Research on Adaptive Selection Algorithm for Multi-model Load Forecasting Based on Adaboost

Kaiwen Zeng¹, *, Jianing Liu¹, Haizhu Wang¹, Zhengjun Zhao² and Chengsheng Wen²

¹ Guangdong Power Grid Power Dispatch Control Center, Guangzhou, 75 Meihua Road, Guangdong, 510000, China
² Guangzhou Zhiguang Automation Co., Ltd, Guangzhou, China

Email: * zengkaiwen@gddd.csg.cn

Abstract. With the development of energy Internet technology and artificial intelligence algorithms, smart grids have an increasing demand for the accuracy of load forecasting. This paper uses Support Vector Machine (SVM), Artificial Neural Network (ANN) and K-Nearest Neighbors (KNN) regression models to forecast electricity consumption of local residents and wind power generation of selected enterprises respectively. The electricity load of residents is divided into workday load and weekend load for comparative analysis. Based on Adaptive Boosting (Adaboost) algorithm, this paper proposes an adaptive selection method which further improves the prediction accuracy of the regression model, by selecting appropriate number of sub-classifiers and setting training weights of training data and sub-classifiers to adjust the distribution of training samples. Simulation results demonstrate that even in scenarios of poor data quality, the prediction accuracy of the model with the proposed Adaboost-based algorithm can be enhanced by up to 10%.

1 Introduction

In recent years, the growth rate of power demand has exceeded the growth rate of primary energy consumption, and electric energy substitution is the basic trend of global terminal energy utilization. In 2018, China's total electricity consumption increased by 8.4% over the previous year, and grid-connected wind power generation increased by 20.1%. It is estimated that China's primary energy consumption structure from 2035-2040 will be dominated by coal, oil and gas and new energy [1]. Load forecasting is an indispensable part of power system planning, operation and dispatch. The accuracy of load forecasting is closely related to the scientific rationality of power grid planning and operation. Improving the accuracy of load prediction can improve the economy and stability of power grid operation and dispatch while ensuring that users can obtain safe and reliable power supply.

Power system load forecasting is generally divided into four types: ultra-short-term load forecasting, short-term load forecasting, medium-term load forecasting and long-term load forecasting. The load forecasting mentioned in this article belongs to short-term load forecasting which refers to forecasting the load of the next day or week. The formulation of daily and weekly power generation plan in power plants is affected by the prediction results.

With the rise of artificial intelligence research trends, researchers have successfully applied SVM [2, 3], ANN [4, 5], Random Forests (RF) [6, 7], Extreme Learning Machine [8] and other machine learning algorithms into predicting power system load. A novel prediction method based on neural network and chaotic intelligent feature selection is proposed in [9]. The proposed feature selection method selects the best candidate input set as the predicted input data. The phase space reconstruction theory of the embedding theorem is used to screen the candidate features, and the correlation between the measured candidate input and the target value is analyzed. The prediction engine uses a Multi-layer Perceptual layer (MLP) to mix Levenberg-Marquardt (LM) and Differential...
Evolution (DE) learning algorithms. This short-term load forecasting method achieves superior performance in the test of PJM and New England power markets when compared with STLF technology. Literature [10] compares the performance of ANN and RF in hotel HVAC energy consumption prediction. The results show that these two models have considerable predictive power and almost the same applicability in building energy applications. [11] shows a new support vector regression algorithm by mixing the three algorithms of Empirical Mode Decomposition (EMD), Genetic Algorithm (GA) and Particle Swarm Optimization (PSO). The regularized structural risk function is optimized by the hybrid algorithm, which effectively improves the accuracy of short-term load forecasting. [12] indicates a load forecasting model based on K-nearest neighbors (KNN). The model only uses limited temperature forecasts and integrates three individual KNNs with different neighborhood rules through weighting factors. The results show that the load prediction results of the KNN algorithm in a limited information environment are still practical.

However, the previous studies mentioned above tend to use a single forecasting model or a simple weighted mixture of several single models. In addition, the selected model training data is sufficient and of high quality. The impact of poor data quality in regression model accuracy is ignored.

This paper selects the precise electricity load data of residential users in a certain area and the wind power data of poor quality in certain enterprises. The load prediction performance of a single SVM, ANN and KNN model in two scenarios is simulated and compared. Then, based on the Adaboost theory, an adaptive selection algorithm is proposed to make adaptive enhancement on the single prediction model respectively for higher accuracy by selecting appropriate number of sub-classifiers and setting training weights of training data and sub-classifiers. By adjusting the distribution of training samples in time, the training process of the model pays more attention to the moment when the prediction error rate is higher. Finally, the results of adaptive enhanced load forecasting in different scenarios are analyzed, and the accuracy and practicability of the proposed adaptive selection forecasting algorithm are verified.

2. Multi-model Load Forecasting Algorithm

In this section, single prediction model of SVM, ANN, and KNN is introduced and applied to forecast residential load and wind power generation.

2.1 SVM Model

The core of the SVM algorithm is to find a hyperplane that separates data into different classes. The algorithm classifies the data by detecting an optimal data. The hyperplane sought must meet the maximum boundary between all classes. The data points closest to the hyperplane are called support vectors.

In the SVM regression model, a kernel function is introduced to linearize the training samples. Through the mapping mechanism, the model converts the more complex vector inner product operations in the high-dimensional space into the operation of the kernel function in the low-dimensional space. The model also introduces a loss function \( \varepsilon \) to describe the deviation between the predicted value and the actual value after regression. The optimization function is defined as follows:

\[
\min \left\{ \frac{1}{2} \|w\|^2 + C \sum \left( \zeta_i + \zeta_i^* \right) \right\} \quad (1)
\]

subject to

\[
g(x) = w \cdot \phi(x) + b \quad (2)
\]

\[
y_i - g(x) \leq \varepsilon + \zeta_i^* \\
g(x) - y_i \leq \varepsilon + \zeta_i \\
\zeta_i, \zeta_i^* \geq 0
\]

where \( x \) denotes the output data, \( g(x) \) denotes the regression function, \( w \) and \( b \) denote the vector parameters, \( \phi(x) \) denotes the nonlinear mapping function, \( \zeta_i \) and \( \zeta_i^* \) denote the slack variables, \( C \) denotes the balance variable, \( \varepsilon \) denotes the loss function.

2.2 ANN Model

ANN is a calculation model that connects artificial neurons with the weights (coefficients) that make up the
neural network. The process of matching experimental data with neural networks is called training. The neural network is trained by iterative adjustment of weights. This paper uses a multi-layer perceptron (MLP) model based on error back propagation algorithm and gradient descent theory to predict short-term load, and the loss function uses a square error function. Assuming that $x$ denotes the input data, $y$ denotes the output data, $k$ and $\kappa$ denote the time coefficient. Figure 1 shows the structure of the ANN prediction model.

![Figure 1. Structure of ANN model](image)

The ANN model needs enough data to describe the connection between the input and the output, and the peak value of the input data affects the accuracy of the output. The main parameters of the ANN model are the size of the input vector, the number of layers and the number of neurons in each layer.

### 2.3 KNN Model

The core of the KNN algorithm is to find the K nearest neighbors to the new sample. The model measures sample similarity through a distance function, and uses most decision rules to classify and regress new samples. The algorithm relies on user-defined examples and user-defined distance metrics. In this paper, the Brute algorithm is adopted to search the neighbors of the target sample, and the weighted average of the nearest neighbors are applied to calculate the final prediction value. The flow chart of KNN algorithm is presented in Figure 2.

![Figure 2. Flow chart of KNN model](image)
3. Adaptive Selection Based on Adaboost

In this section, an adaptive selection algorithm is proposed to adaptively enhance the single prediction model mentioned in section 2 based on Adaboost algorithm. By choosing different numbers of sub-classifiers, setting the weights of the training data and the training weights of sub-classifiers, the distribution of training samples in the current iterative training session can be adjusted in time. Thus, the model pays more attention to the moments when the prediction error rate is high, and further improves the prediction accuracy of the original forecasting model during the training process.

3.1 Adaptive Selection Algorithm

The Adaboost algorithm improves the voting mechanism of algorithm decision-making based on the traditional boosting algorithm, so that the algorithm pays more attention to the error rate of the sample when it uses sample data for training. Based on the improvement method of the voting mechanism in the Adaboost algorithm, the proposed adaptive selection algorithm introduces weight variables to change the distribution of samples according to the error rate. The case study in section 4.2 proves that such improvement is beneficial to improve the accuracy of regression prediction.

In order to explain the process of the algorithm in detail, N sets of sample data \( T \) are introduced in this section. \( w_i \) denotes the single sub-classifier such as SVM, ANN and KNN.

\[
T = \{(x_i, y_i), (x_{i+1}, y_{i+1}), ..., (x_m, y_m)\} \tag{3}
\]

where \( x \) denotes the input data, \( y \) denotes the output data, \( M \) is the total number of single sub-classifiers.

The sample weight \( D_i (i \in N) \), which describes the influence of sample distribution on the classification results of the weak single classifiers, denotes the weight of the \( i \)-th sample set. Moreover, the classifier weight \( \alpha_i \) is applied to describe the weight of the weak single classifier in the final adaptively enhanced classifier.

\[
\alpha_i = \frac{\delta_i}{1 - \delta_i} \tag{5}
\]

where \( \delta_i \) denotes the regression error rate of the \( i \)-th classifier, \( \alpha_i \) denotes the weight of the \( i \)-th classifier.

The training sample weights need to be updated for enabling the classifier \( w_j \) to pay more attention to the data points that the classifier \( w_j \) predicts incorrectly. The sample weight update formula is defined as follows:

\[
D_i^{(i+1)} = \frac{D_i^{(i)}}{\sum_j (D_j^{(i)} \alpha_j^{(i)})} \tag{6}
\]

After the weight update, the iterative loop training enters the next round. Once the loop ends, the various classifiers are combined according to their weights to obtain an adaptively enhanced prediction model. The combination formula is as follows:

\[
W = \sum_j \left( \frac{1}{\alpha_j} \right) w_j(x) \tag{7}
\]

Table 1 shows the pseudocode of the adaptive selection algorithm.

**Table 1. Pseudocode of the adaptive selection algorithm**

|   |   |
|---|---|
| 1 | \( T = \{(x_i, y_i), (x_{i+1}, y_{i+1}), ..., (x_m, y_m)\} \) // Input training data |
| 2 | \( D = \{N, j \in N\} \) Initialize data weight distribution |
| 3 | for \( j = 1 \) to \( M \) do |
| 4 | use data with \( D_i \) weight distribution to train classifier \( w_j \) |
| 5 | update sample weight \( D_i \rightarrow D_i^{(i+1)} \) |
| 6 | increase the weight of data that is incorrectly predicted by the sub-classifier \( w_j \) |
| 7 | Reduce the weight of data that is correctly predicted |
5. by the sub-classifier \( w_j \)
6. \( \text{Calculate the weight of the weak classifier } \alpha_j \)
7. according to the regression error rate

9. end for

10. \( \text{Combine the sub-classifiers into an adaptive enhanced prediction model} \)

4. Case Studies

4.1 Data Processing

The case in this paper collects one-year electricity load data of 27 residential users who participated in the electricity household condition assessment project in region A of China, and the generator output data of the wind farm in region B of China. The start and end time of resident load data is from 00:00 on November 24, 2012 to 23:30 on November 23, 2013. The load data is collected every half hour, and the daily load data includes 48 time points. Since the residential load data in half an hour is relatively small, only about 0.3kWh, in order to improve the prediction accuracy as much as possible, this paper adds the electricity load data of 27 residential users at the same time to obtain a collective electricity load data set. This collective data set is used as the model historical training data set. The start and end time of wind power data is from 0:00 on March 1, 2018 to 23:57 on March 6, 2018. Wind power data is collected every 3 minutes, and daily wind power data includes 480 time points. Since the daily wind power data of the selected wind farm after 12 o'clock has a large degree of missing data with poor data quality, only the data of the first 240 time points of each day are selected for predictive analysis in the following case study.

This paper selects the root mean square error (RMSE) and the prediction error rate \( w \) as the evaluation index of the prediction result. The specific evaluation formula is defined as follows:

\[
\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^{N} \left( \text{predict}[i] - \text{real}[i] \right)^2}
\]

(8)

\[
\text{std} = \frac{\left| \text{predict}[i] - \text{real}[i] \right|}{\text{real}[i]}
\]

(9)

\[
w = \frac{\text{error_num}}{N}
\]

(10)

where \( \text{predict}[i] \) denotes the load forecast value at time \( i \), \( \text{real}[i] \) denotes the actual value of the load at time \( i \). \( \text{std} \) denotes the forecast accuracy index, if \( \text{std} > 0.2 \), the current point is regarded as a prediction error point. \( \text{error_num} \) denotes the number of prediction error points, \( N \) denotes the total number of prediction points.

4.2 Result Analysis

Considering two different scenarios of resident load forecasting and wind power generation forecasting, this paper first uses a single model of SVM, ANN and KNN to make forecast respectively. Resident load forecasting scenarios are subdivided into workday load and weekend load. The SVM model kernel function selects the radial basis kernel function. The ANN model is a three-layer neural network, the number of hidden layer neurons is set to 10, the excitation function uses a modified linear unit function, the weight adjustment uses the gradient descent algorithm, and the maximum number of iterations is set to 10,000. The nearest neighbor \( K \) of the KNN model is set to 5, and the Distance rule is used to determine the weight of the nearest neighbor samples.

Then, in order to further improve the prediction accuracy, this paper proposes the adaptive selection algorithm based on Adaboost to enhance the single model. The number of sub-classifiers \( \text{num_estimators} \) is set to 10 and 30 for comparison. For a single model, \( \text{num_estimators}=0 \). Finally, use the enhanced prediction model to predict the residential load and wind power generation. Figure 3 shows the results of the residents’ workday load forecasting. The blue curve in the figure represents the distribution of the true value of the
predicted point, and the red curve is the distribution of the predicted value. Table 2 indicates the summary of numerical analysis on forecast error rate $\omega$ and RMSE.

**Table 2. Workday load forecast data analysis**

| Model | num_estimators | $\omega$ | RMSE  |
|-------|----------------|---------|--------|
| SVM   | 0              | 0.2083  | 0.9400 |
|       | 10             | 0.2292  | 0.9047 |
|       | 30             | 0.1792  | 0.8986 |
| ANN   | 0              | 0.2708  | 0.7971 |
|       | 10             | 0.3125  | 0.9172 |
|       | 30             | 0.25    | 0.8907 |
| KNN   | 0              | 0.2917  | 1.0908 |
|       | 10             | 0.3333  | 1.1561 |
|       | 30             | 0.2958  | 1.1452 |

![Figure 3. Residential load forecast results (workday)](image)

It can be seen from Figure 3 and Table 2 that the SVM model has the lowest prediction error rate among the whole prediction results of a single model. After adaptively enhancing a single prediction model, the increase in the number of sub-classifiers leads to a decrease in the prediction error rate of the model. When the number of sub-classifiers is 30, the enhanced SVM model performs best, the prediction error rate is reduced by about 3%, the RMSE is reduced by about 4.4%, and the deviation between the prediction result and the true value is further reduced. The effectiveness of the adaptive selection algorithm based on Adaboost is proven.

Figure 4 shows the results of residents’ weekend load forecasting. Table 3 indicates the summary of
numerical analysis on forecast error rate and RMSE.

Table 3. Weekend load forecast data analysis

| Model | num_estimators | w  | RMSE   |
|-------|----------------|----|--------|
| SVM   | 0              | 0.4583 | 1.2126 |
|       | 10             | 0.4375 | 1.1356 |
|       | 30             | 0.4167 | 1.1265 |
| ANN   | 0              | 0.5625 | 1.3679 |
|       | 10             | 0.5417 | 1.3708 |
|       | 30             | 0.5   | 1.3185 |
| KNN   | 0              | 0.5208 | 1.1816 |
|       | 10             | 0.4583 | 1.2378 |
|       | 30             | 0.4374 | 1.1932 |

Figure 4. Residential load forecast results (weekend)

It can be seen from Figure 4 and Table 3 that the prediction error rate of the SVM model is still the lowest among the prediction results of a single model. After adaptive enhancement of a single prediction model, the model prediction error rate is generally reduced, which proves that the prediction results are indeed optimized. However, the error rate of residents’ weekend load forecasts is higher than that of weekday load forecasts. The reason may be that the residents’ electricity load on weekends has strong randomness and volatility. The forecasting difficulty of weekend load is higher than that of working days.
Figure 5. Wind power generation forecast results

Different from the resident load forecasting scenario, the wind power generation scenario is more random. This paper selects poor quality wind power data, which only contains 6 days of data. Figure 5 shows the curve of wind power forecast results. Table 4 indicates the summary of numerical analysis on forecast error rate and RMSE. In this scenario, the results show that the improvement of prediction accuracy is higher than that of the resident load scenario.

Table 4. Wind power generation forecast data analysis

| Model | num_estimators | w    | RMSE    |
|-------|----------------|------|---------|
| SVM   | 0              | 0.5375 | 21.8270 |
|       | 30             | 0.4167 | 17.7900 |
| ANN   | 0              | 0.7500 | 31.9827 |
|       | 30             | 0.6450 | 18.4081 |
| KNN   | 0              | 0.5458 | 21.9744 |
|       | 30             | 0.5375 | 17.5027 |

Figure 5 and Table 4 indicates that the model prediction error rate is reduced generally after adaptive enhancement. The prediction accuracy of the enhanced SVM model is improved by about 12%, and the RMSE is reduced by about 18.5%. The result proves that the adaptive selection algorithm can achieve higher accuracy improvement in the case of poor data quality.

5. Conclusions

This paper focuses on two typical scenarios of residential load forecasting and wind power generation. SVM, ANN and KNN machine learning models are used to predict load and power generation. The adaptive selection algorithm based on Adaboost is proposed to enhance the single prediction model for further improvement of prediction accuracy. The conclusion is summarized as follows:

1) The Adaboost-based adaptive selection algorithm effectively improves the prediction accuracy of a single model. As the number of sub-classifiers increases, the prediction accuracy of the Adaboost-based adaptive selection model is significantly improved.

2) In the scenario with poor data quality, the prediction accuracy of the model with the proposed Adaboost-
based algorithm is improved by up to 10% which indicates a better generalization performance of the adaptive enhancement prediction model than the single model.

Acknowledgement
The work is sponsored by the scientific and innovation project of Guangdong Power Grid Corporation under the grand number GDKJXM20180574.

References
[1] International Energy Agency. World Energy Outlook 2017 [EB/OL]. https://www.iea.org/reports/world-energy-outlook-2017.
[2] Barman, M., Choudhury, N. D., & Sutrathar, S. A regional hybrid GOA-SVM model based on similar day approach for short-term load forecasting in Assam, India. Energy, 2018, 145, 710-720.
[3] Xia, C., Zhang, M., & Cao, J. A hybrid application of soft computing methods with wavelet SVM and neural network to electric power load forecasting. Journal of Electrical Systems and Information Technology, 2018, 5(3), 681-696.
[4] Raza, M. Q., Nadarajah, M., Hung, D. Q., & Baharudin, Z. An intelligent hybrid short-term load forecasting model for smart power grids. Sustainable Cities and Society, 2017, 31, 264-275.
[5] Kumar, S., Mishra, S., & Gupta, S. (2016, February). Short term load forecasting using ANN and multiple linear regression. In 2016 Second International Conference on Computational Intelligence & Communication Technology (CICT) (pp. 184-186). IEEE.
[6] Zhang, R., Xu, Y., Dong, Z. Y., Kong, W., & Wong, K. P. (2016, July). A composite k-nearest neighbor model for day-ahead load forecasting with limited temperature forecasts. In 2016 IEEE Power and Energy Society General Meeting (PESGM) (pp. 1-5). IEEE.
[7] Li, W., Kong, D., & Wu, J. A novel hybrid model based on extreme learning machine, k-nearest neighbor regression and wavelet denoising applied to short-term electric load forecasting. Energies, 2017, 10(5), 694.
[8] Li, T., Qian, Z., & He, T. (2020). Short-term load forecasting with improved CEEMDAN and GWO-based multiple kernel ELM. Complexity, 2020.
[9] Kouhi S, Keynia F, Ravadanegh S N. A new short-term load forecast method based on neuro-evolutionary algorithm and chaotic feature selection[J]. International Journal of Electrical Power & Energy Systems, 2014, 62: 862-867.
[10] Ahmad, M.W.; Moursheed, M.; Rezgui, Y. Trees Vs Neurons: Comparison Between Random Forest and ANN for High-resolution Prediction of Building Energy Consumption. Energy Build. 2017, 147, 77–89.
[11] Fan G F, Peng L L, Zhao X, et al. Applications of Hybrid EMD with PSO and GA for an SVR-Based Load Forecasting Model[J]. Energies, 2017, 10(11): 1713.
[12] Zhang, R., Xu, Y., Dong, Z. Y., Kong, W., & Wong, K. P. (2016, July). A composite k-nearest neighbor model for day-ahead load forecasting with limited temperature forecasts. In 2016 IEEE Power and Energy Society General Meeting (PESGM) (pp. 1-5). IEEE.