Prediction of photovoltaic power generation based on Bayesian neural network with grey correlation

Jiaxiong ZHU¹, Zhigang XIAO² and CHANG FENG³*

¹School of Electronics and Materials Engineering of Leshan Normal university, Leshan 614004, China
² School of Mathematics and Physics of Leshan Normal university, Leshan 614004, China
³ School of Electronics and Materials Engineering of Leshan Normal university, Leshan 614004, China
*Corresponding author’s e-mail: wudxy@lsnu.edu.cn

Abstract. Photovoltaic power generation depends on weather conditions, which has the characteristics of randomness, volatility and intermittence. In order to ensure the security and stability of the grid connected process and make full use of the advantages of photovoltaic power generation, this paper proposes Bayesian optimization algorithm based on grey correlation to predict the power of photovoltaic power generation. The gray correlation selects the similar correlation data from the original data set, and normalizes the data sample set to eliminate the influence of data size caused by different data dimensions. Bayesian optimization algorithm can effectively prevent over fitting by adding regularization term and super parameter \( \alpha, \beta \). Experiments show that the photovoltaic power prediction method based on Bayesian optimization algorithm under grey correlation has higher accuracy and robustness, and can provide more comprehensive information for power grid dispatching and control.

1. Introduction
The output of photovoltaic power generation is unstable and restricted by various factors. The first factor is the electrical and hardware properties of the photovoltaic power station, including the age of the panel, the wear degree of the electrical components, the size of the panel, the difference in the installation process and the performance of the electrical instruments. The second type of factors are environmental factors, mainly including irradiance, atmospheric pressure, relative humidity, temperature, cloud cover and weather type. For a specific photovoltaic power station, the historical power data of the power station is used to measure when establishing the artificial intelligence algorithm model, and the historical power data often contains the internal attributes of the power station, so the external environment factors are the important reasons for the deviation of the artificial intelligence algorithm model.

In reference [1], Barbieri et al. Proposed that using satellite to predict meteorological data and panel temperature can quickly predict photovoltaic power value in a short time.
In reference [2], Agoua et al. Proposed a scheme to solve the problem of data stationarity, and developed an ultra short-term prediction method based on statistical analysis method (to predict the generation power in the next 0 to 6 hours). This method can update the data in real time, and the amount of calculation is low.

In reference [3], Ning Kanghong and others first used extreme learning machine algorithm to train the prediction model, and combined with similar days to eliminate the influence of weather types, the prediction results were greatly improved.

In reference [4], Dai Qian et al. Used Euclidean distance to characterize the similarity between the actual power value and environmental factors, determined the temperature and humidity as the input variables of the model, and then combined with BP neural network and SOM algorithm to predict.

In this paper, gray correlation algorithm is used to select correlation similar data, and then the influence of weather type on prediction results is eliminated. Combined with standard deviation normalization, the influence of dimension on model training is eliminated. Through the Bayesian optimization algorithm to establish the prediction model, and through the experimental comparative analysis, the prediction method has stronger generalization ability and better robustness, and can effectively prevent over fitting.

2. Grey relational analysis

There are many factors that affect the power of photovoltaic power generation. In the actual project, we cannot consider every factor, we can only select the main factors for modeling. Through the correlation analysis, irradiance, ambient temperature and relative humidity have a significant impact on the photovoltaic power generation. Because the actual power is affected by weather type and season, there are no fields of weather type in historical data and weather forecast data, and the data amount is huge. If manually marking weather type, it will waste a lot of time. In order to solve the problem of the influence of weather types, this paper analyzes the historical data through grey correlation, and obtains the data which is consistent with the weather types of the forecast day. These data are called similar correlation data.

Grey correlation analysis is a method to study the correlation between two features. Its essence is to compare and analyze the correlation degree between the geometric shapes formed by two feature sequences.[5-6]

(1) Constructing feature matrix

The values of the eigenvectors of the Characteristic matrix constitute the vector X series, and the eigenvector of the day to be predicted is \(X_0 = [x_0(1), x_0(2), \ldots, x_0(p)]\). The feature vectors of each day in historical data are expressed in \(X_1, X_2, \ldots, X_d\). The \(p \times (d + 1)\) matrix is composed of the eigenvectors of the days to be predicted and the historical data, and the matrix is called the Characteristic matrix.

(2) Matrix solution

The matrix composed of the difference between the day vector to be predicted and the eigenvectors of each historical data is the difference matrix, and the calculation formula is as follows:

\[
\Delta x_i^0(j) = |x_i(j) - x_i(j)|
\]  

Where \(\Delta x_i^0(j)\) is the difference between the j-th value of irradiance and the j-th value of the day to be predicted in the i-th sample of historical data.

(3) The solution of grey relation degree

The correlation coefficient \(r_i^0\) of the daily irradiance to be predicted and the irradiance of the I sample is as follows:
\[ r_{i}^0 = \frac{1}{m} \sum_{j=1}^{m} \min\left\{ \frac{|x_{0}(j) - x_{i}(j)| + \epsilon \max\{|x_{0}(j) - x_{i}(j)|\}}{|x_{0}(j) - x_{i}(j)| + \epsilon \max\{|x_{0}(j) - x_{i}(j)|\}} \right\} \]

Where \( \epsilon \) is the resolution coefficient, usually 0.5. The closer \( r_{i}^0 \) is to 1, the better the correlation.

3. Bayesian neural network

The principle of Bayesian algorithm is a common machine learning algorithm based on Bayesian formula and conditional independence assumption. For the existing sample data, the joint probability distribution is first learned according to the sample. Then, for the given input value, the posterior probability is calculated according to the model and Bayesian theorem.[7-8]

This paper mainly studies Bayesian regularized neural network, which can automatically adjust the regularization coefficient in the training process to meet higher robustness, thus improving the generalization ability of prediction model. Secondly, Bayesian neural network will adjust the weight automatically with the training process, and reduce the weight of the input feature which has little influence, so as to prevent the over complexity of the neural network, simplify the network and prevent the phenomenon of over fitting.

Bayesian neural network avoids over fitting problem of BP neural network by adding regular term, which is as follows:

\[ R(w) = \frac{1}{2} \sum_{i=1}^{m} w_{i}^2 \]  

Where \( w_{i} \) is the ith weight value, \( m \) is the number of weights, and the final optimization objective function is:

\[ y(w) = \beta E_T + \alpha R(w) \]

\[ E_T = \frac{1}{2} \sum_{i=1}^{N} (y_i - y_i)^2 \]

Where,

The training process of Bayesian neural network is as follows:

(1)Initialize the super parameters \( \alpha, \beta \) and the weight value \( w_{i} \) of each connection point, usually making \( \alpha = 0, \beta = 1 \).

(2)The objective function (4) is optimized by BP neural network to obtain the local optimal value.

(3)Calculate the value of grey correlation degree \( r_{i}^0 \).

(4) Update the value of super parameter \( \alpha, \beta \).

(5) Repeat steps (2) to (4) until the requirements are met or the network converges.
In this paper, Bayesian neural network is used to introduce super parameters $\alpha$, $\beta$, and grey correlation degree $r^0$, in the process of building the model, the super parameters can be adjusted in real time according to the training situation, so as to prevent over fitting of the model.

4 Simulation Verification

This paper selects the historical data of Li Tang poverty alleviation photovoltaic power plant in Gan Zi as the basis, and uses the data of the week from May 8, 2020 to May 14, 2020 to forecast. Table 1 shows the prediction accuracy statistics of BP neural network and Bayesian neural network.

| Forecast date | BP neural network (ACC, MAPE) | Bayesian neural network (ACC, MAPE) |
|---------------|-------------------------------|-----------------------------------|
| 2020/5/8      | 0.8810, 0.4238               | 0.9276, 0.2074                   |
| 2020/5/9      | 0.9087, 0.2129               | 0.9494, 0.2088                   |
| 2020/5/10     | 0.9247, 0.1416               | 0.9177, 0.2586                   |
| 2020/5/11     | 0.9216, 0.2065               | 0.9512, 0.1052                   |
| 2020/5/12     | 0.9599, 0.1063               | 0.9812, 0.0469                   |
| 2020/5/13     | 0.9408, 0.2447               | 0.9568, 0.1985                   |
| 2020/5/14     | 0.9629, 0.1725               | 0.943, 0.1609                    |
| Average       | 92.33%, 0.2155               | 94.67%, 0.1695                   |

By analyzing table 1 and Figure 2, it is found that the average accuracy of Bayesian neural network
is higher than that of BP neural network, and the error is correspondingly lower. From the prediction of each day, the accuracy of the prediction results based on Bayesian neural network is higher than that of BP neural network except May 10 and May 14.

Figure 3. Predict power and actual power trend graphs of neural network on rainy day

Figure 4. Predict power and actual power trend graphs of neural network on sunny day

Figure 5. Predict power and actual power trend graphs of neural network on cloudy day
Figure 3, figure 4 and figure 5 are the trend diagrams between the real power generation and the predicted power generation on cloudy, rainy and sunny days respectively. The analysis and comparison show that, on the whole, the performance of cloud and rain forecast is very general, but the performance of Bayesian neural network is still stable in this type of weather, and the fluctuation of accuracy is not obvious. In sunny days, the result curve is smoother and the generalization ability is stronger.

5 Conclusion
In this paper, the improved grey correlation analysis method is used to obtain the correlation similar data, which can improve the quality of training data, reduce useless data, and eliminate the influence of different weather types on the results. By introducing super parameters $\alpha$ and $\beta$, the size of super parameters can control the loss function and regularization variables, and the super parameters can be adjusted in real time according to the training situation in the process of building the model, so as to prevent over fitting of the model. The accuracy of the proposed Bayesian neural network in photovoltaic power prediction has reached the expected level.

Acknowledgements
School level project: technical transformation and capacity increase of 500kW photovoltaic power station in LiTang, GanZi No: 801/204190429.

References
[1] Barbieri FJR, akaruna S, Ghosh A. Very short-term photovoltaic power forecasting with cloud modeling [J]. Renewable & Sustainable Energy Reviews, 2016, 75(8): 445-448.
[2] Agoua. Short-term forecast of generation of electric energy in photovoltaic system [J]. Renewable and Sustainable Energy Reviews, 2018, 81(5): 306-312.
[3] Liu Shirong, Li SONGFENG, Ning Kanghong. Short term prediction of photovoltaic power generation based on extreme learning machine [J]. Control engineering, 2013, 20 (2): 372-376.
[4] Dai Qian, Duan shanxu, Cai Tao. Research on short-term irradiance generation prediction model of photovoltaic system based on clustering identification of weather types [J]. Chinese Journal of electrical engineering, 2011 (34): 881-884.
[5] Huang M, Wang X, Loh P C, et al. Design of LLCL-filter for grid-connected converter to improve stability and robustness [C]. Applied Power Electronics Conference & Exposition, IEEE, 2015.
[6] Zhiheng Z, Weimin W, Zhikang S, et al. Principle and Robust Impedance-Based Design of Grid-tied Inverter with LLCL-Filter under Wide Variation of Grid-Reactance [J]. IEEE Transactions on Power Electronics, 2018:1-1.
[7] Bouzerdoum M, Mellit A, Massi Pavan A. A hybrid model (SARIMA-SVM) for short-term Power forecasting of a small-scale grid-connected photovoltaic plant [J]. Sol Energy, 2013, 98(9): 226-235.
[8] Zhang Jinhua, Cheng Peng, Liu Yongqian, et al. Ultra short term prediction of wind speed and wind power based on phase space reconstruction [J]. Renmin Huanghe, 2012, 34 (7): 117-10.