Short-Term Load Forecasting of Residential Electricity Consumption by Combining BP Neural Network with K-means Clustering Algorithm

Huayu Zhang [a], Jianchang Luo [b]

Chongqing University of Posts and Telecommunications, Chongqing 400065, China

[a]Huayu_zhang@126.com, [b]user_luo@foxmail.com

Abstract

In order to improve the accuracy of load forecasting for residential consumption, a method by combining BP neural network with K-means clustering algorithm is proposed. Firstly, the clustering analysis for electricity using datasets is performed, then, choosing the corresponding data of user and carrying out the forecasting task. Finally, the simulation experiment based on the datasets of 300 households is implemented. The result shows that the new method has a lower prediction error over traditional BP neural network.

Keywords

Load forecasting, BP neural network, K-means algorithm.

1. Introduction

With the rapid development of smart grid, residential electricity consumption is gradually intelligent. At the same time, the behavior of users has some regularity. Therefore, through the analysis of historical electricity consumption data, it is of great significance for grid companies to implement a reasonable and scientific power allocation strategy. Besides, electricity load forecasting is an important part of the smart grid system, and reliable forecasting can also help the efficient operations for electricity resource allocation[1][2]. Aimed at the load forecasting method in smart grid, some scholars carry out a great deal of research work. A power load forecasting method based on RBF neural network is proposed in reference [3], and the problem of over-fitting in training prediction is solved by reducing the number of hidden neurons in neural network. What's more, a smart grid-based energy management testing system is designed in reference [4], realizing the task of real-time energy management and load forecasting. Saleh proposed the method that combining genetic feature-based selector with rough set basic feature selector to select effective features and achieve load forecasting task in reference [5]. Besides, a loT-based deep learning system was proposed to automatically extract the features from the captured data in reference [6], and reached the intent of accurately estimate for future load. The author applied the technique of sparse coding to load forecasting, and used sparse coding to create a model and predict the personal home electricity load in reference [7].

At present, the above research still has some shortages, and the accuracy and stability of load forecasting needs to be further improved. Therefore, a method by combining BP neural network with K-means clustering algorithm[8][9] is proposed in this paper.

2. Short-Term Load Forecasting method of Residential Electricity Consumption

The method by combining BP neural network with K-means clustering algorithm is proposed. Firstly, through the K-Means algorithm, the clustering analysis of electricity data is completed, and the users are classified according to the specified relevant features, at the same time, the user-mode library is established. Then the corresponding model library is selected in the load forecasting stage, and the load forecasting is done based on BP neural network. Through the K-Means algorithm clustering analysis of different electricity using behavior and taking the different characteristics of each

category of users into account, it improves the accuracy of load forecasting and it is an important guide for the formulation of grid dispatch strategy.

The overall framework of the forecasting method is shown as Fig.1.

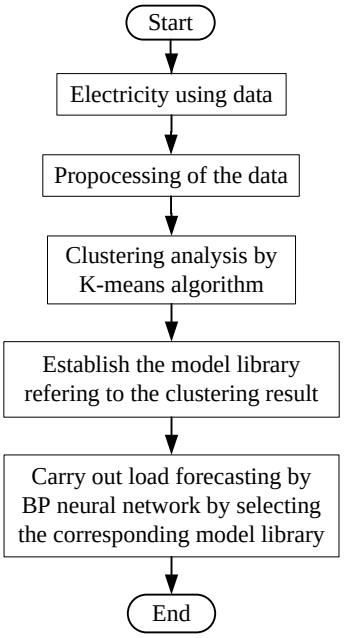


Figure 1 The overall framework of the forecasting method

The specific load forecasting process is as follows.

Step 1 The collection of residential electricity consumption data and data preprocessing.

The smart meter collects the user's actual power consumption data and the corresponding preprocessing on the data is performed, including vacancy processing, abnormal data removal, and the duplicate data removal.

Step 2 The clustering analysis of electricity data based on K-Means Algorithm.

Due to the large amount of data on historical electricity consumption of residents, the analysis of electricity consumption data for each household will increase the overall calculation and reduce the accuracy of prediction. The types of users in residential areas have shown a certain regularity of electricity using. Therefore, each similar type of users are grouped into the same category by K-Means clustering algorithm, and the feature mode library of each type of user is established. When carrying out load forecasting for a certain user, you can select the appropriate model library as a reference, so as to improve the accuracy of prediction.

Step 3 Electricity load forecasting based on BP neural network

When carrying out load forecasting by BP neural network, at first, we need to set the relevant parameters of neural network: the number of BP neural network layers (input layer, hidden layer and output layer), the number of input nodes in input layer, neural network learning and training times, Neural network weight and threshold and prediction error function. Then, input the processed sample dataset, calculate the output of the hidden layer and the output layer according to the initial weight and the threshold value, and calculate the corresponding error by comparing the output of the output layer with the expected value. If the error is within the setting range, the training is ended, and the relevant weight and threshold in the learning process are recorded. Otherwise, the error is propagated backwards to adjust the weights and thresholds, and the training continues until the error satisfies setting the conditions. Finally, the forecast of the electricity load is based on the weights and thresholds that are saved. The flow chart of short-term residential electricity load forecasting by BP neural network is shown as Fig.2.

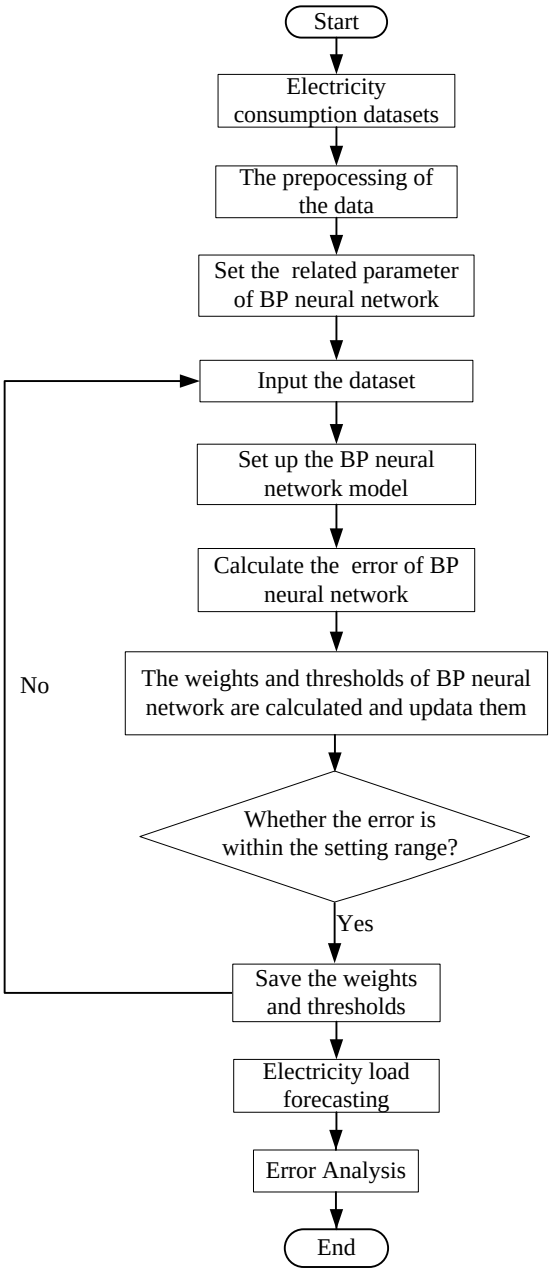
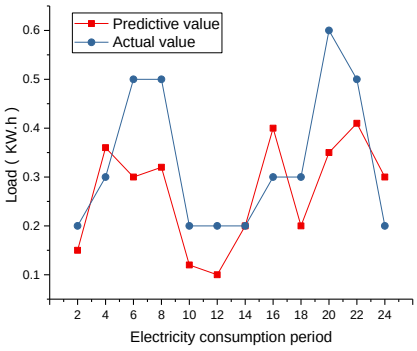


Figure 2 The flow chart of short-term residential electricity load forecasting by BP neural network

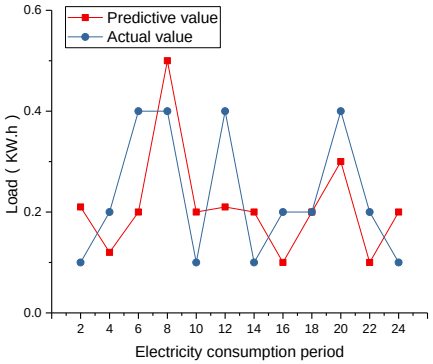
3. Simulation experiment and result analysis

In this section, the following comparative experiments are completed referring to the method by combining BP neural network with K-means clustering algorithm.

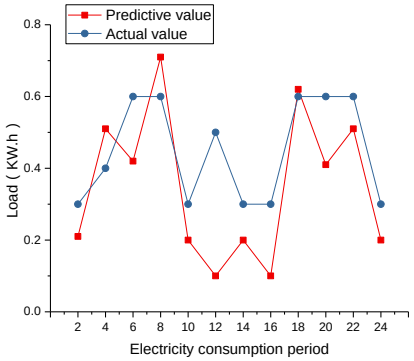
(a) Experiment 1: Select four types of users, and carry out load forecasting simulation respectively. The experiment includes three neurons: input layer, hidden layer and output layer. And the number of input layers is 4, the number of output layers is 1, the number of learning and training is 400, the minimum allowable output error function is 10^{-5}, learning and training rate is 0.05, the number of outer loop training is 500. As shown in Fig.3, the relevant comparison results are obtained according to the simulation experiment.



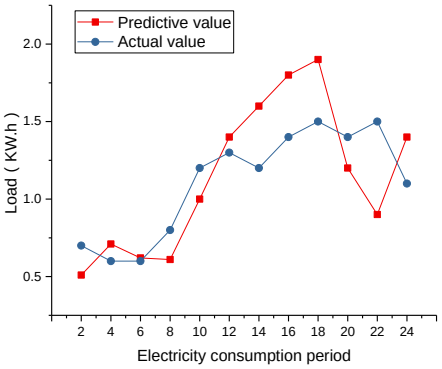
(a) Load forecasting for Class one



(b) Load forecasting for Class two



(c) Load forecasting fo Class three

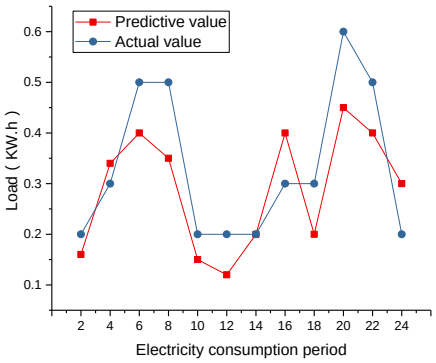


(d) Load forecasting fo Class four

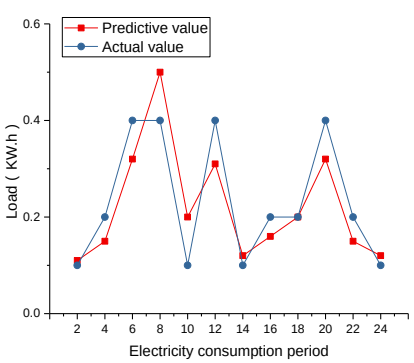
Figure3 The unpredicted BP neural network load forecasting results of all types of users

The unpredicted BP neural network load forecasting results of all types of users are shown in Fig.3. According to the forecasting results, the MAPE of each type of forecasting result can be calculated. The average absolute errors of the prediction results of the first-class users, the second-class users, the third -class users and the fourth-class users are 0.31, 0.583, 0.363 and 0.229 respectively.

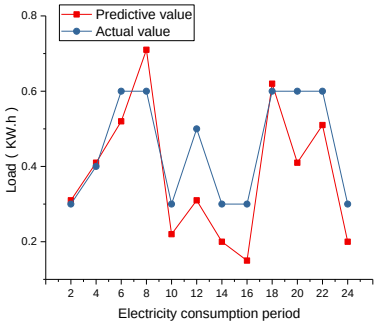
(b) Experiment 2: In the experiment of this part, K-Means clustering analysis is carried out before load forecasting. Select the corresponding historical electricity data as the training datasets according to the characteristics of the current prediction object in the model library. At the same time, the relevant parameters of BP neural network are obtained and load forecasting is eventually achieved. In order to avoid the influence of parameter setting on the prediction result, the parameter of BP neural network is the same as experiment 1. As shown in Fig.4, the load forecasting results that by combining BP neural network with K-means clustering algorithm are obtained.



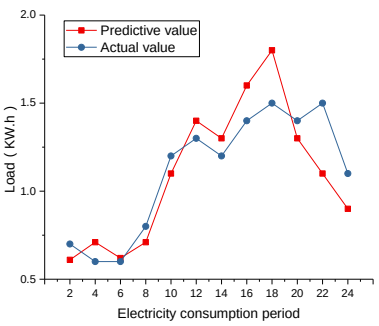
(a) Load forecasting for Class one



(b) Load forecasting for Class two



(c) Load forecasting for Class three



(d) Load forecasting for Class four

Figure 4 The load forecasting results that by combining BP neural network with K-means clustering algorithm

From the analysis of the prediction results in Figure 5.5, it can be seen that the average absolute error MAPE of the predicted results of the first-class users, the second-class users, the third -class users and the fourth-class users are 0.267, 0.306, 0.249 and 0.201 respectively.

According to the simulation results of Experiment 1 and Experiment 2, it can be known that the average absolute error of BP neural network load forecasting results by using K-Means clustering is higher than that of non-clustering method. And the MAPE values of the prediction results for each types of users are decreased by 13.9%, 47.5%, 31.4% and 12.2% respectively. Thus, the K-Means clustering method is effective in improving the accuracy of the prediction results.

4. Conclusion

The method by combining BP neural network with K-means clustering algorithm is proposed in this paper. Firstly, the clustering analysis for electricity using datasets is performed, then, choosing the corresponding data of user and carrying out the forecasting task. Finally, the simulation experiment based on the datasets of 300 households is implemented. The result shows that the new method has a lower prediction error over traditional BP neural network.

References

[1] Kononenko I, Kukar M. Preface - Machine Learning and Data Mining [J]. Journal of Shijiazhuang Vocational Technology Institute, 2016, 2(2):110-114.
[2] Wu X, Kumar V, Quinlan J R, et al. Top 10 algorithms in data mining [J]. Knowledge & Information Systems, 2007, 14(1):1-37.
[3] Skolthanarat S, Lewlomphaisarl U, Tungpimolrut K. Short-term load forecasting algorithm and optimization in smart grid operations and planning[C]// Technologies for Sustainability. IEEE, 2015:165-171.
[4] Thiyagarajan K, Kumar R S. Real Time Energy Management and Load Forecasting in Smart Grid Using CompactRIO [J]. Procedia Computer Science, 2016, 85:656-661.
[5] Saleh A I, Rabie A H, Abo-Al-Ez K M. A data mining based load forecasting strategy for smart electrical grids[J]. Advanced Engineering Informatics, 2016, 30(3):422-448.
[6] Li L, Ota K, Dong M. When Weather Matters: IoT-Based Electrical Load Forecasting for Smart Grid[J]. IEEE Communications Magazine, 2017, 55(10):46-51.
[7] Yu C N, Mirowski P, Ho T K. A Sparse Coding Approach to Household Electricity Demand Forecasting in Smart Grids[J]. IEEE Transactions on Smart Grid, 2017, 8(2):738-748.
[8] Wu D, Zhang Y, Yang F, et al. Improved k-means algorithm based on optimizing initial cluster centers. Icic Express Letters, 2013, 7(3):991-996.
[9] Li S, Zeng G, Ke Q, et al. Fast approximate k-means via cluster closures [J]. 2017, 157(10):3037-3044.