

Application of Federated Learning Algorithm Based on K-Means in Electric Power Data

Weimin He and Lei Zhao*

State Grid Jiangsu Electric Power Co., Ltd., Nanjing, 210022, China

*Corresponding Author: Lei Zhao. Email: hgfs2009@qq.com

Received: 03 June 2022; Accepted: 03 July 2022

Abstract: Accurate electricity forecasting is the key basis for guiding the power sector to arrange operation plans and guaranteeing the profitability of electric power companies. However, with the increasing demand of enterprises and departments for data security, the phenomenon of “Isolated Data Island” becomes more and more serious, resulting in the accuracy loss of the traditional electricity prediction model. Federated learning, as an emerging artificial intelligence technology, is designed to ensure data privacy while carrying out efficient machine learning, which provides a new way to solve the problem of “Isolated Data Island” in terms of electricity forecasting. Nonetheless, due to the popularity of smart meters, the collected electricity data presents the characteristics of uneven distribution and huge data volume, so it is difficult to apply the electric quantity prediction model generated only by federated learning in practice. To solve this problem, a clustering federated learning method (C-FL) is proposed to protect data privacy while improving the accuracy of power prediction. Firstly, C-FL uses K-means algorithm to cluster power data locally in power enterprises, and then builds accurate power forecasting models for each class of power data combined with other local clients through federated learning. A large number of experimental results show that the clustering federated learning method proposed in this paper is superior to the existing federated learning models in terms of the accuracy of electric power forecasting.

Keywords: Electricity forecast; federated learning; K-means; data security

1 Introduction

Electric quantity forecasting is an important work for electric power enterprises, which is of irreplaceable significance for making strategic development plan, marketing strategy, rational resource allocation and promoting economic and social development of electric power grid [1]. At the same time, accurate prediction of electric power is also the focus of power enterprise development, which has a direct impact on the rationality of transmission, generation and distribution of electric power. Therefore, in the development process of electric power enterprises, we must pay attention to the prediction of electric power, actively use advanced prediction technology, improve the accuracy



This work is licensed under a Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

of electric power prediction, and lay a solid foundation for the stable, orderly and coordinated economic development of electric power enterprises [2]. At present, the commonly used methods of electric quantity prediction include grey prediction method, electric quantity output benefit method, electric elasticity coefficient method and regression analysis method [3]. These methods can predict the electric quantity effectively when the data is completely shared.

In June 2017, the State promulgated the Cyber Security Law of the People's Republic of China, proposing that "construction and operation units shall take technical measures and other necessary measures to maintain the integrity, confidentiality and availability of data". In May 2018, the State issued the Information Security Technology Personal Information Security Specification, which regulates the use and sharing of personal information, making it clear that "in principle, personal information shall not be shared, and when it is really needed to be shared, full attention should be paid to risks, and the identity and data security capability of the recipient should be evaluated". In 2019, the Law of the People's Republic of China on Network Data Security has entered the formulation stage, which will further clarify the legal responsibility of data security. These bills show that personal information protection and data security are getting more and more national attention, the demand for data security compliance will be further increased, and the network security market is also expected to continue the rapid growth. However, the phenomenon of "Isolated Data Island" is becoming more and more obvious, which restricts the value output of power data and makes it difficult for traditional data sharing technology to meet the demand. Therefore, how to alleviate the problem of "Isolated Data Island" in power enterprises and establish a high-precision electricity prediction model through each "Isolated Data Island" while protecting data security is an urgent problem to be solved.

Federated learning is an emerging artificial intelligence technology, first proposed by Google in 2016, its design goal is to carry out efficient machine learning among multiple participants or multiple computing nodes on the premise of ensuring information security in big data exchange, protecting the privacy of terminal data and members' data, and ensuring legal compliance. The pioneer of this technology in China is WeBank, and it takes the lead in using federated learning to solve the problem of cross-department and cross-enterprise data fusion, and makes use of NVIDIA GPU resources from TenCent cloud. The federated learning technology has been used to solve the difficult problem of loans for more than 70% small and micro enterprises without historical credit information, to support corporate lending has more than 1 billion. Therefore, federated learning is applied in the field of power big data, which can carry out multi-party data calculation with high security on the premise of ensuring that all internal data of the parties do not leave the local, and carry out effective prediction of electric power to some extent.

However, the existing power prediction model based on federated learning still cannot be directly applied to the actual, the main reason is that when local power data is processed by power enterprise for federated learning, due to the power customer implements different marketing strategies, it is necessary to make an accurate grouping of electric power customers. Otherwise, the parameters of the model updated by alternating communication and operation between controlling parties are not accurate enough, and the gap between the obtained model and the non-federated model which aggregates the data is too large [4]. Because federated learning has the characteristics of solving the problem of "Isolated Data Island", it is better for training data sets with uniform data distribution, but this paper is based on the fact that the distribution density of power customer data is not balanced, and presents a clustering federated learning method based on K-means to build electricity consumption forecast model. This method firstly clusters these unevenly distributed power data with a large amount of data locally, and then federated learning is carried out on the power data of each cluster, which will improve the local power data training gradient, then, the precision of power forecasting model under federated

learning is optimized. K-means clustering algorithm, as a commonly used data mining technique for customer classification, proposes a weighted clustering criterion. For a given sample set, the sample set is divided into K clusters according to the distance between the samples, let the points in the cluster as close together as possible, and let the distance between the clusters as large as possible. The clustering results applied to the power data with federated learning show that the clustering algorithm is suitable for the actual operation data and achieves the effect of clustering compactness, better clustering results can also ensure that the subsequent federated learning process for each class of power data combined with machine learning to build accurate power forecasting model, and then predict more accurate power data, eventually bring higher returns to power supply enterprises. In the simulation environment, this paper compares the accuracy of the power forecasting model under the traditional federated learning, a large number of experimental results show that the clustering federated learning method proposed in this paper is superior to the existing federated learning models in terms of the accuracy of electric quantity prediction.

2 Related Work

Power forecasting is an important part of the power generation plan of the power system. It is the necessary prerequisite for the reasonable arrangement of power generation, transmission and electric energy distribution, and the basis of the operation of the power system. Therefore, it is of great significance to improve the accuracy of the power forecasting. A gate-controlled circulation unit (GRU) model combined with STL decomposition method is proposed by the College of Computer Science and Technology of Shanghai Electric Power University, using the STL (Seasonal-Trend decomposition procedure based on Less) decomposition method [5] to decompose the raw power data into three parts, reducing the interference between different parts, Global GRU layer and Local GRU [6] are also designed and used, to ensure better capture of global and local linear independence, the effective prediction within the time series is realized [7,8]. Its Chinese Xinjiang electric power co., LTD., based on the improved MGM, has realized a relatively reliable long-term forecast of provincial electricity consumption [9,10]. Combined with the total electricity consumption data of the whole society in a province, it analyzes the change and development trend of relevant influencing factors [11], through the analysis of correlation between factors and the current economic and social development needs, the correction formula of the background value in the traditional MGM(1, m) model is improved [12], makes to power data structure model of the improved samples of the sequence, The final experiment predicted the result of the total electricity consumption of the whole society in a certain province from 2017 to 2018, showing an accurate result with an average error of less than 1% [13]. In order to improve the accuracy of short-term load forecasting, the Economic and Technological Research Institute of State Grid Liaoning Electric Power Co. , Ltd. has proposed a multi-scale information fusion convolutional neural network model based on deep learning technology [14], the full convolution network structure and causal logic constraints are introduced to enhance the feature expression of time series, and the residual network structure is designed to increase the depth of the network, so as to increase the receiving region of the output neurons and improve the prediction accuracy [15,16].

As the sound of personal information protection and data security becomes higher and higher, the demand of data security compliance further increases, the government has issued a number of data privacy protection policies in recent years. Data becomes more secure, but the phenomenon of "Isolated Data Island" becomes more and more obvious, which restricts the value output of power data and makes traditional data sharing technology difficult to meet the demand, accurate power forecasting is a huge challenge. In view of the "Isolated Data Island" problem among various systems

and the security problem of the electric power data in the development of the artificial intelligence application in the electric power Internet of things, the implementation of cross-industry cross-service data centralized training existing practical problems and policy bottlenecks [17,18], Beijing Zhongdian Puhua Information Technology Co., Ltd. introduced the federated learning technology, and based on the architecture of ubiquitous electric Internet of things, designed strategies and basic processes for using federated learning in various application circumstances [19]. The College of Computer Science and Technology of Jilin University proposed a COVID-19 chest CT image segmentation method based on federated learning in response to the reality that the patient sample data were small, distributed in different institutions and not shared with each other. The experimental results showed that the test effect of this method was good after training, and the DICE index could reach 63.26%, which was helpful for the diagnosis of COVID-19 [20].

The distribution of data greatly affects the performance of federated learning algorithm. Although the federated learning can be used to predict the electricity quantity when the data are not visible to each other, because of the uneven distribution of the electricity data and the huge amount of data, there is a problem of precision loss in the traditional federated learning model. If this problem is not solved, the precision of power forecasting will be very low. Therefore, this article launches the research aiming at this aspect.

3 Electricity Forecasting Model Based on Clustering Federated Learning

3.1 Overview of Federated Learning

As digital technology enters a period of rapid development, technologies such as big data and artificial intelligence are experiencing explosive growth [21,22], which brings new opportunities for upgrading and transformation of traditional business forms, on the other hand, it inevitably brings new challenges to data and network security, and the problem of “Isolated Data Island” is one of the key challenges. Vertically, the top companies in the industry monopolize a huge amount of data, which is often hard for small companies to access, leading to ever widening hierarchies and gaps between firms. Horizontally, companies in different industries at the same level, because of the blocking of the system and business, it is difficult to realize the exchange and integration of data and information. Against this background, the federated learning technology pioneered by Google came into being. Federated learning is a concept pioneered by Google research in 2016 [23–25]. This technique completes the joint modeling without sharing the data. Specifically, data owned by individual data owners (individuals/enterprises/organizations) will not leave the local, by using the method of parameter exchange under the encryption mechanism in the federal system (that is, under the condition of not violating the data privacy law), a global sharing model is jointly established, and the established models only serve the local target in their respective regions. As an end-side artificial intelligence algorithm, federated learning has the advantages of protecting data privacy and high operational reliability, and has gradually become one of the development trends of model training.

3.2 K-means Algorithm

The K-means algorithm divides data into predetermined class number K on the basis of minimizing the error function, and uses distance as the similarity evaluation index, that is, the closer the two objects are, the greater the similarity [26].

1. Clustering criterion function

$$E = \sum_{i=1}^k \sum_{j=1}^{n_i} \|x_{ij} - m_i\|^2 \quad (1)$$

where n_i is the number of samples for the i^{th} class; x_{ij} represents the j^{th} sample in the i^{th} class; m_i represents the cluster center of Class i .

2. Algorithm process

- 1) K objects are selected randomly from the data set (N) as initial cluster centers;
- 2) Calculate the distance from each sample to each cluster center, and assign the object to the nearest cluster;
- 3) Recalculate the centers of K clusters after all objects have been allocated:

$$m_i = \frac{1}{n_i} \sum_{j=1}^{n_i} x_{ij} \quad (2)$$

- 4) Compared with the K cluster centers obtained from the previous calculation, if the cluster centers change, rotate 2), otherwise rotate 5);
- 5) When the center of mass does not change, the cluster results are stopped and output.

The K-means algorithm can often be applied to data sets that are small in dimension and small in value, such as grouping the same objects from a randomly distributed set of objects. It can be used in the following scenarios:

- 1) Document classifier: Classify documents into different categories based on tags, topics, and document content.
- 2) Customer segmentation: Clustering helps marketers improve their customer base (working within their target area) and further segment customers based on their purchase history, interest, or activity monitoring.
- 3) Automated clustering of IT alerts: Large enterprise IT infrastructure technology components (such as networks, storage, or databases) generate a large number of alerts. Because alert messages can point to specific operations, they must be filtered manually to ensure the priority of the follow-up process. Clustering the data can help to understand the alarm category and the average repair time, and help to predict the future failure.

3.3 C-FL Method Framework

Combined with the advantages of traditional federated learning to protect data security, this paper designs a clustering federated learning (C-FL) power prediction method based on K-means, which mainly includes three processes, as follows: Fig. 1 shows the process of power data clustering by using K-means algorithm in C-FL; Fig. 2 shows the process of using clustered power data with other local clients through federated learning to train multiple models for power prediction in C-FL; Fig. 3 shows the process of using multiple models to predict power consumption by calculating Euclidean distance in C-FL. 3.3.1–3.3.3 will describe the three processes in detail.

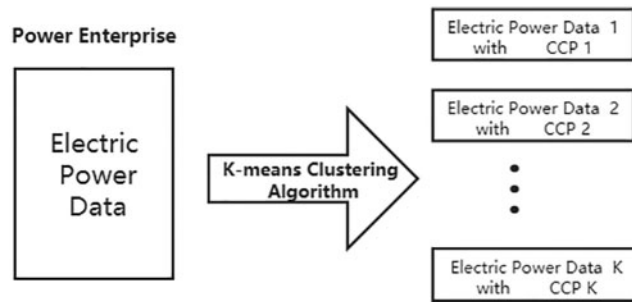


Figure 1: C-FL clustering operation process

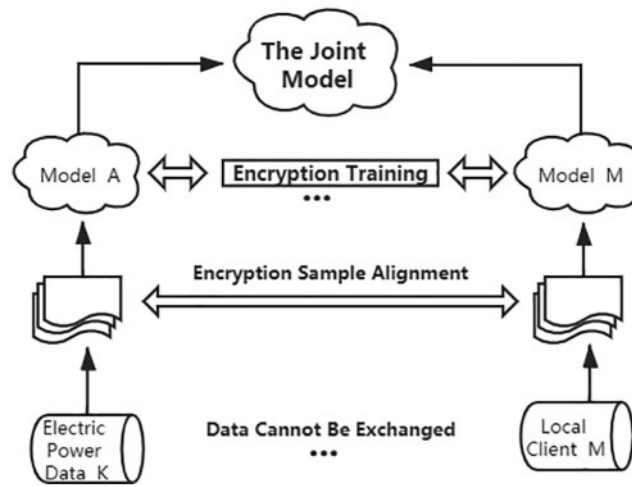


Figure 2: C-FL multi-model training process

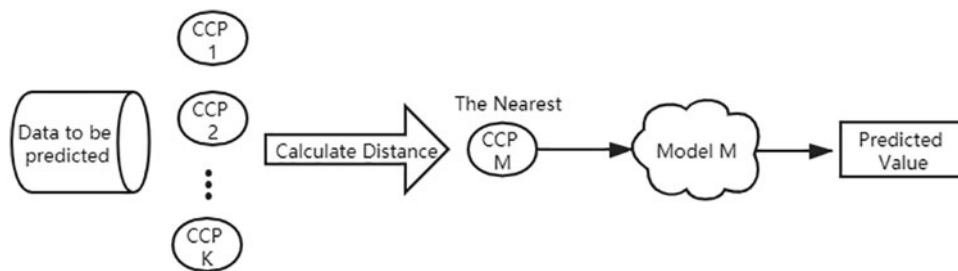


Figure 3: C-FL multi-model usage process

3.3.1 Clustering

Because federated learning has the characteristic of solving the problem of “Isolated Data Island”, it has a good effect on training data sets with uniform data distribution, but the paper is based on the fact that the distribution density of power customer data is not balanced, a federated learning technique based on K-means clustering algorithm is proposed.

Before Federated learning, power enterprises cluster the stored power data locally by using K-means clustering algorithm to generate K electric power data and K cluster center points (marked as CCP K). Because of the characteristic of K-means clustering algorithm, the size of K value should be specified, that is, how many classes the data is clustered into. For the selection of K values, it is important to scale each cycle within the same range so that the size of the energy load does not interfere with the selection of the cluster.

3.3.2 C-FL Multi-model Training

Clustering operation has classified local power data into K class which is most suitable while ensuring data security. Next, we need to unite M local clients which hope to participate in this joint modeling, performing federated learning with K classes of power data separately. A class of electric power data sets up a power forecasting model jointly by many parties, to generate multi-model, in order to achieve the goal of subsequent accurate forecasting.

Because of the particularity of electric power enterprises in China, the power forecasting model based on horizontal cluster federated learning is of no practical significance. Therefore, this paper introduces the training phase of C-FL multi-model with the example of vertical clustering federated learning model.

The nature of vertical clustering federated learning is a combination of features that can be used in scenarios where users overlap more and feature overlap less. Its essence is the association of features of cross-users in different formats, such as power company A and local weather bureau B. Suppose A and B want to jointly train a power forecasting model, and their business systems have their own data. In addition, company A has the label data that the model needs to predict. For data privacy and security reasons, A and B cannot exchange data directly. In order to ensure the confidentiality of data in the training process, the third-party cooperator C is introduced. Here, we assume that collaborator C is honest and does not collude with either A or B. The training process is divided into two steps:

Step 1: Encrypt the sample alignment, the sample data is the class 1 power data in A and the data needed in B;

Step 2: Align samples for model encryption training:

- 1) The public key sent by third party C to A and B to encrypt the data to be transmitted;
- 2) A and B calculate the intermediate results of their respective features and encrypt the interaction to obtain the respective gradients and losses;
- 3) A and B respectively calculate the encrypted gradient and add a mask to send to C, while B calculates the encrypted loss and sends it to C;
- 4) C decrypts the gradient and loss and sends it back to A and B. A and B remove the mask and update the model.

By iterating the above steps until the federated learning loss function converges, the training process for the class 1 power data is completed, and the No. 1 power forecasting model is obtained. And so on, take the same way to the rest of the $K-1$ class power data respectively, and get the No. 2 ... K electric quantity forecast model, thus produces the electric quantity forecast multi-model.

3.3.3 C-FL Multiple Model Usage

Based on the above-mentioned training method of clustering federated learning model, a multi-model for power forecasting is generated. Multiple models are used to predict the electricity quantity, and the accurate results are obtained while the data security is guaranteed.

The joint forecasting step is divided into three steps:

- 1) The Euclidean distance between the predicted data and the K data center is calculated respectively. The Euclidean distance formula is:

$$d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (3)$$

- 2) The data to be predicted is put into model M , which is the nearest to it.
- 3) Get the predicted value from model M .

4 Experimental Evaluation

4.1 Experimental Setup

The dataset used in this experiment contains 9,568 data points collected from combined cycle power plants over six years (2006–2011). Its data labels include the local average hourly ambient temperature (T), ambient pressure (AP), relative humidity (RH), and power generation (PE). After investigation, the power output of the power plant is affected by the factors such as low and peak season, the data of power output shows the characteristics of uneven distribution and huge and complicated data. We separate the electricity factor data from the weather factor data and use the time-based tags to simulate two independent local client data for federated learning, as local client A and local client B.

The data stored by each local client is divided into training set and test set by random partition method. The accuracy of the experimental model is judged by the Mean Absolute Error (MAE) and Root Mean Square Error (RMSE).

Mean Absolute Error (MAE) is the mean of the absolute error between the predicted value and the observed value. MAE is a linear score in which all individual differences have equal weights on the average. The mathematical formula is:

$$MAE = \frac{1}{m} \sum_{i=1}^m |y_i - \tilde{y}_i| \quad (4)$$

where, y_i represents the true value, \tilde{y}_i represents the test value, m for the number of samples.

Root Mean Square Error (RMSE) is the sample standard deviation of the difference between the predicted and observed values (called residuals). The RMSE is used to indicate how discrete the sample is. The smaller the RMSE is, the better the nonlinear fitting. RMSE can better reflect the actual situation of prediction error. Its mathematical formula is:

$$RMSE = \sqrt{\frac{1}{m} \sum_{i=1}^m (y_i - \tilde{y}_i)^2} \quad (5)$$

where, y_i represents the true value, \tilde{y}_i represents the test value, m for the number of samples.

In this experiment, we compared the predictive models of two learning methods:

- 1) Traditional Federated Learning (FL): Use the training set data in A and B, the power generation prediction model is trained through the linear regression of the traditional federated learning model.
- 2) Cluster Federation Learning (C-FL): First, use the K-means clustering algorithm locally to cluster the power data of the training set in A, and then combine the power data of each

cluster with the meteorological data in B, and use the linear regression training model through federated learning mode to obtain multiple models to predict power generation.

4.2 Experimental Procedure

Pre-processing the power data in A, and the local training set power data is clustered by K-means Algorithm. The experiment uses elbow method to compare the experimental accuracy under different K values. The results are shown in Fig. 4. The core metric for determining the optimal K value using the elbow method is SSE (Sum of The Squared Errors):

$$SSE = \sum_{i=1}^k \sum_{p \in C_i} |p - m_i|^2 \quad (6)$$

where, C_i is the i^{th} cluster, p is the sample point in C_i , m_i is the center of mass of C_i (the mean of all samples in C_i).

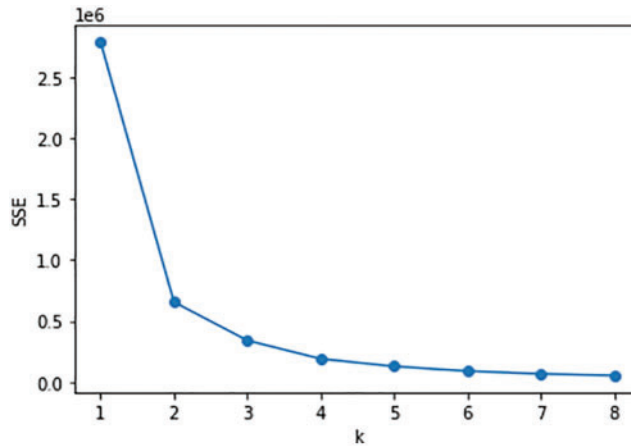


Figure 4: SSE under different K values

SSE is the clustering error of all samples, which represents the clustering effect. With the increase of clustering number K , the sample division will be more refined, the degree of aggregation of each cluster will gradually increase, so SSE will naturally gradually become smaller. When K is less than the number of true clusters, SSE will decrease greatly because the increase of K will greatly increase the degree of aggregation of each cluster, and when K reaches the number of true clusters, the return on the degree of aggregation obtained by increasing K will quickly decrease, so SSE's decline will plummet and then flatten out as the value of K continues to increase, meaning that the graph of SSE and K is an elbow, and the corresponding K value of the elbow is the real clustering number of the data. As can be seen from the graph, the K value for the elbow is 2 (with the highest curvature), so the optimal number of clusters for this data set should be 2.

Next, this paper first builds a power forecasting model through traditional federated learning combined with A and B with the training set power data without clustering, and then the power data of the training set clustered into two classes are combined with the meteorological data in B respectively, and the power prediction model is built through federated learning, and two power prediction models are obtained.

Finally, for the electric quantity prediction model without clustering federation learning, the test set data is directly put into the model to get the predicted value. For the electric quantity prediction model with clustering federation learning, the distance between the test data and the center point of two clusters is compared by Euclidean distance formula, and the data of the test set is put into the electric quantity prediction model of the center point nearest to it to get the predicted value.

4.3 Experimental Results and Analysis

The models generated by the FL method and the C-FL method are used to predict the electricity quantity separately, and the results are as follows:

The test results under the FL method are shown in Fig. 5.

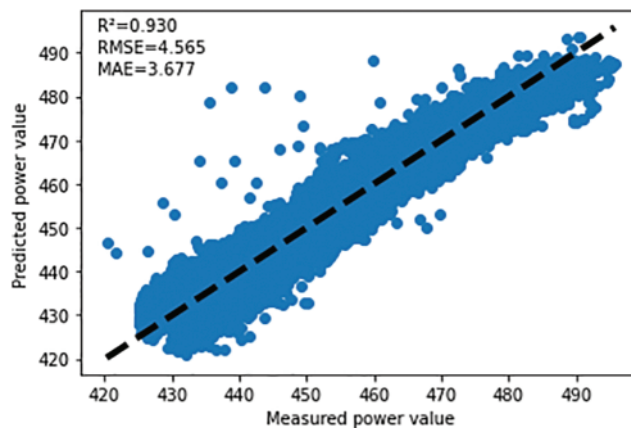


Figure 5: FL prediction results

The test results under the C-FL method are as follows, among which the prediction results of the first prediction model are shown in Fig. 6, and the prediction results of the second prediction model are shown in Fig. 7.

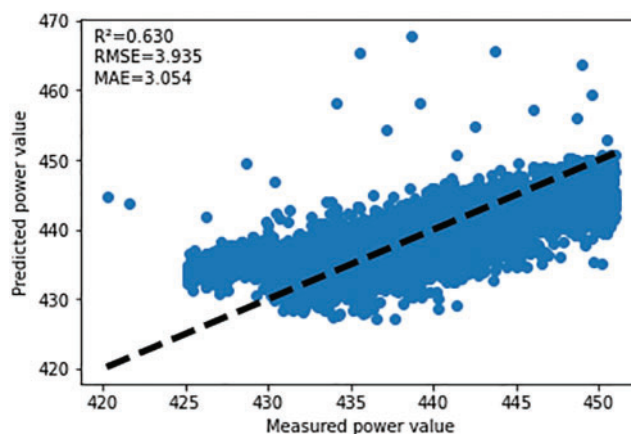


Figure 6: Prediction results of the first prediction model

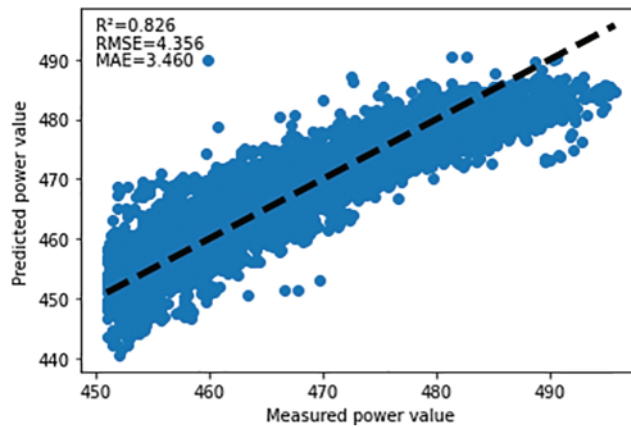


Figure 7: Prediction results of the second forecast model

As can be seen from the above three figures, because the FL method aggregates all power data, the R-Square value of its model is greater than the two models of the C-FL method, so we cannot use R-square to measure the quality of the models built by the FL and C-FL approaches. But compared with FL, the first model's MAE of the C-FL method decreased by 0.623 and the RMSE decreased by 0.63; the second model's MAE of the C-FL method compared with FL decreased by 0.217, and the RMSE decreased by 0.209. Through the analysis of the above data, we can know that although the two models generated by clustering federated learning (C-FL) only use part of the data for training, but the prediction accuracy of each model is much higher than that of training together, this may be because we cluster the disordered power data before the federated learning, and then get the multi models and the use of multiple models greatly improves the accuracy of federated learning in power forecasting. Traditional federated learning (FL) completes joint modeling without data sharing, successfully solving the thorny problem of "Isolated Data Islands", but if the data characteristics of the federated learning side are very different, if you directly perform federated learning on each local data, the accuracy of the generative model prediction may not be satisfactory. Clustering federated learning (C-FL) based on K-means also completes the training of the model under the premise of data security. Although the time of data processing before federated learning increases due to clustering, the original problem of decreasing accuracy and prediction accuracy of the federated learning model caused by excessively large differences in data features of one side is greatly improved.

5 Conclusion

When federated learning is performed on an unevenly distributed electricity data set with a large amount of data, in order to solve the problem of low accuracy and inaccurate prediction of the trained power prediction model, this paper uses the K-means algorithm to first cluster the local power data. Then the power data of each cluster is combined with multiple parties to perform federated learning to obtain a multi-model for power prediction. If you want to predict the electricity later, you can use the prediction data label to classify, and use the model that best matches it to predict, so as to achieve a better prediction effect. Research has verified that the clustering federated learning method (C-FL) adopted in this paper retains the advantages of traditional federated learning to maintain data security, and at the same time, the federated learning effect is better when faced with data with large differences in characteristics such as electricity data, the obtained electricity consumption prediction model is

more accurate. In the future, we can consider improving the clustering efficiency of the clustering algorithm in federated learning.

Acknowledgement: First of all, I would like to sincerely thank Professor Wu Jintao of Nanjing University of Information Science and Technology for his detailed guidance on this paper. Through the conversation with Professor Wu, I got a lot of useful suggestions and found the right direction. In addition, I am also very grateful to Professor Xu Zhanyang of Nanjing University of Information Science and Technology for his strong support for this work. He has provided a lot of material help for this research. Finally, I would like to thank my colleagues for their encouragement and support.

Funding Statement: The authors received no specific funding for this study.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

- [1] S. R. Khuntia, J. Rueda and M. Meijden, "Forecasting the load of electrical power systems in Mid-and long-term horizons: A review," *IET Generation Transmission & Distribution*, vol. 10, no. 16, pp. 3971–3977, 2016.
- [2] X. F. Wang, H. L. Su, T. L. Song and Q. F. Huang, "Differential user baseline load forecasting based on load subdivision," *Power Engineering Technology*, vol. 37, no. 6, pp. 33–38, 2008.
- [3] J. N. Zheng and M. Fan, "Discussion on several commonly used electricity prediction methods," *Hydropower and Pumped Storage*, vol. 15, pp. 65–67, 2015.
- [4] F. Xie, J. L. Bian, N. Wang and Q. Zheng, "Federal learning application in the field of artificial intelligence in the power internet of things," *China High and New Technology*, vol. 2019, no. 23, pp. 18–21, 2016.
- [5] R. B. Cleveland, W. S. Cleveland, J. E. Mcrae and I. Terpenning, "STL: A seasonal-trend decomposition," *Journal of Official Statistics*, vol. 6, no. 1, pp. 3–73, 1990.
- [6] X. Y. Gao, Y. Wang, Y. Gao, C. Z. Sun, W. Xiang *et al.*, "Short-term load forecasting model of GRU network based on deep learning framework," in *2018 2nd IEEE Conf. on Energy Internet and Energy System Integration (EI2)*, Beijing, China, IEEE, pp. 1–4, 2018.
- [7] T. P. Xue, "Analysis on the improved algorithm of medium and long-term load forecasting in power system," *Journal of Shanghai University of Electric Power*, vol. 31, no. 3, pp. 255–257, 2015.
- [8] J. G. Li, S. J. Zhou and H. J. Li, "Short-term electric quantity prediction method based on GRU and STL decomposition," *Journal of Shanghai University of Electric Power*, vol. 36, no. 5, pp. 415–420, 2020.
- [9] Z. Z. Shen, "Medium and long-term power load forecasting based on improved grey model," M.S. dissertation, University of Xihua, Chengdu, 2016.
- [10] D. P. Wang, "Grey forecasting model and application research of medium and long-term power load forecasting," Ph.D. dissertation, Huazhong University of Science and Technology, Wuhan, 2013.
- [11] W. Y. Liu, D. Y. Men, J. F. Liang and W. Z. Wang, "Monthly load forecasting based on grey relational degree and least squares support vector machine," *Power System Technology*, vol. 36, no. 8, pp. 228–232, 2012.
- [12] Y. S. Huang and C. Y. Jia, "Research on grey power load prediction based on particle swarm optimization and BP neural network," *Journal of State Grid Institute of Technology*, vol. 17, no. 5, pp. 6–11, 2014.
- [13] J. Yu, G. K. Yu, H. H. Guan, G. L. Gao and J. Ren, "Medium and long-term forecast of provincial electricity consumption based on improved MGM," *Sichuan Electric Power Technology*, vol. 43, no. 5, pp. 73–78, 2020.
- [14] M. L. Zhang, Z. R. Song, Y. Liang, Z. Shi and P. Ye, "Load prediction of urban prospective space based on saturated load density," *Journal of Shenyang University of Technology*, vol. 40, no. 1, pp. 12–18, 2008.
- [15] G. Zheng, B. Liu, Y. Zhou, D. Liu and G. Q. Mu, "Short-term power load prediction based on neural networks," *Journal of Xi'an University of Technology*, vol. 18, no. 2, pp. 126–130, 2002.

- [16] Y. L. Xu, Z. K. Wu, H. Y. Zhu, B. B. Wang and Z. F. Deng, "Short-term power load prediction based on multi-scale convolutional neural network," *Journal of Shenyang University of Technology*, vol. 42, no. 6, pp. 618–623, 2020.
- [17] B. X. Fang, Y. Jia, A. P. Li and R. Jiang, "Survey of big data privacy protection technology," *Big Data*, vol. 2, no. 1, pp. 1–18, 2016.
- [18] Q. Yang, Y. Liu, T. J. Chen and Y. X. Tong, "Federated machine learning: Concepts and applications," *ACM TIST*, vol. 10, no. 2, pp. 1–9, 2019.
- [19] J. Z. Wang, L. W. Kong, Z. C. Huang, L. J. Chen, Y. Liu *et al.*, "Overview of federated learning algorithms," *Big Data*, vol. 6, no. 6, pp. 64–82, 2020.
- [20] S. S. Wang, J. Y. Chen and Y. N. Lu, "COVID-19 chest CT image segmentation based on federated learning and blockchain," *Journal of Jilin University*, vol. 51, no. 6, pp. 2164–2173, 2021.
- [21] Y. Lecun, Y. Bengio and G. Hinton, "Deep learning," *Nature*, vol. 512, no. 7553, pp. 436–444, 2015.
- [22] J. Z. Wang, Z. C. Huang and J. Xiao, "Artificial intelligence enables fintech," *Big Data*, vol. 4, no. 3, pp. 114–119, 2018.
- [23] J. Konecny, H. B. McMahan, D. Ramage and P. Richtarik, "Federated optimization: Distributed machine learning for on-device intelligence," arXiv preprint arXiv:1610.02527, 2016.
- [24] J. Konecny, H. B. McMahan, F. X. Yu, A. T. Suresh and D. Bacon, "Federated learning: Strategies for improving communication efficiency," arXiv preprint arXiv:1610.05492, 2016.
- [25] H. B. McMahan, E. Moore and D. Ramage "Federated learning of deep networks using model averaging," arXiv preprint arXiv:1602.05629, 2016.
- [26] D. H. Zhai, J. Yu, F. Gao, L. Yu and F. Ding "Study on K-means text clustering algorithm for selecting initial cluster center by maximum distance method," *Application Research of Computers*, vol. 31, no. 3, pp. 714–719, 2014.