Image Processing, Electronics and Computers
L. Trajkovic et al. (Eds.)
© 2024 The Authors.
This article is published online with Open Access by IOS Press and distributed under the terms of the Creative Commons Attribution Non-Commercial License 4.0 (CC BY-NC 4.0).
doi:10.3233/ATDE240493

User Load Prediction Combining Micro Meteorological Monitoring and Long Short-Term Memory Networks

Kangyi LI[1], Xiaofen YANG, Shousheng ZHAO, Zeshi XU, Daiming ZHOU and Qiang FAN
State Grid Zhejiang Electric Power Co., Ltd., Hangzhou, Zhejiang 310007, China

Abstract. Influenced by meteorology, production, holiday factors and so on, user loads fluctuate randomly. General linear regression methods are difficult to describe the nonlinear trends of load changes. Machine learning algorithms can be used to fully explore the correlation between data before and after the sequence. In order to improve the predictive performance of the model, a user load prediction method combining micro meteorological monitoring and deep learning algorithms is proposed. Firstly, the mRMR method is used to select variables from the input meteorological and load historical data, reducing the input dimension, thereby reducing the complexity of the model and improving the efficiency of the prediction algorithm. Then, the dimensionality reduced variable historical data is input into LSTM to establish an mRMR-LSTM user load prediction model. The load forecasting experiment was conducted using historical data of users from six industries in the city. The results showed that, while ensuring the efficient operation of the model, the model has a good predictive effect on user loads in most industries.

Keywords. Micro meteorological monitoring, industry user load, load forecasting, neural networks, deep learning, long short-term memory networks

1. Introduction

Unlike system level loads, user loads have strong randomness, a small base, and are strongly influenced by meteorological conditions, production plans, and other factors, making it difficult to obtain good prediction results for user loads. When establishing a prediction model for user load time series, the load values of the first few time periods and the micro meteorology at the prediction time will have varying degrees of impact on the load values at the prediction time, and general linear equations are difficult to describe the nonlinear laws of load changes. Therefore, it is necessary to fully explore the correlation between the data before and after the sequence using machine learning algorithms. Traditional machine learning algorithms such as artificial neural networks or support vector machines have shortcomings such as low running efficiency in training model parameters and network structure, and shallow perceptrons are difficult to characterize high-dimensional data features, which makes traditional algorithms unsuitable for applications with high input dimensions and large load data. Therefore, it is necessary to study prediction algorithms with high computational efficiency, small

[1] Corresponding Author: Kangyi Li, email: lkangy11@163.com.

memory consumption, and the ability to process high-dimensional features to ensure the timeliness and efficiency of load forecasting. Currently, deep learning, as a research hotspot in the fields of data mining and artificial intelligence, has received widespread attention and application in areas such as speech processing, machine translation, and image recognition. Researchers in the power industry have applied deep learning theories to fields such as power quality analysis, load forecasting, fault diagnosis, and user electricity consumption pattern classification, and have achieved good analysis results.

Gu based on the centroid similarity clustering load pattern algorithm to mine typical load patterns in historical load data, and groups them according to similarity [1]. Then, a radial basis neural network load prediction model is established for each typical load pattern. Based on the analysis of load characteristics, Cai combined the advantages of AR, SVM and other models to establish a refined load forecasting management system for typical large users in the region [2]. In the model, the effects of different date types, meteorological factors, and production plans on the load are considered. Luo used wavelet denoising and decision tree algorithm to establish a user load combination prediction model, fully mining the historical electricity consumption patterns of users, and establishing prediction models for users with different electricity consumption patterns, thereby reducing the complexity of the model [3]. Wang used the parallel computing model MapReduce and the memory parallel computing framework Spark to analyze big data on the power user side, and proposes a parallel load forecasting method based on the random forest algorithm, which shortens the load forecasting time and improves the processing ability of the random forest algorithm for large data [4]. Zheng combined big data technology to construct a load impact model for distribution transformers, predict short-term grid supply loads, analyze residential vacancy rates, and analyze industry capacity utilization, which improves the accuracy of load forecasting [5]. In addition, based on the analysis of demand side response, Giacometto established a load forecasting model and compared the prediction accuracy of traditional linear regression forecasting models with artificial intelligence forecasting models [6]. Li used extreme learning machines to establish a user load prediction model and achieved good prediction results [7]. Xiao established a denoising stack auto-encoder network and a deep confidence network based on a deep learning architecture to achieve the classification and recognition of power quality disturbance signals [8]. Compared to traditional classification and recognition algorithms, deep learning algorithms avoid the impact of improper feature extraction on classification accuracy, and also have good noise resistance. Some related works applied models such as deep auto-encoder networks, convolutional neural networks, and classification deep learning neural networks to the study of time series prediction, verifying the advantages of deep learning feature processing and classification accuracy [9-11]. In the field of user electricity consumption pattern classification, Lin combined a large number of user load curves into four typical electricity consumption patterns based on clustering algorithms, which are bimodal, tri modal, stationary, and peak avoiding [12]. Then, a sparse self coding deep learning model was used to classify the clustering results, achieving the goal of identifying user electricity consumption patterns.

Long short-term memory networks (LSTM) is a deep learning algorithm invented by scientists Hochreiter and Schmidhuber in 2006, which successfully solves the problems of gradient vanishing and exploding in recurrent neural networks (RNNs) when dealing with long time series problems [13]. LSTM achieves information forgetting and retention through control units, thus being able to effectively process and predict time series with long intervals or delays [14-18]. These article used LSTM to process user load time series and achieve short-term prediction of user load. On the other hand, the user load prediction algorithm in this article requires meteorological and load measurement data and statistical data as inputs, involving numerous input variables that are not independent of each other. To ensure the efficiency and robustness of prediction modeling, we need to first select the input variables. The mRMR method [19] uses mutual information theory to select input variables from historical load data and meteorological data, so as to have a significant correlation with the load value at the time of prediction, while minimizing the redundancy between input variables. Becoming the preferred method for selecting input variables in this article.

In response to the problems of low operational efficiency and low prediction accuracy in traditional machine learning methods for user load forecasting, this paper proposed a user load forecasting method that combines micro meteorological monitoring and deep learning algorithms. Based on data preprocessing of meteorological and load data, it deeply explores the correlation between multivariate data such as meteorological, production planning, and industry characteristics and load fluctuations, which can effectively improve the accuracy of photovoltaic power forecasting.

2. Data Preprocessing

In the process of collecting meteorological data and user load, abnormal measurement sensors, data transmission interruptions, and line maintenance can cause meteorological data anomalies. Abnormal data cannot reflect the actual changes in weather and load, resulting in a lower correlation between meteorological data and user load, and therefore cannot be selected as input variables. Therefore, it is necessary to identify and correct the abnormal data obtained from various meteorological stations.

Abnormal data can be divided into two types: distorted data and missing data. Distorted data includes shock load data and spiked load data, where shock load is mainly caused by sudden events, major events in certain social life, or random factors in the electricity market model; Burr load data shows sudden increases or decreases between several load points in adjacent time periods of the curve. Missing data is often due to a malfunction in a measurement unit or other related components during the data collection process. Missing data is usually recorded as a null value or zero, but the zero value is not the actual load value.

The commonly used methods for identifying and correcting abnormal data include the following:

2.1. Empirical Correction Method

This method identifies and corrects load data by relevant judging based on long-term accumulated experience which has shortcomings such as strong subjectivity, cumbersome operation, and low accuracy.

2.2. Curve Permutation Method

This method directly removes obvious abnormal load data or replace it with the previous normal load curve. However, it has significant limitations. From a statistical perspective, the direct elimination method sacrifices sample size to obtain complete information, which wastes load data resources and discards a large amount of information hidden in these curves.

2.3. Threshold Discrimination Method

This method observes the general variation patterns of loads in different industry attributes, sets the range of load fluctuations according to time periods, determines the maximum and minimum load widths, and then screens the various load data.

2.4. Horizontal and Vertical Comparison Method

These two types of methods are the most widely used and their effectiveness has also been recognized by researchers. The so-called horizontal comparison refers to comparing and analyzing the loads at each point and their adjacent loads from the perspective of a single daily load curve. If the difference is greater than a certain threshold, it is considered that the data is abnormal, the so-called vertical comparison refers to analyzing and comparing the loads at various points from the perspective of multiple historical load curves of the user. When the deviation exceeds the set threshold, it is judged as abnormal data and replaced by the mean.

2.5. Interpolation Correction Method

This method uses interpolation algorithms to identify and correct load data.

2.6. Probability and Statistics Method

The main application of this method is the confidence interval method in statistics, which assumes that the multi day load data in the same period is approximately normally distributed. Multiple historical load data are used as samples for probability statistical analysis to estimate the expected values and variances of the two normal distribution models in a specific time period. Based on the set confidence level, the confidence interval of the load level in that time period is estimated.

3. mRMR-LSTM User Load Forecasting

3.1. Input Variable Selection Based on mRMR Method

The Input Variable Selection method proposed in this article used mutual information theory to select input variables from historical load data and meteorological data, so as to find out significant correlation between historical data and the predicted load value at the time of prediction, while minimizing the redundancy between input features. In the theory of information entropy, the amount of information shared among multiple variables is represented by mutual information values. The formula for mutual information value is as follows [20,21]:

$$I(X,Y)=\sum_{i=1}^{n}\sum_{j=1}^{m}p(x_i,y_j)\log_2\left(\frac{p(x_i,y_j)}{p(x_i)p(y_j)}\right) \tag{1}$$

In the formula, n represents the total number of data for input variable X, and m represents the total number of data for input variable Y.

The larger the mutual information value, the stronger the correlation between variables; otherwise, the weaker the correlation. Based on mutual information values and combined with different evaluation indicators, multiple variable selection methods have been formed [22]. Among them, the maximum correlation minimum redundancy method (mRMR) uses mutual information values to measure the correlation and redundancy between input variables. Using the maximum correlation index to increase the correlation between the selected input variables and the prediction target, and using the minimum redundancy index to minimize the correlation between the selected input variables, the reduction of input variables is achieved [19].

The measurement indicators for maximum correlation and minimum redundancy are defined as follows:

$$\max D(S,c), D=\frac{1}{\mathrm{N}}\sum_{x_i\in S}I(x_i;c) \tag{2}$$

$$\min D(S), \mathrm{R}=\frac{1}{\mathrm{N}^2}\sum_{x_i,x_j\in S}I(x_i;x_j) \tag{3}$$

In the formula: S is set of input variables, N is the number of input variables, and c is the prediction target. D is the correlation between input variables x_i and prediction target c in S. R is redundancy between variables in S. mRMR defines optimization function and perform iterations to maximize the correlation between input features and targets, and minimize the redundancy between input features.

mRMR avoids redundant information while selecting input variables with high correlation. Therefore, the mRMR algorithm is first used to sort all input variables. Secondly, based on the working mode of the wrapper, the input variable set sorted by mRMR is added to the selected feature set one by one through a forward search strategy. Then, an LSTM prediction model is established for the selected feature set, and the optimal input variable is determined by comparing the prediction error.

3.2. Long Short-Term Memory Networks

Due to the problem of vanishing and exploding gradients during the training process, RNN is unable to capture the impact of long-range output on the current output, which limits its widespread application and development. Long short-term memory network is an improvement of RNN, which effectively controls long-distance information by adding new unit states in the hidden layer, and solves the problems of gradient vanishing and exploding.

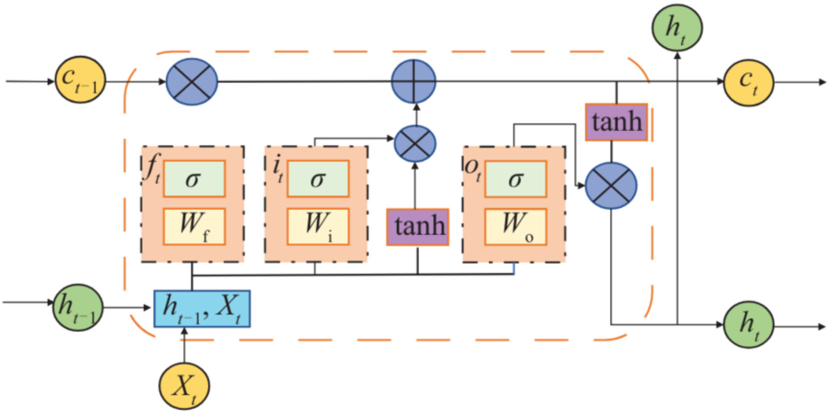

Figure 1. Network Structure LSTM.

The network structure of LSTM is shown in Figure 1. The LSTM network mainly includes input gates, output gates, and forget gates. The forget gate determines whether useful information is retained, the input gate determines what information is stored in the memory unit, and the output gate determines the next hidden state. LSTM uses error backpropagation algorithm to update and adjust model parameters. The calculation process of error backpropagation is shown in formula (4).

$$\begin{cases} i_t = \sigma(W_{ix}x_t + W_{im}m_{t-1} + W_{ic}c_{t-1} + b_i) \\ f_t = \sigma(W_{fx}x_t + W_{fm}m_{t-1} + W_{fc}c_{t-1} + b_f) \\ c_t = f_t c_{t\text{-}1} + i_t g(W_{ix}x_t + W_{im}m_{t-1} + W_{ic}c_{t-1} + b_c) \\ o_t = \sigma(W_{ox}x_t + W_{om}m_{t-1} + W_{oc}c_{t-1} + b_0) \\ m_t = o_t b_y \end{cases} \tag{4}$$

In the formula: i_t is the input gate; W - Weight matrix; x_t is the input data for the current time step t; c_t and m_t serve as activation carriers for cellular state and memory blocks; b is the bias term; f_t is the forgetting gate.

3.3. mRMR-LSTM Model Algorithm Flow

Due to the high dimensionality, rich information, and complexity of the original historical data used as input variables for the model, load fluctuation related information such as meteorological factors and holiday factors are hidden in the complex correlations of the data. By using the mRMR method to select input variables, the complexity of the model can be reduced, that is, the dimensionality can be reduced. Then, the reduced historical data of the input variables can be input into the LSTM network to continuously explore temporal data dependencies and achieve accurate user load prediction. The algorithm flow of the mRMR-LSTM model is shown in Figure 2.

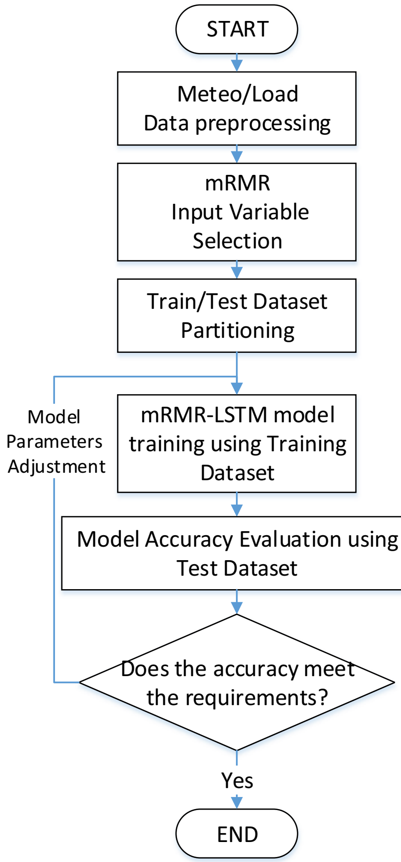

Figure 2. Algorithm flowchart of mRMR-LSTM model.

The specific process of the mRMR-LSTM model algorithm is as follows:

- Fill in missing values in meteorological and load historical data, adjust data sampling intervals, and complete data preprocessing. The processed historical data is used as a candidate input variable.
- Using the mRMR method to select candidate input variables and reduce dimensionality.
- Divide the selected input variable dataset into a training set and a testing set.
- Initialize the model parameters, place the training set into the mRMR-LSTM model for training, and then use the test set for testing. Iteratively run until the accuracy meets the requirements.

4. Experimental Verification and Result Analysis

To verify the effectiveness and accuracy of the prediction model in this article, user load data from 6 industries in a prefecture level city in China from August 1, 2017 to June 30, 2018, as well as meteorological data from local meteorological stations during the same period, were selected as model training data. Load data and meteorological data from July 1, 2018 to July 31, 2018 were selected as test data for prediction, and the model was evaluated. Meteorological data mainly includes temperature, humidity, rainfall, total horizontal radiation, etc. The data sampling interval is 1 hour, and the running environment is Python 3.7 and Pytorch 1.5.

4.1. Evaluating Indicator

The evaluation of predictive models usually has the following three error indicators, and the error statistical formula is:

Mean Absolute Percentage Error, MAPE:

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\frac{|y_i - \hat{y}_i|}{y_{avg}} \times 100\% \tag{5}$$

Root Mean Square Error, RMSE:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(y_i - \hat{y}_i\right)^2} \tag{6}$$

Mean Absolute Error, MAE:

$$MAE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i|} \tag{7}$$

In the formula: *n* represents the total number of test points; y_i and $\hat{y}_i$ are the true and predicted values of user load at the *i*-th test point, respectively, and y_{avg} are the mean of *n* true values.

4.2. Analysis of Industry User Load Forecast Results

This article collected 1-year historical monitoring data from 131 meteorological stations near the city. Through identification and correction, meteorological stations with severe missing data were eliminated, and micro meteorological data from 36 meteorological stations were obtained. The information of each meteorological station is shown in Table 1. Each user selects meteorological data from the nearest weather station based on longitude and latitude as candidate input variables.

Table 1. Meteorological Station Information

No	St. Code	No	St. Code	No	St. Code	No	St. Code
1	M4755	10	M3581	19	M4758	28	M3551
2	M3531	11	M3591	20	M6702	29	M6738
3	M3557	12	M3592	21	M6711	30	M3559
4	M4702	13	M4701	22	M9012	31	M6735
5	M4707	14	M4703	23	M9013	32	M9007
6	M4752	15	M4705	24	M9014	33	M3544
7	M4754	16	M4708	25	M9016	34	M3553
8	M5562	17	M4734	26	M6701	35	M6779
9	M6707	18	M4735	27	M6770	36	M9009

Display the load forecasting performance of the top ranked user with existing data in each industry. The user information table is shown in Table 2.

Table 2. User Information

ID	User Name	Consumption Category	Capacity (KVA)	St. No	Industry
110002873218	university	Teaching	235880	M6707	education
110004581626	Real Estate Development Co.	Commercial	54600	M3557	real estate
110003385830	Property Management Co.	Commercial	49140	M6779	retail
300012245842	Construction and Development Co.	Non residential lighting	57000	M6735	Research and experimental development
110000000508	Hospital	Non residential lighting	70720	M3544	Health care
300001098956	New Materials Co.	Large industrial electricity	202000	M3559	Chemical product manufacturing

The results of user load forecasting for various industries are shown in Figure 3-Figure 8, and the accuracy evaluation indicators are shown in Table 3.

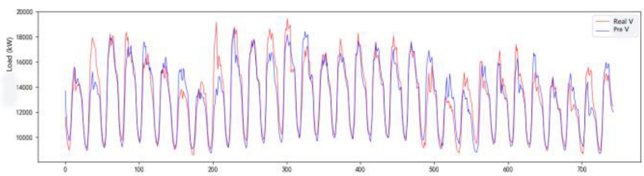

Figure 3. Load Forecast Results for University

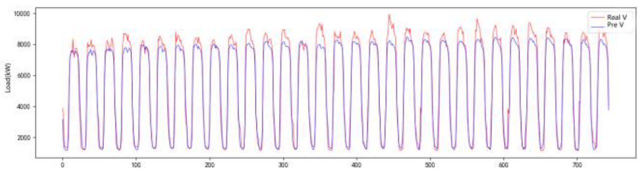

Figure 4. Load forecast results of Real Estate Development Co.

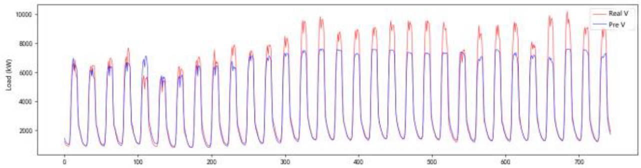

Figure 5. Load forecast results of Property Management Co.

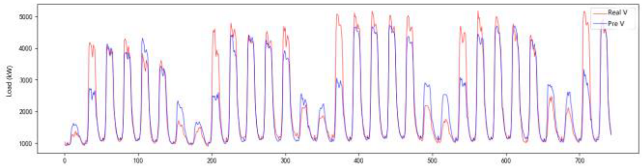

Figure 6. Load forecast results of Construction and Development Co.

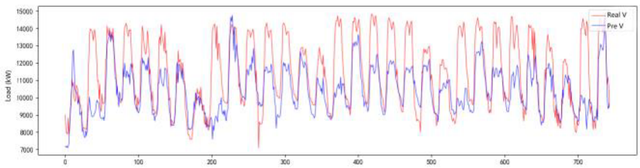

Figure 7. Load forecast results of Hospital

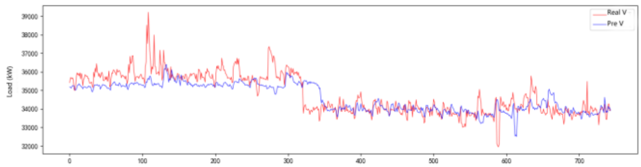

Figure 8. Load forecast results of New Materials Co.

Table 3. Accuracy Evaluation Indicators

User ID	RMSE	MAE	MAPE(%)
110002873218	1059.41	766.89	5.78
110004581626	574.00	423.08	7.55
110003385830	741.68	438.47	11.01
300012245842	501.56	278.63	12.43
110000000508	1706.07	1236.65	10.83
300001098956	669.59	479.71	1.38

From the analysis of the results, it can be seen that the model has a good predictive effect on user load in most industries. However, for some industries, such as chemical raw materials and chemical product manufacturing, the predictive effect on users is not ideal. This is mainly for users with high volatility, and more samples are needed to learn their load variation patterns.

5. Conclusion

The prediction method combining micro meteorological monitoring and mRMR-LSTM proposed in this article can effectively improve the level of industry user load prediction. Experimental analysis was conducted on historical user data from six industries in a prefecture level city in China, and the following main conclusions were drawn:

- The use of mRMR method for input variable selection can effectively reduce model complexity, i.e. reduce dimensions and improve model computational efficiency.
- The mRMR-LSTM model proposed in this article takes into account the impact of meteorological fluctuations and industry characteristics on load fluctuations, achieving high-precision industry load forecasting. Among the 6 industries, the average absolute percentage error of users in 5 industries is less than 13%. The predictive performance of users in the chemical raw material and chemical product manufacturing industry is not ideal, mainly for users with high volatility, and more samples are needed to learn their load variation patterns.

Acknowledgment

This paper is supported by Science and Technology Project of State Grid Zhejiang Electric Power Co. Ltd. of China (B311SX23000C).

References

[1] Gu Y, Zhang S, and Feng J. Multi-model Fuzzy Synthesis Forecasting of Electric Power Loads for Larger Consumers. Trans China Electrotechnical Society. 2015 Dec; 30(23): 110-115, doi:10.19595/j.cnki.1000-6753.tces.2015.23.014.

[2] Cai J, Luo D, Zhou X, et al. Load Forecasting Management System About Large Consumer. Power Demand Side Management. 2012 Feb; 14(04): 7-10+24.

[3] Luo Y, Cheng J, Wang Y, et al. Wavelet De-noising and Decision Tree Based Load Forecasting of Large Consumers. Southern Power System Technology. 2016 Oct; 10(10): 37-42. doi:10.13648/j.cnki.issn1674-0629.2016.10.006.

[4] Wang D, Sun Z, Big Data Analysis and Parallel Load Forecasting of Electric Power User Side. Proceedings of the CSEE. 2015 Feb; 35(03): 527-537. doi:10.13334/j.0258-8013.pcsee.2015.03.004.

[5] Zheng H, Jin N, Ji C, et al. Analysis Technology and Typical Scenario Application of Electricity Big Data of Power Consumers. Power System Technology. 2015 Nov; 39(11): 3147-3152. doi:10.13335/j.1000-3673.pst.2015.11.020.

[6] Giacometto F, Cardenas J J, Kampouropoulos K, et al. Load Forecasting in the User Side Using Wavelet-ANFIS. //IECON 2012 - 38th Annual conference of the IEEE Industrial Electronics Society: IECON 2012, Montreal, Quebec, Canada, 25-28 October 2012, pages 881-1822, [v.2].:IEEE, 2012: 1049-1054.

[7] Li Y, Guo P, Li X, Short-term load forecasting based on the analysis of user electricity behavior. Algorithms, 2016 Apr; 9(4): 80. doi:10.3390/a9040080.

[8] Xiao L, Classification of Power Quality Disturbances Based on Deep Learning, Nanchang University, 2017.

[9] Alizamir M, Shiri J, Fard AF, et al. Improving the accuracy of daily solar radiation prediction by climatic data using an efficient hybrid deep learning model: Long short-term memory (LSTM) network coupled with wavelet transform. Engineering Applications of Artificial Intelligence: The International Journal of Intelligent Real-Time Automation, 2023, 123(Pt.A): 106199-1-106199-24. doi:10.1016/j.engappai.2023.106199.

[10] Dujardin J, Lehning M. Wind-Topo: Downscaling near-surface wind fields to high-resolution topography in highly complex terrain with deep learning. Quarterly Journal of the Royal Meteorological Society, 2022, 148 (744 Pt.A): 1368-1388, doi:10.1002/qj.4265.

[11] Simon V, Kirien W, Sjoerd D, et al. Statistical Postprocessing of Wind Speed Forecasts Using Convolutional Neural Networks. Monthly weather review. 2021 Apr; 149(4): 1141-1152, doi:10.1175/MWR-D-20-0219.1.
[12] Lin J, Application of Recognition Model for Power Load in Cluster Ensemble and Deep Learning, South China University of Technology, 2014.
[13] Kong W, Dong Z Y, Jia Y, et al. Short-term residential load forecasting based on LSTM recurrent neural network. IEEE Transactions on Smart Grid, 2017 Feb; 99(2): 1-11.
[14] Tyass I, Khalili T. Wind Speed Prediction Based on Statistical and Deep Learning Models. International Journal of Renewable Energy Development. 2023 Mar; 288-291.
[15] Limouni T, Yaagoubi R. Univariate and Multivariate LSTM Models for One Step and Multistep PV Power Forecasting, Aug 2022; 815-820.
[16] Trigo-González M, Cortés-Carmona M, et al. Photovoltaic power electricity generation nowcasting combining sky camera images and learning supervised algorithms in the Southern Spain. Renewable Energy. April 2023; 206:251-262.
[17] Ajith M, Martínez RM. Deep learning based solar radiation micro forecast by fusion of infrared cloud images and radiation data. Applied Energy, July 2021; 294:234-243.
[18] Agga A, Abbou A, Labbadi M., et al. Short-term self consumption PV plant power production forecasts based on hybrid CNN-LSTM, convlstm models. Renewable Energy. March 2021; 177: 101-112.
[19] Li Y, Gu X. Feature Selection for Transient Stability Assessment Based on Improved Maximal Relevance and Minimal Redundancy Criterion. Proceedings of the CSEE. Dec 2013; 33(34): 179-186+27. doi:10.13334/j.0258-8013.pcsee.2013.34.024.
[20] Tang Y, Gui W, Peng T, et al. Prediction method for dissolved gas concentration in transformer oil based on variable selection of mutual information. Chinese Journal of Scientific Instrument. Jul 2013; 34 (07): 53-59. doi:10.19650/j.cnki.cjsi.2013.07.008.
[21] Zhao T, Wang L, Zhang Y, et al. Relation Factor Identification of Electricity Consumption Behavior of Users and Electricity Demand Forecasting Based on Mutual Information and Random Forests. Proceedings of the CSEE. Feb 2016; 36 (03): 604-614. doi:10.13334/j.0258-8013.pcsee.2016.03.003.
[22] Han M, Liu X. Stepwise Input Variable Selection Based on Mutual Information for Multivariate Forecasting. Acta Automatica Sinica. Jun 2012; 38 (06): 999-1006.