_{ox}, *and* W_{oh} are the matrix weights obtained by multiplying the current input value x_t by the previous unit output h_{t-1} of the relevant gate, respectively. b_f, b_i, b_c, and b_o represent the bias value and σ the sigmoid function. The input gate i_t, the forget gate f_t, the previous state value $\tilde{C}_{t-1}$, and the current neuron candidate value $\tilde{C}_t$ are used to calculate the new state value $\tilde{C}_{t+1}$. Equation (7, 8) can be used to determine the output value h_t after the new state value has been determined [59].

$$C_{t+1} = f_t * xC_{t-1} + i_t * \tilde{C}_t \tag{7}$$

$$h_t = o_t * tanh(S_t) \tag{8}$$

In this study, a multilayer LSTM network is designed to estimate active power. Table 2 shows the details of the designed architecture. The first layer contains 64 cell units and uses the ReLU activation function. The second layer contains 32 cell units and uses the ReLU activation function. The third and fourth layers contain a dense layer and an output dense layer, respectively. The dense layers are fully connected and contain 16 and 1, neuron. The output of the model produces a single numerical value estimate.

TABLE II
LSTM STRUCTURE PARAMETERS

Layer	Output shape	Parameter
LSTM	(0,0,64)	19200
LSTM	(0,0,32)	12416
Dense	(0,16)	528
Dense	(0,1)	17
Total Parameter		32,161

3.5. Error Metrics

A range of statistical techniques were employed to assess the DL-based architecture's prediction. In this context, Equation (9) adjusted R-squared (R^2), Equation (10) mean squared error (MSE), Equation (11) root mean squared error (RMSE), and Equation (12) mean absolute error (MAE) metrics were used to evaluate the discrepancy between predicted and actual values [60].

$$R^2 = \frac{\left(\sum_{i=1}^{N}(x_i^* - \overline{x_i^*})(x_i - \overline{x_i})\right)^2}{\sum_{i=1}^{N}(x_i^* - \overline{x_i^*})^2 \sum_{i=1}^{N}(x_i - \overline{x_i})^2} \tag{9}$$

In Equation 9, the R^2 value ranges from 0 to 1, with a higher value indicating a better predictive performance of the model. N is the number of data points, x is the dependent variable, x_i^* is the independent variable, $\overline{x_i^*}$ is the mean value of the independent variable, and x_i^* is the mean value of the dependent variable. In Equation 10, MSE is a statistical measure of how much error a regression model's predictions make relative to the actual data. In Equation 11, the standard deviation in prediction errors is represented by RMSE, and a lower value denotes a better model. In Equation 12, the absolute difference between the variables' expected and actual values is measured by the MAE [60, 61].

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(x_i - x_i^*)^2 \tag{10}$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i^* - x_i)^2} \tag{11}$$

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

Lower MSE, RMSE, and MAE values indicate that the model makes better predictions. For all three equations, N represents the number of data points, x_i represents the actual values, and x_i^* represents the expected output.

4. RESULTS AND DISCUSSION

This section analyzes the performance results obtained from the proposed LSTM-based DL model.

EUROPEAN JOURNAL OF TECHNIQUE, Vol.13, No.2, 2023

4.1. Experimental Settings

In this study, the data was tested using Python 3.10.12, TensorFlow 2.12, and a 64-bit system with a 2199 MHz 4-core processor and 32 GB of memory.

4.2. Hyperparameter and Optimization Techniques

The dataset was divided into training (60%), validation (20%), and testing (20%) subsets. According to this ratio, 22692, 7564, and 7564 were divided into three sets and used for training, validation, and testing, respectively. Hyperparameters are the settings that affect the performance results of the model. In order to determine these settings, the model was tested with different parameters, and the best-performing settings were selected. The learning coefficient of our model was initialized at a rate of 1e-3, and the coefficient was tried to be improved with the Adam optimizer. The training was set to 100 epochs. Table 3 shows the hyperparameters used for DL-based time series analysis.

TABLE III
TRAINING HYPERPARAMETERS

Hyperparameter	Parameter
Learning rate	1e-3
Optimizer	Adam
Batch size	32
Loss function	MSE
Number of epochs	100
Re-scaling	MinMaxScale [0,1]

Test data is used to evaluate the accuracy of the proposed prediction model. The regression graph obtained from the test data set using the LSTM architecture is shown in Figure 5. It is seen that the actual values and the values predicted by the architecture are gathered on the regression line. It is clear that the proposed architecture has high prediction accuracy.

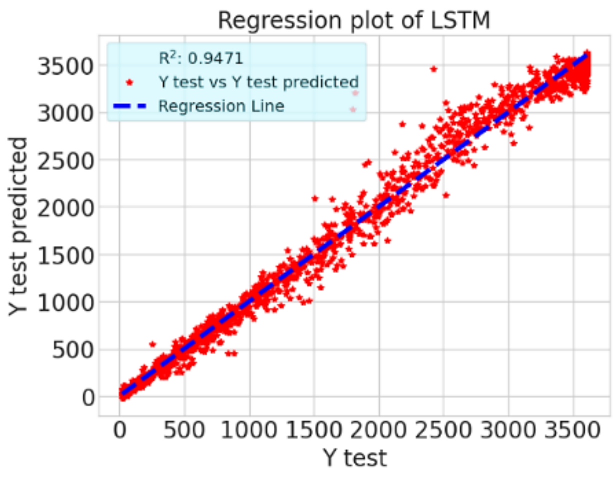

Figure 5. Regression plot of test dataset

The theoretical power curve is the graph of the power indicator expected from the turbine under ideal conditions. The prediction graph of the proposed model is consistent with the theoretical power curve graph. This shows that the model has good prediction performance. There is a direct proportionality between wind speed and actual power up to the turbine decommissioning speed point. Figure 6 shows the turbine's active power, the theoretical power, and the predicted power values obtained using the proposed method. When the graph is analyzed, the estimated power curve, the actual active power curve, and the theoretical power curve have a similar distribution.

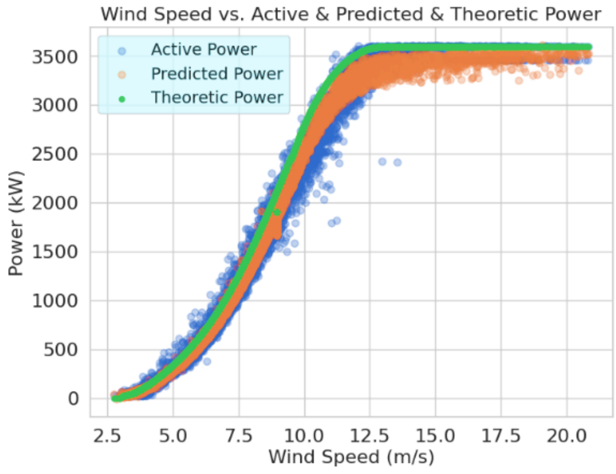

Figure 6. Graph of theoretical, active, predicted power and wind speed

Figure 7 shows the actual active power and the predicted power values by the proposed architecture for the date range 01.12.2018–05.12.2018 on the time axis graph. The actual data and the predicted data are given in the same figure. The proposed model performed well by overlapping with the actual value.

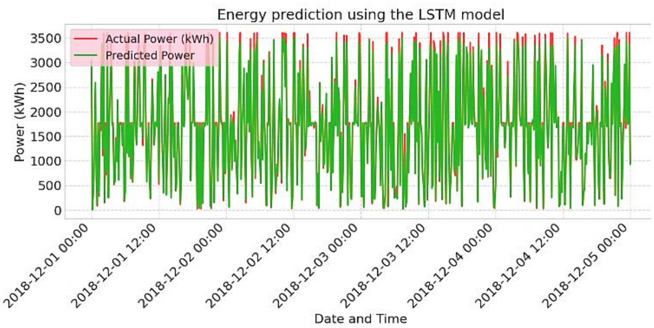

Figure 7. Time slice of predicted and active power

In this study, the performance of the model was evaluated according to the indicators described in Section 3.5. According to the results presented in Table 4, the proposed method has achieved high performance with an R^2 value of 96.10% on the training dataset. In addition, MAE, MSE, and RMSE values are 0.0190, 0.0034, and 0.0584, respectively. In addition, the proposed architecture achieved an R^2 score of 94.71% on the test dataset. This shows that the model is not overfitting and can capture the connection between the data in the newly encountered test dataset well. The MAE, MSE, and RMSE values in the test dataset are 0.0047, 0.0685, and 0.9471, respectively, and a good prediction result is obtained with low error metrics.

TABLE IV
PERFORMANCE RESULTS OF THE MODEL FROM THE TRAINING AND TEST DATASET

	Training Dataset	Testing Dataset
MAE	0.0190	0.0226
MSE	0.0034	0.0047
RMSE	0.0584	0.0685
R^2	0.9610	0.9471

Training time (s)	44.02

When the results are examined, the model is able to analyze the data well and shows a successful prediction capability. This is due to the ability of the LSTM-based machine learning method to capture long-term dependencies. Compared to classical machine learning-based methods, LSTM uses a special mechanism called memory cells. The cells have the ability to store previous knowledge and use it later. This allows the model to make predictions based on previous data.

5. CONCLUSION

Wind energy forecasting is an important component of energy management systems. In this study, an LSTM-based architecture for active power energy forecasting is proposed using time series data from a wind turbine. The anomalous data in the dataset is extracted by dividing it into frames. Then the cleaned data is used to feed the LSTM-based architecture. The results and performance metrics show the high success rate of the model. LSTM is a method with high prediction performance, especially in large datasets, due to its ability to capture long-term dependencies. By utilizing this, the proposed DL-based LSTM method has achieved high prediction accuracy.

For energy forecasting, LSTM-based methods can be used to achieve high accuracy in forecasting. However, the result can be improved by using different architectures. In addition, the LSTM model is a computationally expensive method due to its complexity. In our study, we used a multilayer LSTM model. These architectures are capable of successfully capturing complex relationships between data. However, increasing the number of layers may increase the computational cost. In future work, we plan to design fewer-layer architectures without degrading performance. In this way, we aim to reduce the computational cost.

REFERENCES

[1] M. Saglam, C. Spataru, and O. A. Karaman, “Electricity demand forecasting with use of artificial intelligence: The case of Gokceada Island,” Energies, vol. 15, no. 16, p. 5950, 2022. https://doi.org/10.3390/en15165950

[2] Ş. Fidan and H. Çimen, “Rüzgâr Türbinlerinde Tork ve Kanat Eğim Açısı Kontrolü,” Batman Üniversitesi Yaşam Bilimleri Dergisi, vol. 11, pp. 12–26, 2021. Retrieved from https://dergipark.org.tr/en/pub/buyasambid/issue/63446/880791

[3] Yilmaz, M. (2018). Real measure of a transmission line data with load fore-cast model for the future. Balkan Journal of Electrical and Computer Engineering, 6(2), 141-145. https://doi.org/10.17694/bajece.419646

[4] Yilmaz, M. (2017, March). The Prediction of Electrical Vehicles' Growth Rate and Management of Electrical Energy Demand in Turkey. In 2017 Ninth annual IEEE green technologies conference (GreenTech) (pp. 118-123). IEEE. https://doi.org/10.1109/GreenTech.2017.23

[5] M. Saglam, C. Spataru, and O. A. Karaman, “Forecasting electricity demand in Turkey using optimization and machine learning algorithms,” Energies, vol. 16, no. 11, p. 4499, 2023. https://doi.org/10.3390/en16114499

[6] Z. Niu, Z. Yu, W. Tang, Q. Wu, and M. Reformat, “Wind power forecasting using attention-based gated recurrent unit network,” Energy (Oxf.), vol. 196, no. 117081, p. 117081, 2020. https://doi.org/10.1016/j.energy.2020.117081

[7] L. Donadio, J. Fang, and F. Porté-Agel, “Numerical weather prediction and artificial neural network coupling for wind energy forecast,” Energies, vol. 14, no. 2, p. 338, 2021. https://doi.org/10.3390/en14020338

[8] S. Hanifi, X. Liu, Z. Lin, and S. Lotfian, “A critical review of wind power forecasting methods—past, present and future,” Energies, vol. 13, no. 15, p. 3764, 2020. https://doi.org/10.3390/en13153764

[9] L. Liu and Y. Liang, “Wind power forecast optimization by integration of CFD and Kalman filtering,” Energy Sources Recovery Util. Environ. Eff., vol. 43, no. 15, pp. 1880–1896, 2021. https://doi.org/10.1080/15567036.2019.1668080

[10] J. M. González-Sopeña, V. Pakrashi, and B. Ghosh, “An overview of performance evaluation metrics for short-term statistical wind power forecasting,” Renew. Sustain. Energy Rev., vol. 138, no. 110515, p. 110515, 2021. https://doi.org/10.1016/j.rser.2020.110515

[11] V. Cerqueira, L. Torgo, and C. Soares, “Machine learning vs statistical methods for time series forecasting: Size matters,” 2019. https://doi.org/10.48550/arXiv.1909.13316

[12] I. Aydin, S. B. Celebi, S. Barmada, and M. Tucci, “Fuzzy integral-based multi-sensor fusion for arc detection in the pantograph-catenary system,” Proc. Inst. Mech. Eng. Pt. F: J. Rail Rapid Transit, vol. 232, no. 1, pp. 159–170, 2018. https://doi.org/10.1177/0954409716662090

[13] Ö. A. Karaman, “Performance evaluation of seasonal solar irradiation models—case study: Karapınar town, Turkey,” Case Stud. Therm. Eng., vol. 49, no. 103228, p. 103228, 2023. https://doi.org/10.1016/j.csite.2023.103228

[14] A. Çalışkan, S. Demirhan, and R. Tekin, “Comparison of different machine learning methods for estimating compressive strength of mortars,” Constr. Build. Mater., vol. 335, no. 127490, p. 127490, 2022. https://doi.org/10.1016/j.conbuildmat.2022.127490

[15] Öztekin, A., & Erçelebi, E. (2019). An efficient soft demapper for APSK signals using extreme learning machine. Neural Computing and Applications, 31, 5715-5727. https://doi.org/10.1007/s00521-018-3392-6

[16] S. B. Çelebi and B. G. Emiroğlu, “Leveraging deep learning for enhanced detection of Alzheimer’s disease through morphometric analysis of brain images,” Trait. Du Signal, vol. 40, no. 4, pp. 1355–1365, 2023. https://doi.org/10.18280/ts.400405

[17] T. Chen and C. Guestrin, “XGBoost: A Scalable Tree Boosting System,” in Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2016. https://doi.org/10.1145/2939672.2939785

[18] Demir, S., & Sahin, E. K. (2023). Predicting occurrence of liquefaction-induced lateral spreading using gradient boosting algorithms integrated with particle swarm optimization: PSO-XGBoost, PSO-LightGBM, and PSO-CatBoost. Acta Geotechnica, 18(6), 3403-3419. https://doi.org/10.1007/s11440-022-01777-1

[19] S. Hochreiter, “The vanishing gradient problem during learning recurrent neural nets and problem solutions,” Internat. J. Uncertain. Fuzziness Knowledge-Based Systems, vol. 06, no. 02, pp. 107–116, 1998. https://doi.org/10.1142/S0218488598000094

[20] Yilmaz, A., & Poli, R. (2022). Successfully and efficiently training deep multi-layer perceptrons with logistic activation function simply requires initializing the weights with an appropriate negative mean. Neural Networks, 153, 87-103. https://doi.org/10.1016/j.neunet.2022.05.030

[21] J. F. Torres, A. Galicia, A. Troncoso, and F. Martínez-Álvarez, “A scalable approach based on deep learning for big data time series forecasting,” Integr. Comput. Aided Eng., vol. 25, no. 4, pp. 335–348, 2018. https://doi.org/10.3233/ICA-180580

[22] J. Zhang, J. Yan, D. Infield, Y. Liu, and F.-S. Lien, “Short-term forecasting and uncertainty analysis of wind turbine power based on long short-term memory network and Gaussian mixture model,” Appl. Energy, vol. 241, pp. 229–244, 2019. https://doi.org/10.1016/j.apenergy.2019.03.044

[23] S. Zhang, Y. Wang, M. Liu, and Z. Bao, “Data-based line trip fault prediction in power systems using LSTM networks and SVM,” IEEE Access, vol. 6, pp. 7675–7686, 2018. https://doi.org/10.1109/ACCESS.2017.2785763.

[24] S. Fidan, H. Oktay, S. Polat, and S. Ozturk, “An artificial neural network model to predict the thermal properties of concrete using different neurons and activation functions,” Adv. Mater. Sci. Eng., vol. 2019, pp. 1–13, 2019. https://doi.org/10.1155/2019/3831813

[25] I. Aydin, O. Yaman, M. Karakose, and S. B. Celebi, “Particle swarm based arc detection on time series in pantograph-catenary system,” in 2014 IEEE International Symposium on Innovations in Intelligent Systems and Applications (INISTA) Proceedings, 2014. https://doi.org/10.1109/INISTA.2014.6873642

[26] S. F. Stefenon, L. O. Seman, V. C. Mariani, and L. dos S. Coelho, “Aggregating prophet and seasonal trend decomposition for time series forecasting of Italian electricity spot prices,” Energies, vol. 16, no. 3, p. 1371, 2023. https://doi.org/10.3390/en16031371

[27] M. Ş. Üney and Ö. A. Karaman, "Load Frequency Control (LFC) of a Microgrid using PSCAD/EMTDC Simulation Program," Adıyaman Üniversitesi Mühendislik Bilimleri Dergisi, vol. 8, no. 15, pp. 328–342, 2021. https://doi.org/10.54365/adyumbd.939716

[28] Ş. Fidan, M. Cebeci, and A. Gündoğdu, "Extreme Learning Machine Based Control of Grid Side Inverter for Wind Turbines," Tehnički vjesnik, vol. 26, pp. 1492–1498, 2019. https://doi.org/10.17559/TV-20180730143757

[29] B. Zazoum, "Solar photovoltaic power prediction using different machine learning methods," Energy Rep., vol. 8, pp. 19–25, 2022. https://doi.org/10.1016/j.egyr.2021.11.183

[30] W. Zou, C. Li, and P. Chen, "An inter type-2 FCR algorithm based T–S fuzzy model for short-term wind power interval prediction," IEEE Trans. Industr. Inform., vol. 15, no. 9, pp. 4934–4943, 2019. doi: 10.1109/TII.2019.2910606.

[31] P. Du, J. Wang, W. Yang, and T. Niu, "A novel hybrid model for short-term wind power forecasting," Appl. Soft Comput., vol. 80, pp. 93–106, 2019. https://doi.org/10.1016/j.asoc.2019.03.035

[32] X. Yuan, Q. Tan, X. Lei, Y. Yuan, and X. Wu, "Wind power prediction using hybrid autoregressive fractionally integrated moving average and least square support vector machine," Energy (Oxf.), vol. 129, pp. 122–137, 2017. https://doi.org/10.1016/j.energy.2017.04.094

[33] S. Agatonovic-Kustrin and R. Beresford, "Basic concepts of artificial neural network (ANN) modeling and its application in pharmaceutical research," J. Pharm. Biomed. Anal., vol. 22, no. 5, pp. 717–727, 2000. https://doi.org/10.1016/S0731-7085(99)00272-1

[34] B. Birecikli, Ö. A. Karaman, S. B. Çelebi, and A. Turgut, "Failure load prediction of adhesively bonded GFRP composite joints using artificial neural networks," J. Mech. Sci. Technol., vol. 34, no. 11, pp. 4631–4640, 2020. https://doi.org/10.1007/s12206-020-1021-7

[35] H. A. N. Kubilay, G. Öztürk, and A. Aslan, "Yapay Sinir Ağları Kullanarak Yüzey Pürüzlülüğü Tespiti," International Conference on Pioneer and Innovative Studies, vol. 1, pp. 487–492, 2023.

[36] H. Sun, C. Qiu, L. Lu, X. Gao, J. Chen, and H. Yang, "Wind turbine power modelling and optimization using artificial neural network with wind field experimental data," Appl. Energy, vol. 280, no. 115880, p. 115880, 2020. https://doi.org/10.1016/j.apenergy.2020.115880

[37] Yilmaz, A., Simsek, C., Tozlu, B. H., Aydemir, O., & Karaveliogl u, Y. (2022). Selection of suitable sensors of the electronic nose used for classification of myocardial infarction, stable coronary artery disease and healthy individuals. Selcuk University Journal of Engineering Sciences, 21(1), 39-43. https://sujes.selcuk.edu.tr/sujes/article/view/597

[38] S. B. Çelebi and B. G. Emiroğlu, "A novel deep dense block-based model for detecting Alzheimer's disease," Appl. Sci. (Basel), vol. 13, no. 15, p. 8686, 2023. https://doi.org/10.3390/app13158686

[39] Şimşek, C., Yilmaz, A., Tozlu, B. H., Aydemir, Ö., & Karavelioğlu, Y. (2022). Classification of Cardiovascular Diseases Using Electronic Nose Dataset with Artificial Neural Network Classifier. Avrupa Bilim ve Teknoloji Dergisi, (38), 479-483. https://doi.org/10.31590/ejosat.1165991

[40] S. B. Çelebi and B. G. Emiroğlu, "Alzheimer Teşhisi için Derin Öğrenme Tabanlı Morfometrik Analiz," Iğdır Üniversitesi Fen Bilimleri Enstitüsü Dergisi, vol. 13, no. 3, pp. 1454–1467, 2023. https://doi.org/10.21597/jist.1275669

[41] F. Emmert-Streib, Z. Yang, H. Feng, S. Tripathi, and M. Dehmer, "An introductory review of deep learning for prediction models with big data," Front. Artif. Intell., vol. 3, 2020. https://doi.org/10.3389/frai.2020.00004

[42] R. DiPietro and G. D. Hager, "Deep learning: RNNs and LSTM," in Handbook of Medical Image Computing and Computer Assisted Intervention, Elsevier, 2020, pp. 503–519. https://doi.org/10.1016/B978-0-12-816176-0.00026-0

[43] S. Zaheer et al., "A multi parameter forecasting for stock time series data using LSTM and deep learning model," Mathematics, vol. 11, no. 3, p. 590, 2023. https://doi.org/10.3390/math11030590

[44] M. Fazil, S. Khan, B. M. Albahlal, R. M. Alotaibi, T. Siddiqui, and M. A. Shah, "Attentional multi-channel convolution with bidirectional LSTM cell toward hate speech prediction," IEEE Access, vol. 11, pp. 16801–16811, 2023. https://doi.org/10.1109/ACCESS.2023.3246388.

[45] V. Rai, G. Gupta, S. Joshi, R. Kumar, and A. Dwivedi, "LSTM-based adaptive whale optimization model for classification of fused multimodality medical image," Signal Image Video Process., vol. 17, no. 5, pp. 2241–2250, 2023. https://doi.org/10.1007/s11760-022-02439-1

[46] Ö. A. Karaman, "Prediction of wind power with machine learning models," Appl. Sci. (Basel), vol. 13, no. 20, p. 11455, 2023. https://doi.org/10.3390/app132011455

[47] X. Luo, D. Zhang, and X. Zhu, "Deep learning based forecasting of photovoltaic power generation by incorporating domain knowledge," Energy (Oxf.), vol. 225, no. 120240, p. 120240, 2021. https://doi.org/10.1016/j.energy.2021.120240

[48] H. Chen and X. Chang, "Photovoltaic power prediction of LSTM model based on Pearson feature selection," Energy Rep., vol. 7, pp. 1047–1054, 2021. https://doi.org/10.1016/j.egyr.2021.09.167

[49] Z. Ma et al., "Application of hybrid model based on double decomposition, error correction and deep learning in short-term wind speed prediction," Energy Convers. Manag., vol. 205, no. 112345, p. 112345, 2020. https://doi.org/10.1016/j.enconman.2019.112345

[50] T. Ouyang, H. Huang, Y. He, and Z. Tang, "Chaotic wind power time series prediction via switching data-driven modes," Renew. Energy, vol. 145, pp. 270–281, 2020. https://doi.org/10.1016/j.renene.2019.06.047

[51] Kaggle.com. [Online]. Available: https://www.kaggle.com/datasets/berkerisen/wind-turbine-scada-dataset,. [Accessed: 28-Oct-2023]

[52] P. Schober, C. Boer, and L. A. Schwarte, "Correlation coefficients: Appropriate use and interpretation," Anesth. Analg., vol. 126, no. 5, pp. 1763–1768, 2018. https://doi.org/10.1213/ANE.0000000000002864

[53] V. Hodge and J. Austin, "A survey of outlier detection methodologies," Artif. Intell. Rev., vol. 22, no. 2, pp. 85–126, 2004. https://doi.org/10.1023/B:AIRE.0000045502.10941.a9

[54] S. Shokrzadeh, M. Jafari Jozani, and E. Bibeau, "Wind turbine power curve modeling using advanced parametric and nonparametric methods," IEEE Trans. Sustain. Energy, vol. 5, no. 4, pp. 1262–1269, 2014. https://doi.org/10.1109/TSTE.2014.2345059.

[55] S. G. K. Patro and K. K. Sahu, "Normalization: A Preprocessing Stage," 2015. https://doi.org/10.48550/arXiv.1503.06462

[56] Bilal, M. A., Wang, Y., Ji, Y., Akhter, M. P., & Liu, H. (2023). Earthquake Detection Using Stacked Normalized Recurrent Neural Network (SNRNN). *Applied Sciences*, *13*(14), 8121. https://doi.org/10.3390/app13148121

[57] S. Hochreiter and J. Schmidhuber, "Long short-term memory," Neural Comput., vol. 9, no. 8, pp. 1735–1780, 1997. https://doi.org/10.1162/neco.1997.9.8.1735.

[58] F. Shahid, A. Zameer, and M. Muneeb, "A novel genetic LSTM model for wind power forecast," Energy (Oxf.), vol. 223, no. 120069, p. 120069, 2021. https://doi.org/10.1016/j.energy.2021.120069

[59] A. T. Mohan and D. V. Gaitonde, "A deep learning based approach to reduced Order Modeling for turbulent flow control using LSTM neural networks," 2018. https://doi.org/10.48550/arXiv.1804.09269

[60] T. Chai and R. R. Draxler, "Root mean square error (RMSE) or mean absolute error (MAE)-Arguments against avoiding RMSE in the literature," Geoscientific model development, vol. 7, no. 3, pp. 1247–1250, 2014. https://doi.org/10.5194/gmd-7-1247-2014

[61] Öztekin, A., & Erçelebi, E. (2016). An early split and skip algorithm for fast intra CU selection in HEVC. Journal of Real-Time Image Processing, 12, 273-283. https://doi.org/10.1007/s11554-015-0534-2

BIOGRAPHIES

Selahattin Barış Çelebi, received his bachelor's degree in computer engineering from Selçuk University, Konya. He received his master's degree in computer engineering from the department of computer engineering at Fırat University, Elazığ. Subsequently, he obtained his doctoral degree in computer engineering from Kırıkkale University, Kırıkkale. His research interests are in the fields of artificial intelligence, data science, machine learning, image processing.

Ömer Ali Karaman received the M.Sc. in DTC control of asynchronous motor in 2010 and the Ph.D. degree in harmonics active power filter in Fırat University in 2018. He is a Full Asst.Professor in Electronics and automation department in Batman University. His research focuses on renewable energy systems, energy and solar radiation forecasting.