

Wind Power Forecasting in Short-Term Using Fuzzy K-Means Clustering and Bagging Neural Network

R. Praveena, K. Dhanalakshmi, M.S. Thanabal

Wind power forecasting is the basic and essential task in the wind power plant operation. Wind power is one of the fast-growing feasible power resources and it will be viewed as the additional substitute for conservative power produced from the non-sustainable power source. The use of the conservative source of electricity will be reduced with the help of wind power forecasting (WPF). The fluctuation and instability of the wind will be difficult for collecting the original data and it will be the major impact for forecasting the accuracy. Last few years many researchers use different data mining technique that has been applied in various prediction system that produced good forecasting accuracy. To improve the accuracy of the forecasting and reduce the computational complexity for a hybrid method is proposed that consist of fuzzy k-means clustering and Bagging Neural Network (BNN). In historical days there is a lot of similarities, fuzzy k-means clustering is used for clustering the similar days it consisting of the detailed information about historical data and weather conditions. In order to avoid the overfitting problem, a bagging algorithm to be incorporated with backpropagation neural network. For proving the efficiency, the proposed hybrid method will be evaluated in real wind farm which will provide better forecasting accuracy and also expected to reduce the computational complexity when compared with other existing wind power forecasting approaches.

Keywords- Bagging Neural Network (BNN); Ensemble Learning; Fuzzy K-means clustering; Wind Power Forecasting (WPF).

I. INTRODUCTION

Wind power is one of the developing sustainable power resources. It provided electrical supply with high efficiency around worldwide. However, the random behaviour of the wind will affect reliability and the scalability of the grid and it also difficult for collecting the training sample. The reliable and the absolute wind power forecasting will provide a possible solution to this issue. It also reduces the cost of the reeling resource. For prediction of wind power generation, a huge number of WPF approaches are discussed. The most commonly used wind power forecasting method is classified into three types are the physical method, statistical method, combined method [1]. The physical method can provide

physical information such as geographical location and meteorological condition of the particular wind farm and it does not need any historical information for installing the new wind farm. The numerical wind forecasting will help in providing a better prediction result [2].

The statistical method helps in determining the similarity between the variable and the wind power such as wind speed, wind pressure, wind direction etc [3]. The combined method has taken the advantage of physical and the statistical method to provide better forecasting accuracy. The physical method has used the specified wind power forecasting model of the wind farm without the requirement of the huge amount of data.

For obtaining the correct design detail information about the wind farm and the wind turbine to be needed [4]. But it will be difficult in collecting the physical characteristic in short term. When compared with the physical method, the statistical method can need a large amount of historical information for building the best forecasting model. The statistical method uses the time series model. This model is used for prediction of the wind power. The time series models are of many types some of them are autoregressive, AR integrated model [5].

This model will help in the modification of the weight dynamically based on the autoregressive model. The short-term wind forecasting uses the time series model which will provide a better prediction result. This method uses the continuous sequence and it does not require similar data.

Many researchers are concentrated on the data mining approaches [6]. For short-term wind forecasting the neural network is used. It also consists of some other approaches. The Neural Network (NN) can handle the composite relationship between the weather condition and historical information [7]-[8]. The historical data are used by the neural network which will be the major cause for forecasting the wind power. Some of the variables give as the inputs are wind pressure and wind direction [9]. One of the important method used for training the input data is the back propagation neural network (BPNN). The BPNN can easily formulate and it also helps in improving the potential. The backpropagation can initially take the weight and bias value as the input. The overfitting problem may arise because of the too many attributes are used as the training data [10]. One of the major impacts in reducing the forecasting accuracy is that it does

Revised Version Manuscript Received on 10, September 2019.

R. Praveena, PG Scholar, Department of Computer Science & Engineering, PSNA College of Engineering & Technology, Dindigul, Tamilnadu, India. (Email id:ramalingam.praveena@gmail.com)

K. Dhanalakshmi, Professor, Department of Computer Science & Engineering, PSNA College of Engineering & Technology, Dindigul, Tamilnadu, India.

M.S. Thanabal, Professor, Department of Computer Science & Engineering, PSNA College of Engineering & Technology, Dindigul, Tamilnadu, India.

not consider the selection of training samples.

For improving the accuracy of the forecasting, the clustering method is used which helps in the selection of training samples [11]. The k-means clustering algorithm is used for clustering the similar data and it also classified in too many categories [12]. The bagging Neural network will help in reducing the overfitting problem [13]. The Fuzzy K-means clustering will helps in clustering the similar days and it will provide the accurate predicting result [14].

In this paper, fuzzy k-means clustering and bagging neural network are used to forecast the wind power in short period of time.

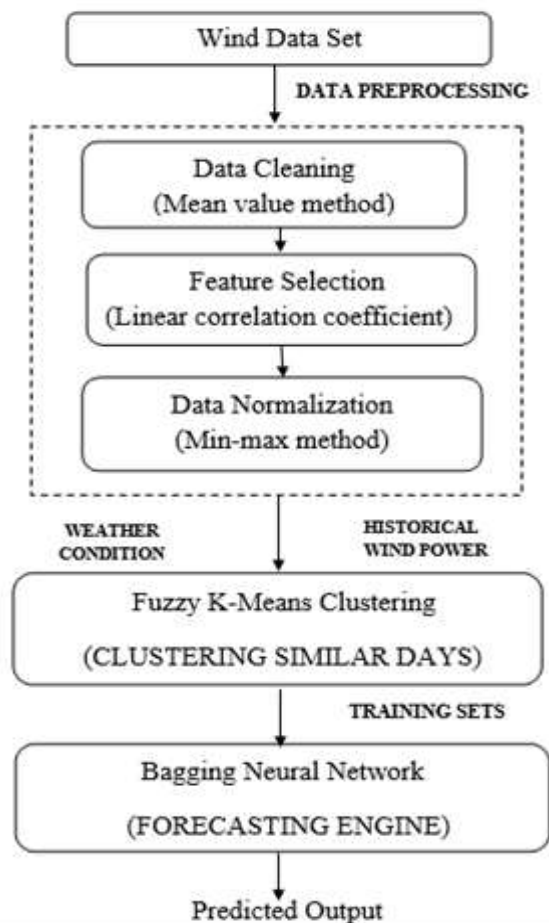


Fig. 1 Architecture of the Wind Power Forecasting Approach

II. WIND POWER FORECASTING APPROACHES

The data mining methods are widely used in many applications and it helps in solving many problems like classification, clustering and prediction problems. The proposed system of Fuzzy K-means clustering and Bagging Neural Network (BNN). In fig 1 shown the proposed wind power forecasting approach. First, the data preprocessing helps in removing the unreasonable data and the feature selection helps in selecting the relevant feature and then normalize the selected feature. Second, the Fuzzy k-means clustering is used clustering the normalized data. Finally, the bagging algorithm is used for the prediction of the wind power and it also solves the overfitting problem.

A. A. DATA PREPROCESSING

The dataset consists of a large number of training samples. It may contain some unreasonable and noisy data. However,

the dataset consists of too many features it reduces the computational complexity. For reducing the features, some feature selection techniques are used. Finally, normalization has been done for the selected data. It will help in improving accuracy.

Data Cleaning

The original sample may consist of some noisy, unreasonable and incomplete data. The unusable data to be removed with the help of data cleaning. After removing the data where will empty space. For filling the empty space the mean value method is used.

Feature Selection

In these, there are huge amount of input variables are available in the dataset. For selecting the best feature linear correlation coefficient is used. It will find the similarity between the data in random. The correlation coefficient value to be between 1 are selected feature and other feature are not correlated features. The linear correlation coefficient can find some relevant features. The selected features are wind speed, wind direction, wind temperature, wind power.

Data Normalization

Normalization has been done for the selected features. It can improve the forecasting accuracy and it also helps in convert the raw data in to the most useful data. Therefore, normalization can be done by applying min-max method. It will be expressed by using the formula $=(x-min)/(x-max)$. The value of to be mapped between [0, 1]. The original data to be represented as x, minimum and maximum value to be represented as min and max value of the data.

B. CLUSTERING BASED ON SIMILARITY

Wind turbine produced a huge number of data when it operating casually. Because of the discontinuity of the wind where will be difficult in collecting the wind data. Then it will be difficult in obtaining better forecasting accuracy. For improving the forecasting accuracy fuzzy k-means clustering is considered to be the good method. This method helps in clustering similar days.

Fuzzy k-means clustering gives accurate forecasting result. It is an extension of the k-means clustering. K-means clustering means hard clustering in which each point of information belongs to one cluster centre whereas Fuzzy k-means clustering means soft clustering in which each point of information belongs to two or more cluster centre. It helps in clustering similar days and classifies the wind data into some category based on historical power data and weather condition.

The fuzzy set will help in calculating the membership value of each data or instance. The value of the fuzzy set to correspond to the cluster centre. Any value can set to a number of fuzzy sets based on its need. The one important point is that the fuzzy set number and the number of the cluster to be same. The fuzzy k-means clustering uses the fuzzification attributes are called m. The m value to be within the range [1, N] and it will determine the degree of the

fuzziness of the cluster. The m value to be set to 1 then the effect will be the crisp clustering of points into the clusters. If them value increases the degree of fuzziness between the data or the point increases.

The k-means clustering and the fuzzy k-means clustering are similar. The following process of the fuzzy k-means clustering is 1) First initialize centroid matrix with a random variable. 2) Calculate distance using Euclidean distance of all object with centroids. 3) Calculate the membership matrix. 4) Compute new cluster centroids and repeat till stopping criteria is set and again goes to step 2.

III. BAGGING NEURAL NETWORK

Data mining approach use the Neural Network (NN) for prediction. It will be considered as the more effective method for forecasting of wind power. One of the most commonly used prediction algorithms in NN is Back Propagation Neural Network (BPNN). It consists of three layers. The first layer is the input layer, the second layer is the hidden layer, the third layer is the output layer. The BPNN has some principle that will be shown in fig 2. It has two processes. In these, the first process is back propagation, the neuron in each layer can affect only in the next layer. The second process is forward propagation, it cannot obtain the expected output in the output layer. In these, the gradient descent method is used for calculating the weight vector space and where is the need for a dynamic search for the set of weight for reducing the error function.

In fig 2, the input attributes are [p1,p2,p3.....pn], the weight [q1,q2,q3.....qn] is between the first layer and the second layer. The weight [r1,r2,r3.....rn] is between the second layer and the third layer. The expected outcome is Y(t).

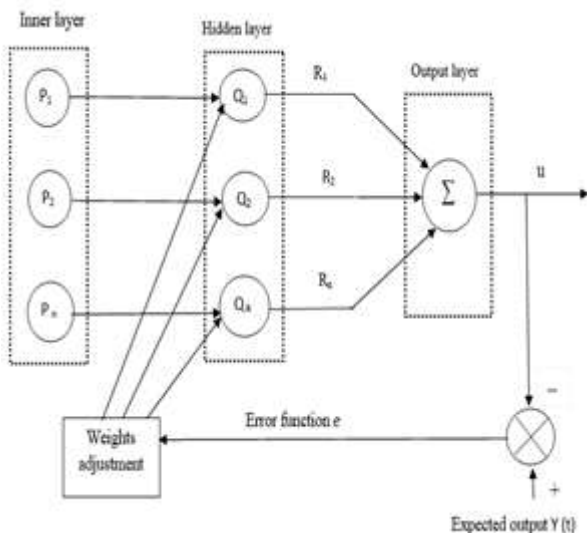


Fig. 2 Back Propagation Neural Network Model

To handle the nonlinear problem sigmoid function to be used as the transfer function. It will be expressed as

$$d = f(m) = \frac{1}{1 + e^{-m}} \quad (1)$$

The second layer output h can be formulated as

$$h = f_1(q_i p_i) \quad (2)$$

The third layer output can be formulated as

$$u = f_2(\sum c_i z_i) \quad (3)$$

Thus, the forward propagation process to be finished.

The function u and Y(t) produces the error signal e. It can be calculated as

$$e = \frac{1}{2} \sum (y_i(t) - u_i)^2 \quad (4)$$

The gradient descent method helps in the modification of the weight q_i and also helps in minimizing the error function e. For updating the weight we need to calculate the partial deviation. In this process there are several changes in the weight. It took lots of time for each modification.

When comparing with other algorithms like support vector machine and linear aggregation the back propagation neural network can obtain better forecasting result while dealing with huge amount of data. One of the disadvantages is that the overfitting problem will be arises in the BPNN. In order to solve the above problem, the Bagging Neural Network is used. In order to enhance the stability and also for improving the forecasting accuracy an ensemble learning method is used. In ensemble learning method bagging algorithm (bootstrap aggregation) as a kind of learning method and it helps in improving the performance of BPNN. The accuracy is poor because of the weak learner. The weak learner to be trained by the training sample continuously. It produces a sequence of predicting accuracy.

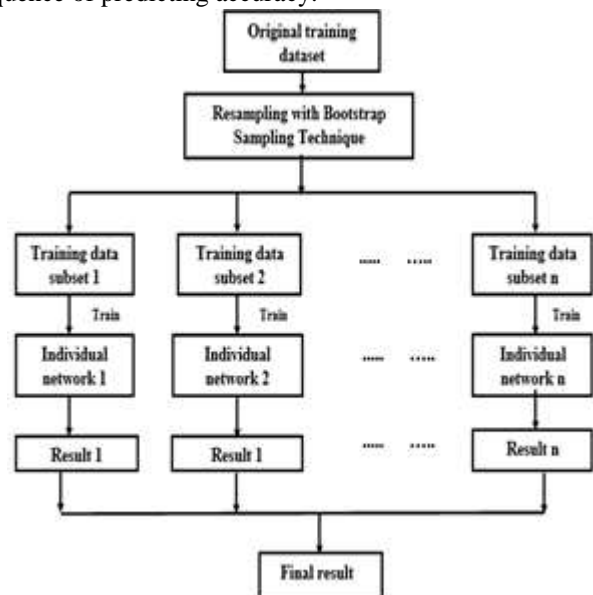


Fig. 3 Process of the Bagging Algorithm

Figure 3 shows the following process of the bagging algorithm. 1) trained set can be taken as the input. 2) the trained set to be resampled with the help of bootstrap sampling technique. 3) the sample to be divided into subsets. 4) each subset to be trained by the individual network. 5) the result consisting of forecasting model m₁, m₂, m_k is

WIND POWER FORECASTING IN SHORT-TERM USING FUZZY K-MEANS CLUSTERING AND BAGGING NEURAL NETWORK

produced. 6) finally, the forecasting model M is chosen by calculating the average value for the result. In these process, same importance will be given to all forecasting models

By applying this technique can improve the accuracy of the forecasting and also reduce the computational complexity

V. CONCLUSION

In this paper, bagging neural network and the fuzzy k-means clustering is the data mining technique proposed to provide a better forecast of the wind power. By applying this technique, the historical days are clustered based on the weather condition and historical power data. The bagging algorithm helps in solving the instability problem and it helps in improving the accuracy of the forecasting and also reduce the computational complexity.

REFERENCES

1. J. Jung and R. P. Broadwater, "Current status and future advances for wind speed and power forecasting," *Renew. Sustain. Energy Rev.*, vol. 31, no. 2, pp. 762-777, Mar.2014.
2. D. Carvalho, A. Rocha, M. Gómez-Gesteira, and C. Santos, "A sensitivity study of the WRF model in wind simulation for an area of high wind energy," *Environ. Model. Softw.*, vol. 33, no. 7, pp. 23-34, Jul. 2012
3. A. M. Foley, P. G. Leahy, A. Marvuglia, and E. J. McKeogh, "Current methods and advances in forecasting of wind power generation" *Renew. Energy*, vol. 37, no. 1, pp. 1-8, Jan.2012
4. Q. Xu et al., "A short-term wind power forecasting approach with adjustment of numerical weather prediction input by data mining," *IEEE Trans. Sustain. Energy*, vol. 6, no. 4, pp. 1283-1291, Oct.2015.
5. I. Colak, S. Sagiroglu, M. Yesibudak, E. Kabalci, and H. I. Bulbul, "Multi-time series and-time scale modeling for wind speed and wind power forecasting-Part I: Statistical methods, very short-term and short-term applications," in *Proc. IEEE ICRERA*, Palermo, Italy, Nov. 2015, pp.209-214.
6. C. Wan, Z. Xu, P. Pinson, Z. Y. Dong, and K. P. Wong, "Probabilistic forecasting of wind power generation using extreme learning machine," *IEEE Trans. Power Syst.*, vol. 29, no. 3, pp. 1033-1044, May2014.
7. M. Lei, L. Shiyan, J. Chuanwen, L. Hongling, and Z. Yan, "A review on the forecasting of wind speed and generated power," *Renew. Sustain. Energy Rev.*, vol. 13, no. 4, pp. 915-920, May2009.
8. Yuan-Kang, Po-En-Su, "Stratification-Based Wind Power Forecasting in a High-Penetration Wind Power System Using a Hybrid Model", *IEEE*, June-2016, Vol.53, NO. 3.
9. K. Bhaskar and S. N. Singh, "AWNN-assisted wind power forecasting using feed-forward neural network," *IEEE Trans. Sustain. Energy*, vol. 3, no. 2, pp. 306-315, Apr.2012
10. S. Li, P. Wang, and L. Goel, "Wind power forecasting using neural network ensembles with feature selection," *IEEE Trans. Sustain. Energy*, vol. 6, no. 4, pp. 1447-1456, Oct.2015.
11. X. Zhang, R. Wang, T. Liao, T. Zhang, and Y. Zha, "Short-term forecasting of wind power generation based on the similar day and Elman neural network," in *Proc. IEEE SSCI*, Cape Town, South Africa, Dec. 2015, pp.647-650.
12. E. Erdem, J. Shi, and Y. Peng, "Short-term forecasting of wind speed and power-A clustering approach," in *Proc. ISERC*, Toronto, ON, Canada, May 2014, pp.1-8.
13. Wenbin Wu and Mugen Peng, Senior Member, "A Data Mining Approach Combining K-Means Clustering With Bagging Neural Network for Short-Term Wind Power Forecasting" *IEEE internet of things journal*, vol. 4, no. 4, august 2017.
14. Prema V., Uma Rao K. (2015), "Development Of Statistical Time Series Models For Solar Power Prediction", *Renewable Energy*, Vol.83, Pp.100-109.