Journal of Computer Science and Digital Technology, 2025, 1(1), 33-42
https://doi.org/10.71204/rpvp8658

Wind Farm Forecasting Based on MLSTM

Peng Zhao [1,2], Junye Yang [1,2,*], Senhao Zhang [1,2]

[1] School of Computer Science and Technology, Taiyuan Normal University, Jinzhong 030619, China; 1721862373@qq.com

[2] Shanxi Key Laboratory of Intelligent Optimization Computing and Blockchain Technology, Taiyuan Normal University, Jinzhong 030619, China; 1721862373@qq.com

*** Correspondence:**
Junye Yang

1721862373@qq.com

Received: 5 April 2025 /Accepted: 15 April 2025 /Published online: 21 April 2025

Abstract

With the increasing importance of wind energy in the global energy structure, wind power forecasting has become one of the key technologies to ensure grid stability and improve energy dispatch efficiency. This paper uses the wind speed and wind power data of US wind farms in 2012 to predict and compare wind speed and wind power of different LSTM variant models (including traditional LSTM, xLSTM, sLSTM and mLSTM) by sampling every 5 minutes. The research focuses on comparing the performance of each model in predicting wind speed under the same power output conditions. The experimental results are evaluated by three common evaluation indicators: MAE , MSE and R^2. The results show that the mLSTM model performs best in wind speed forecasting, with better accuracy and stability than other LSTM variants. The research in this paper provides a new method to improve the accuracy of wind power forecasting and provides effective decision support for the operation and management of wind farms.

Keywords: Wind Power Prediction; Wind Speed; LSTM; mLSTM; Time Series Prediction

1. Introduction

As a clean and renewable energy source, wind power has become an important part of the global energy transformation. With its abundant wind energy resources, the wind power industry plays an increasingly important role in the energy structure of each country. In order to optimize the operation of wind farms, ensure the stability of the power grid and improve the efficiency of power dispatching, accurate prediction of wind power has become a hot topic of research. However, due to the nonlinear relationship between wind speed and wind power and the complex time series characteristics, traditional wind power prediction methods such as regression analysis and support vector machine (SVM) face great challenges.

In recent years, long short-term memory network (LSTM) has been widely used in wind power prediction due to its excellent time series data modeling ability. LSTM can effectively capture the dependencies in long time series and avoid the gradient vanishing problem encountered by traditional neural networks when processing time series data. Although LSTM has achieved certain results in wind power prediction, its model performance still has room for improvement. To solve this problem, this paper proposes a multi-layer LSTM (mLSTM) model and compares it with traditional LSTM, xLSTM and sLSTM to improve the accuracy of wind speed and wind power prediction.

This study uses wind speed and wind power data from US wind farms in 2012 to focus on analyzing the prediction effects of mLSTM and other LSTM variant models on wind speed (especially wind speed at 100 meters, in m/s) under the same power output conditions. By using evaluation indicators such as MAE, MSE, and R^2, we found that the mLSTM model performed best in wind speed prediction. The results of this study provide a more accurate prediction method for the efficient operation of wind farms and provide new ideas for the application of deep learning in the wind power field.

2. Related Work

The literature on wind power plant prediction encompasses a variety of methodologies and advancements aimed at improving the accuracy and reliability of forecasting wind energy generation. A significant focus has been on the application of artificial intelligence techniques, particularly neural networks, to enhance prediction capabilities.

The operational challenges posed by the variability of wind energy generation have also been addressed in the literature. Duque et al (2011) discussed the optimization of hydro-pumped storage systems to mitigate imbalances caused by wind power fluctuations, indicating the necessity of integrating different energy sources to ensure grid stability. Liu et al (2012) introduced a method for short-term wind power prediction utilizing neural networks trained on historical data, specifically wind speed and direction. This foundational work set the stage for subsequent studies that aimed to refine and adapt neural network approaches for various contexts. For instance, Lungu et al (2016) developed a two-step forecasting solution tailored for small wind farms in hilly regions of Romania, emphasizing the need for accurate hourly predictions to meet regulatory requirements.In addition to neural networks, other innovative techniques have emerged. Ghoushchi et al (2021) presented an extended fuzzy wavelet neural network approach, which incorporates weather and power plant parameters to predict output. This method showcases the versatility of fuzzy logic in addressing the complexities of wind power forecasting.Advancements in forecasting techniques have also been explored through hybrid methodologies. Pang et al (2021) proposed a hybrid forecasting method that integrates multiple renewable energy sources, demonstrating improved accuracy over traditional methods in a clustered renewable energy system. This approach highlights the importance of considering the interplay between different energy generation types in enhancing overall forecasting performance.Moreover, the integration of environmental features into forecasting models has been a notable trend. Wang et al (2021) explored ultra-short-term wind power forecasting by employing environmental feature decomposition and reconstruction techniques, specifically using the Ensemble Empirical Mode Decomposition (EEMD) method combined with an improved Random Forest algorithm. This study underscores the potential of advanced data processing techniques to enhance prediction accuracy.LSTM (Long Short-Term Memory Network) is a neural network model designed specifically for time series data, which can effectively capture long-term dependencies in data. For wind power forecasting, data such as wind speed and wind direction usually show obvious time series characteristics, and MLSTM can learn this time series pattern better than traditional neural networks or other regression models, thereby improving the accuracy of prediction. Wind power generation is not only affected by wind speed and wind direction, but also by a variety of environmental factors such as temperature, air pressure, and humidity. The MLSTM model has a strong nonlinear fitting ability and can learn these complex relationships through a multi-layer structure, which is more advantageous than traditional linear or shallow neural network methods (such as the simple neural network in Liu et al (2012)). By introducing a multi-layer structure, the MLSTM model can model at different time scales. For example, short-term wind speed fluctuations and long-term climate change patterns can be effectively captured in the same model. Some traditional methods, such as the integrated method of multiple energy sources proposed by Pang et al (2021), can improve prediction accuracy, but usually rely more on manually designed features, while MLSTM automatically extracts features at different time scales through self-learning. Since the generation of wind energy is greatly affected by weather conditions, the output of wind farms is highly uncertain and volatile. The MLSTM model can adaptively update its internal state to handle this uncertainty, especially in the face of environmental changes or sudden changes in wind speed. Its dynamic update capability is more flexible and efficient than traditional rule-based or experience-based models (such as the environmental feature decomposition method of Wang et al (2021)). The MLSTM model works well in processing single wind farm forecasts, and can also show good scalability in multi-wind farm system or cluster forecasts. For example, for the problem of integrating multiple renewable energy sources mentioned in Pang et al (2021), MLSTM can handle the synergy between different energy sources by adjusting the network structure and input data without the need for additional complex feature engineering or manually designed combination methods.

3. Data Description

The data set used in this paper comes from the 2012 wind speed data and wind power data of the US wind farm, which are sampled every 5 minutes as shown in Table 1. The data is sampled every 5 minutes and contains records of wind speed (unit: m/s) and wind power (unit: MW). In order to analyze the relationship between wind speed and wind power, the wind speed in the data set is measured at an altitude of 100 meters.

Table1. Table showing some data sets

Year	Month	Day	Hour	Minute	power (MW)	wind speed at 100m (m/s)
2012	1	1	0	0	15.998000000000001	14.569
2012	1	1	0	5	16	14.908
2012	1	1	0	10	16	15.173
					...	
2012	1	1	0	15	16	15.149000000000001
2012	1	1	0	45	16	15.463000000000001

Table2. Data Field Description

Fields	Field Description
Year、Month、Day、Hour、Minute	Year, month, day, hour, minute
power (MW)	Electricity generated
wind speed at 100m (m/s)	Wind speed

4. Model Introduction

The multiplicative long short-term memory network (mLSTM) is an important recurrent neural network architecture that plays a key role in sequence modeling tasks, especially in areas such as time series. The mLSTM used in this paper is a hybrid model that combines the long short-term memory (LSTM) and multiplicative recurrent neural network architectures. Its structure can be regarded as a specially designed directed graph that processes sequence data through a gating mechanism and a unique hidden layer transformation method. When building an mLSTM model, the first thing involved is the initialization of parameters, including the setting of various weight matrices and biases. Then, the input sequence data will enter the network in sequence for processing according to the time step.

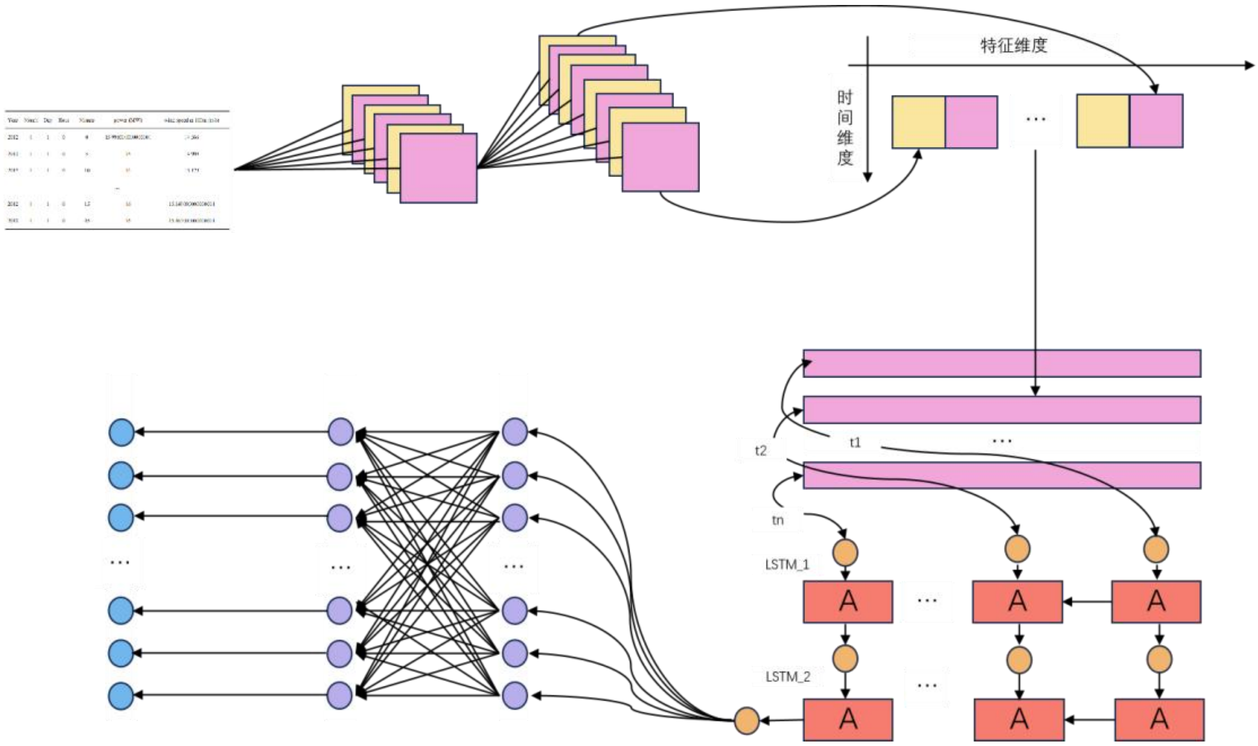

Figure 1. LSTM architecture

Input layer: The input layer receives input data for each time step in the sequence. Each input element corresponds to a node in the input layer, which passes the data to the subsequent hidden layer

Hidden layer: The hidden layer of mLSTM plays a core role in the entire model. It combines the gating unit of LSTM and the factorized hidden weight mechanism of multiplicative RNN (mRNN). mLSTM has input gates, output gates, and forget gates similar to LSTM, and these gating units control the flow of information between the internal states of the network. At the same time, the intermediate state of mRNN is introduced m_t:

$$m_t = (W_{mx}x_t) \odot (W_{mh}h_{t-1}) \tag{1}$$

It is connected to each gating unit, adding flexibility to the calculation of the hidden layer. For example, when calculating the input gate i_t, output gate O_t, and forget gate f_t, in addition to receiving input from the input layer x_t and the previous hidden state h_{t-1}, it also combines the intermediate state m_t for calculation, that is:

$$f_t = \sigma(W_{ix}x_t + W_{im}m_t) \tag{2}$$

$$o_t = \sigma(W_{ox}x_t + W_{om}m_t) \tag{3}$$

$$f_t = \sigma(W_{fx}x_t + W_{fm}m_t) \tag{4}$$

The update of the internal state c_t also depends on these gated units and intermediate states:

$$c_t = f_t \odot c_{t-1} + i_t \odot tanh\left(\hat{h}_t\right) \tag{5}$$

The final output of the hidden state is $h_t = tanh\,(c_t) \odot c_t$, in:

$$\hat{h}_t = W_{xt}x_t + W_{hm}m_t \quad (6)$$

This complex computational structure enables mLSTM to better capture long-term dependencies in the sequence while flexibly transforming according to different inputs.

Output layer: The output layer generates corresponding outputs according to specific tasks. When training mLSTM models, the commonly used method is the gradient-based optimization algorithm. First, the input sequence data is fed into the network for forward propagation to obtain the predicted output of the network. Then, the predicted output is compared with the true label to calculate the loss function, and the commonly used loss function is the cross entropy loss function. Next, the gradient of each parameter to the loss is calculated through the back-propagation algorithm, and the weights and biases are updated according to the gradient to minimize the loss function. During the training process, some regularization techniques, such as variational dropout, can also be used to prevent overfitting and improve the generalization ability of the model. mLSTM shows unique advantages in sequence modeling tasks. It can provide different cyclic transformation functions for each possible input, making it more expressive in autoregressive density estimation.

5. Experimental Analysis

The experiment used the 2012 wind speed and wind power dataset of US wind farms, with the goal of time series forecasting. After the data was loaded, the data was normalized using MinMaxScaler to scale it to the range of [0,1]. Normalization can help the model converge faster and improve training results, especially when using neural networks. Then, the dataset was divided into a training set and a test set, with 80% used as training data and 20% used as test data. This is the standard training-testing dataset division method.

Next, the dataset was converted into input sequences and target outputs using a sliding window method. Each input sequence contains 9 time steps of data (the first 8 data points), and the target output is the next 9th data point. This processing method helps the model learn the rules and trends in time series data.

During the training process, all models used mean squared error (MSE) as the loss function and used Adam optimizer for optimization. During training, each model iterated on the training set multiple times and gradually adjusted the model parameters to minimize the loss function.

During training, the training loss is calculated once for each epoch, and the parameters of the model are updated through backpropagation. The training loss curve of each model is plotted to help observe the convergence of the model during training. By comparing the training losses of different models, it is possible to evaluate which models are easier to converge and which models have better training effects. After the model training is completed, each model is evaluated using the test set. During the evaluation process, the model predicts the test data based on the knowledge learned during the training process, and compares the predicted results with the actual test data. In order to ensure the fairness of the evaluation, the prediction results of all models will be inverse normalized to convert the data back to the original range for intuitive comparison.

During the evaluation process, a comparison chart of the model's prediction results and the real data is plotted to help observe the model's prediction effect. The figure shows the prediction curve and the actual data curve of each model, further verifying the prediction ability of each model. As shown in Figure 2.

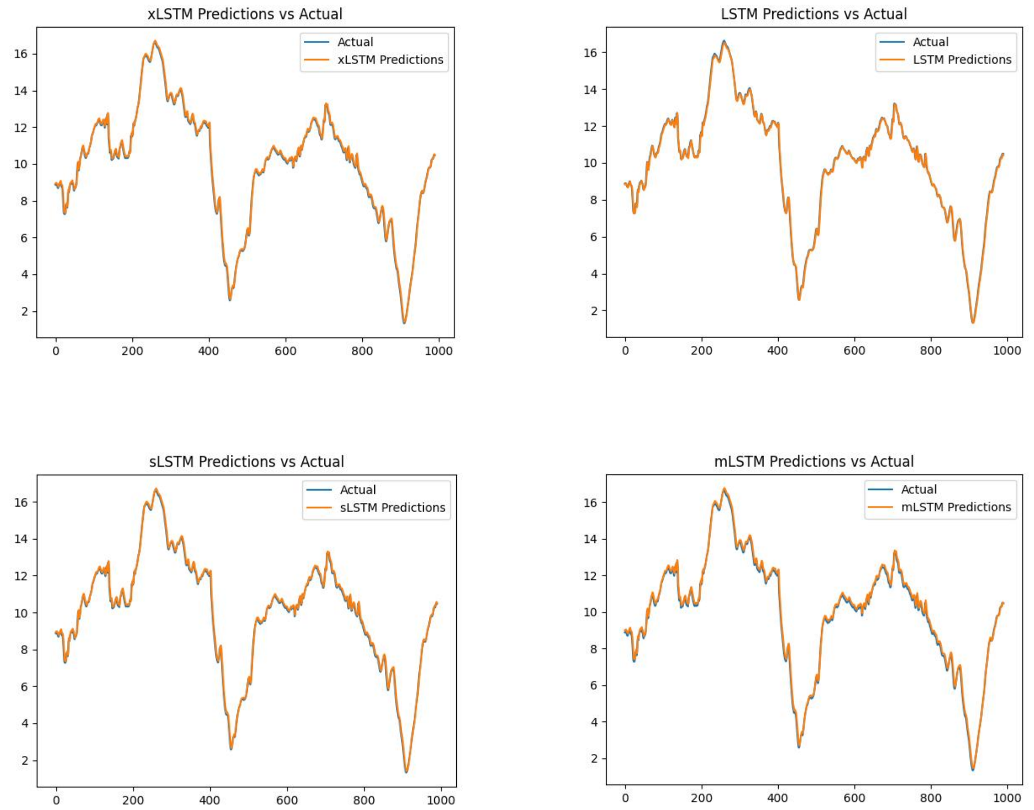

Figure 2. Data visualization of xLSTM, LSTM, sLSTM, and mLSTM prediction results

In this experiment, we used four different types of LSTM models (xLSTM, LSTM, sLSTM, mLSTM) for time series forecasting and evaluated the performance of the models using mean square error (MSE), mean absolute error (MAE), and coefficient of determination (R^2).

Table3. MAE, MSE and R^2 comparison tablea

Model	MAE	MSE
XLSTM	0.876338891355896	0.7941865921020508
LSTM	0.11889254301786423	0.022091850638389587
SLSTM	0.1066385880112648	0.020274749025702477
MLSTM	0.10501305013895035	0.019756944850087166

The experimental results show that the mLSTM model performs best, with the lowest MAE and MSE, and its R^2 value is close to 1, indicating that mLSTM has the strongest prediction ability in this experiment and is the most suitable model choice. mLSTM performs very well in terms of both prediction accuracy and fitting ability. The LSTM and sLSTM models also perform very well, with low MAE and MSE values, and R^2 close to 1, which is excellent and can almost achieve the best prediction effect. Compared with mLSTM, the performance of sLSTM and LSTM is slightly inferior, but they are still strong competitors. Although the xLSTM model has a high R^2 value, indicating that the model can capture most of the changes in the data, its MAE and MSE values are large, indicating that the prediction error is large and the prediction effect is not as good as other models.

The experimental results show that multi-layer LSTM (mLSTM) has significant advantages in wind speed and wind power prediction, and can more accurately capture complex patterns in the data. Compared with the traditional LSTM model, mLSTM improves the learning ability and prediction effect of the model by increasing the number of layers.

6. Conclusions

This paper proposes a method based on multi-layer long short-term memory network (mLSTM) and compares it with traditional LSTM, xLSTM and sLSTM models, mainly for the prediction of wind speed and wind power in US wind farms. Through experimental verification, mLSTM shows superior performance in wind speed prediction, especially in predicting wind speed (100 meters height, unit: m/s), achieving the best MAE, MSE and R^2 values.

Specifically, the mLSTM model can better capture the multi-level time series characteristics in wind speed data by stacking multiple LSTM layers, thereby improving the prediction accuracy and stability. In contrast, although xLSTM and sLSTM also achieved good performance in some cases, their prediction accuracy is slightly inferior to mLSTM. The results of this study provide a more efficient and accurate solution for wind power prediction. By adopting multi-layer LSTM (mLSTM), we not only improve the accuracy of wind speed prediction, but also provide more reliable data support for power dispatch and optimization of wind farms. This is of great significance for optimizing the operating efficiency of wind farms, reducing costs and ensuring the stability of power grids.

Although this study has achieved remarkable results in wind speed prediction, there are still some limitations. For example, the selection and preprocessing of data may have a certain impact on the performance of the model. In the future, richer data sets and data enhancement methods can be tried. In addition, although the mLSTM model performs well in wind speed prediction, how to combine it with other types of prediction models (such as convolutional neural networks, reinforcement learning, etc.) to improve the global prediction ability of wind power is still a direction worthy of in-depth exploration.

Future work can focus on the following aspects: Data diversity: Collect wind farm data in different regions and under different climatic conditions to enhance the generalization ability of the model. Hybrid model: Combine mLSTM with other deep learning models to explore its application in multi-task prediction. Real-time prediction system: Optimize the training and inference speed of the model for the real-time prediction needs of wind farms to achieve more efficient deployment.

Author Contributions:

Conceptualization, P.Z., J.Y., S.Z.; methodology, P.Z., J.Y., S.Z.; software, P.Z., J.Y., S.Z.; validation, P.Z., J.Y., S.Z.; formal analysis, P.Z., J.Y., S.Z.; investigation, P.Z., J.Y., S.Z.; resources, P.Z., J.Y., S.Z.; data curation, P.Z., J.Y., S.Z.; writing—original draft preparation, P.Z., J.Y., S.Z.; writing—review and editing, P.Z., J.Y., S.Z.; visualization, P.Z., J.Y., S.Z.; supervision, P.Z., J.Y., S.Z.; project administration, P.Z., J.Y., S.Z.; funding acquisition, P.Z., J.Y., S.Z. All authors have read and agreed to the published version of the manuscript.

Funding:

Not applicable.

Institutional Review Board Statement:

Not applicable.

Informed Consent Statement:

Not applicable.

Data Availability Statement:

This dataset comes from the "United States Wind Farm Data - 2012". The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author(s).

Conflict of Interest:

The authors declare no conflict of interest.

References

Duque, Á. J., Castronuovo, E. D., Sánchez, I., & Usaola, J. (2011). Optimal operation of a pumped-storage hydro plant that compensates the imbalances of a wind power producer. Electric power systems research, 81(9), 1767-1777.

Foley, A. M., Leahy, P. G., Marvuglia, A., & McKeogh, E. J. (2012). Current methods and advances in forecasting of wind power generation. Renewable energy, 37(1), 1-8.

Lungu, I., Căruțașu, G., Pirjan, A., Oprea, S. V., & Bara, A. (2016). A two-step forecasting solution and upscaling technique for small size wind farms located in hilly areas of Romania. Studies in Informatics and Control, 25(1), 77-86.

Wessel, A., Stock, S., & von Bremen, L. (2017). Wind, the Wild Boy in the Power Plant Family: How Grid Integration of Offshore Wind Farms is Achieved. Sea–Wind–Power: Research at the first German offshore wind farm Alpha Ventus, 171-186.

Medved, D., Conka, Z., Pavlik, M., Zbojovsky, J., Kolcun, M., & Ivancak, M. (2019). Prediction of electricity production in island operation under the different wind generation modes. Przeglad Elektrotechniczny, 95(7), 150-154.

Ghoushchi, S. J., Manjili, S., Mardani, A., & Saraji, M. K. (2021). An extended new approach for forecasting short-term wind power using modified fuzzy wavelet neural network: A case study in wind power plant. Energy, 223, 120052.

Pang, S., Zheng, Z., Luo, F., Xiao, X., & Xu, L. (2021). Hybrid forecasting methodology for wind power-photovoltaic-concentrating solar power generation clustered renewable energy systems. Sustainability, 13(12), 6681.

Wang, Z., Zhang, B., Mobtahej, M., Baziar, A., & Khan, B. (2021). Advanced reactive power compensation of wind power plant using PMU data. IEEE Access, 9, 67006-67014.

Gers, F. A., Schmidhuber, J., & Cummins, F. (2000). Learning to forget: Continual prediction with LSTM. Neural computation, 12(10), 2451-2471.

Yuan, F., Huang, X., Zheng, L., Wang, L., Wang, Y., Yan, X., Gu, S., & Peng, Y. (2025). The Evolution and Optimization Strategies of a PBFT Consensus Algorithm for Consortium Blockchains.Information,16(4), 268. https://doi.org/10.3390/info16040268

Huang, Z., Xu, W., & Yu, K. (2015). Bidirectional LSTM-CRF models for sequence tagging. arXiv preprint arXiv:1508.01991.

Zhou, C., Sun, C., Liu, Z., & Lau, F. (2015). A C-LSTM neural network for text classification. arXiv preprint arXiv:1511.08630.

Graves, A., & Schmidhuber, J. (2005). Framewise phoneme classification with bidirectional LSTM networks. In Proceedings. 2005 IEEE International Joint Conference on Neural Networks, 4, 2047-2052.

Liu, Z., Gao, W., Wan, Y. H., & Muljadi, E. (2012). Wind power plant prediction by using neural networks. In 2012 IEEE energy conversion congress and exposition (ECCE), 3154-3160.

Wang, Y., Huang, M., Zhu, X., & Zhao, L. (2016). Attention-based LSTM for aspect-level sentiment classification. In Proceedings of the 2016 conference on empirical methods in natural language processing, 606-615.

Wang, F., Jiang, J., & Zhang, L. (2021). Research on Ultra-Short-Term Wind Power Forecasting Based on Refactored Representation of Environmental Features. In IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology, 375-379.