International Journal of Gaming and Computer-Mediated Simulations
Volume 16 • Issue 1

Residential Electricity Consumption Prediction Method Based on Deep Learning and Federated Learning Under Cloud Edge Collaboration Architecture

Wei Wang, State Grid Hebei Marketing Service Center, China
Xiaotian Wang, State Grid Hebei Marketing Service Center, China
Xiaotian Ma, State Grid Hebei Marketing Service Center, China*
https://orcid.org/0009-0001-1419-2141
Ruifeng Zhao, State Grid Hebei Marketing Service Center, China
Heng Yang, Beijing Tsingsoft Technology Co., Ltd., Beijing, China

ABSTRACT

Traditional residential electricity prediction methods have problems, such as difficulty in ensuring user privacy and poor convergence speed due to the influence of Different Residential Electricity Consumption (REC) habits. A REC prediction method based on Deep Learning (D-L) and Fed-L under the Cloud Edge Collaboration (CEC) architecture is proposed to address the above issues. First, based on the CEC architecture, combining edge computing and cloud computing center server, the overall model of REC prediction is built. Then, Federated Learning (Fed-L) and D-L model Empirical Mode Decomposition - Long Short-Term Memory (EMD-LSTM) were introduced on the edge side, and the edge side Fed-L depth model was personalized by using EMD-LSTM. Finally, aggregation of edge side models was achieved in the cloud by receiving encrypted model parameters from the edge side and updating and optimizing all edge side models. The results show that the proposed method has the highest REC prediction accuracy, reaching 96.56%, and its performance is superior to the other three comparative algorithms.

KEYWORDS

Cloud Edge Collaboration, Federated Learning, EMD, LSTM, Residential Electricity Consumption Prediction, Simulation

DOI: 10.4018/IJGCMS.336846 *Corresponding Author

1. INTRODUCTION

The reliability and economy of power system operation are important guarantees for the healthy development of both society and the economy. Before a significant breakthrough in energy storage technology, the electricity generation and consumption are roughly equal (Hong et al., 2020; Wang et al., 2022; Chodakowska et al., 2021). Therefore, the planning and operation of the power system, as well as the daily scheduling work of the power system, require REC prediction as the decision-making basis, which has important research significance (Rendon-Sanchez & De Menezes, 2022).

With the continuous development of new energy grids, the combination of new energy and REC has become increasingly close. REC prediction has higher value, higher difficulty, and new technological foundation in the new era. When the power grid consumes intermittent, random, and fluctuating new energy generation, it needs to have a certain reserve capacity. However, when the grid's capacity is insufficient, serious waste of resources, such as wind and solar power, may occur (Phyo & Jeenanunta, 2021; Barman & Dev Choudhury, 2022; Tran et al., 2021). With the further development of electricity market reform, the number of users in the electricity market is gradually increasing. The power market adopts the market mechanism to improve the efficiency of various users in the power sector. Accurate and reliable REC prediction is an important basis for the business decisions of users, such as power companies and retailers. Therefore, as the number of market participants increases, the number of stakeholders who need to make decisions based on REC predictions gradually increases (Li et al., 2022; Mustaqeem & Kwon, 2022). In addition, with the further opening of the electricity market, the impact of REC's predictive power on the profitability of relevant power companies and retailers is gradually increasing. Therefore, REC prediction has a more important position in the new market environment (Ahmed et al., 2020; Huang et al., 2021; Basiri et al., 2021). Moreover, the accuracy of REC prediction results will directly affect the quality of urban power grid planning work (Shahid, Zameer, Muneeb et al., 2021; Wang et al., 2020; Kong et al., 2019). Due to the complex and diverse factors that affect the prediction of residential electricity REC, simply considering the characteristics of the energy consumption system itself is completely insufficient in the energy Internet mode. On the basis of in-depth analysis of electrical characteristics, taking into account social, meteorological, economic, and other influencing factors is necessary to ensure the accuracy and stability of prediction accuracy (Chen et al., 2022; Sheng et al., 2021).

In addition, accurate REC prediction can ensure better trading in the dynamic electricity trading market and is also an important indicator for future planning of the power system. Accurately providing predicted load data for the corresponding period plays a crucial role in generating and selling electricity, formulating reasonable electricity prices, and ensuring the normal production and life of residents. However, the level of REC is influenced by local economic development level, human activities, REC characteristics, and climate conditions, which undoubtedly increase the difficulty of predicting electricity REC (Bento et al., 2021; Gao et al., 2022).

At present, scholars have conducted more research on precise load prediction algorithms for total household electricity demand, but there has not been in-depth research on precise interaction of terminal devices based on residential demand response, multidimensional evaluation models, interaction strategies under multiple factors, identification and simulation of key loads for residents, and other related contents. This cannot provide effective scientific research or theoretical support for the construction of residential energy Internet (Alfieri & De Falco, 2020; Xie et al., 2020; Shahapure & Nicholas, 2020). Due to the lack of data resources for the REC situation and the integration with key data in the substation area, both the depth and breadth of research and the support for practical applications need to be strengthened. It is far from meeting the application requirements to support the interaction between the power grid and the residential intelligent REC, nor can it support the full mode simulation of the source network load storage under the large power grid. Therefore, based on accurate REC prediction research, it can quickly make up for the current shortcomings and provide good theoretical support for residential intelligent REC, which is of great significance in ensuring the smooth operation of the power system and economic benefits (El-Hendawi & Wang, 2020; Zhang & Li, 2021).

This article proposes a fast and efficient prediction method with good user privacy performance for residential electricity consumption prediction. The main research content is as follows:

(1) Section 2 analyzes existing research and literature to clarify the advantages and disadvantages of different methods.
(2) Section 3 constructs a residential electricity consumption prediction model based on cloud edge collaborative architecture. Detailed introductions are given to the sub modules of the model, including the overall model architecture, personalized federated learning based on deep learning, EMD-LSTM model, federated learning deep model, and the specific process of residential electricity consumption prediction.
(3) The performance of the proposed method is validated through simulation experiments in Section 4, including training the model to select better model parameters and comparative analysis with other models.

2. RELATED RESEARCH

The further reform of the electricity market, the increasing randomness of REC behavior, and the stable implementation of smart grids and ubiquitous IoT have respectively brought more urgent needs, higher challenges, and new technological foundations to the research of electricity REC prediction. Based on the rich historical data provided by the smart grid, researching REC prediction methods with higher prediction accuracy and stronger explanatory performance is of great significance for improving the operational and commercial decision-making efficiency of various users.

A REC prediction method with dynamic mirror descent function was constructed based on the DL framework (Han & Wang, 2023), which can be adaptively applied to individual REC prediction. However, this method has significant differences in prediction accuracy for different seasons. Shabbir (2021) conducted comparative analysis on different types of REC prediction models based on the same dataset and ultimately selected the grassroots model with better performance. However, this method does not provide a novel resident fit prediction method. A single REC prediction model was constructed by combining k-nearest neighbor algorithm and feature selection (Forootani & Rastegar & Sami, 2022). However, this method did not fully consider the factors that affect load changes. Gonzalez (2023) proposed a deep network model based on DNN that can be used for adaptive REC prediction. However, as the number of residents changes, the error will also continue to increase. Park (2023) proposed a REC prediction model that can aggregate remote training models by introducing FL. However, the model cannot truly reflect the actual load changes of different users. Hou (2021) proposed an adaptive REC prediction method by identifying the clustering structure to aggregate the regularity of REC. However, the prediction error of this method is relatively high. Wu (2022) proposed a transfer learning based REC prediction method using a boosting framework to transfer and select multiple transfer models. However, when the power load exhibits irregular fluctuations, the REC prediction accuracy of this method will decrease.

The structure of historical residential electricity data is very complex, and the characteristics of residential electricity itself determine that these data have highly variable characteristics. Most existing methods are based on deep learning to construct residential electricity consumption prediction models, which may not be able to freely use historical data of residential electricity consumption during training, making it difficult to make effective predictions in the process of residential electricity consumption prediction. FL is a distributed machine learning with privacy protection function, and each user's original private dataset is stored on the edge side. A REC prediction method based on DL and FL under the CEC architecture is proposed to address the issues of difficult user privacy assurance, poor convergence speed, and prediction accuracy in traditional REC prediction methods. Compared with traditional methods, there are the following innovative points:

(1) CEC architecture is built by combining edge computing and cloud computing, and REC prediction model is built on this basis.
(2) By introducing FL for model training on the edge side and then collecting, the privacy of user data during transmission is ensured.
(3) A FL edge side model for REC prediction was constructed based on EMD-LSTM, which personalized the FL and reduced the error of REC prediction.

3. BUILDING A REC PREDICTION MODEL BASED ON CEC ARCHITECTURE

3.1 Overall Model Architecture

According to data predictions from relevant scholars, data in various fields around the world will show rapid growth in the future. Therefore, edge computing, which breaks the traditional centralized data processing, has emerged. Edge computing is located between the terminal side and the cloud data center, which is closer to the data source geographically. Here, we consider using the CEC architecture to construct a REC prediction model.

The CEC system model mainly consists of user terminal side, multiple edge nodes, and cloud side. The user's REC data is collected and transmitted to the edge node through a data collection device. The edge node processes or transmits the deployed service according to the service deployment requirements. The edge node and cloud computing center are configured with computing and storage resources that match their computing capabilities. Among them, cloud computing centers are server clusters with powerful storage and computing capabilities. However, due to the distance between cloud computing centers and user terminals, it will generate higher latency and network communication costs. The computing resources and hardware devices of edge nodes are limited, but their distribution is closer to the user terminal side. In the CEC environment, cloud computing centers are mainly responsible for providing services from the edge side. Edge computing is mainly to process real-time data and provide the cloud computing center with the services it needs.

The overall structure of the REC prediction model based on CEC is as follows.

In Fig. 1, the CEC architecture is applied to REC prediction, and FL and DL are also applied in the prediction process. The model includes three parts:

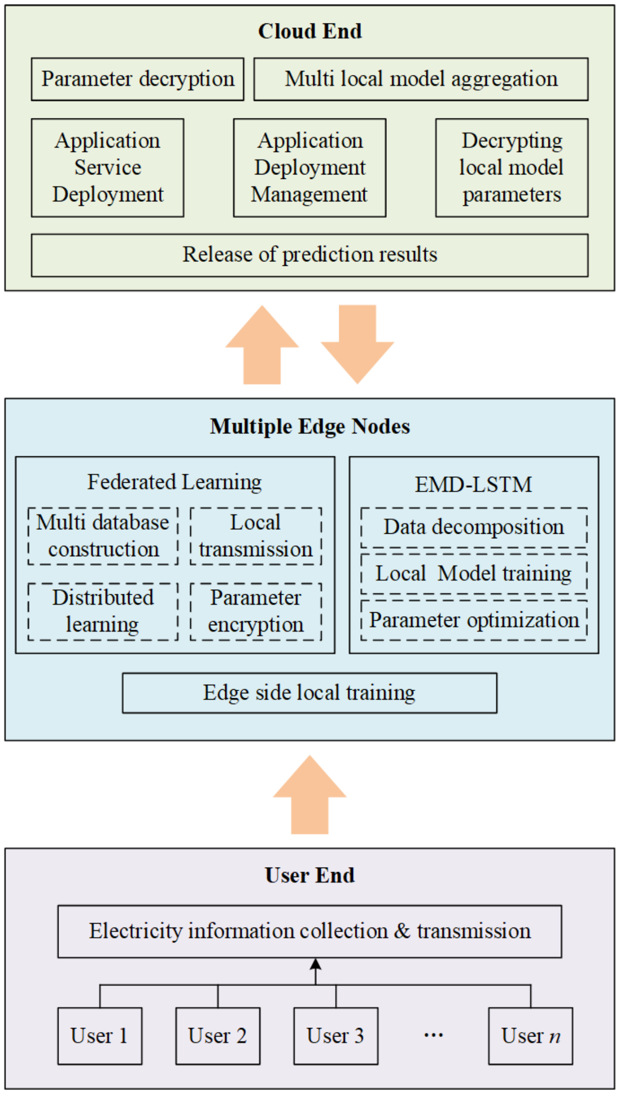
(1) User side. At the user end, service data can be collected through power information collection and transmission devices, and the collected data can be sent to the nearest edge node based on the user's location.
(2) Edge nodes. There are many edge nodes, depending on the size of the predicted REC range. Each edge node conducts edge side training based on FL and DL models after receiving user data. The trained model parameters will be encrypted and sent to the cloud.
(3) In the cloud. The cloud will receive the trained model parameters from all edge nodes, and aggregate all edge side models based on these parameters. Through continuous updates and optimization, the aggregated model is formed and used for online prediction of REC.

3.2 Personalized FL Based on DL

3.2.1 FL

FL is an emerging machine learning algorithm framework originally used to solve the problem of Android mobile end users updating language prediction models on the edge side. The design goal of FL is to ensure privacy and data security during the big data exchange process under the premise of

Figure 1. REC prediction model based on CEC

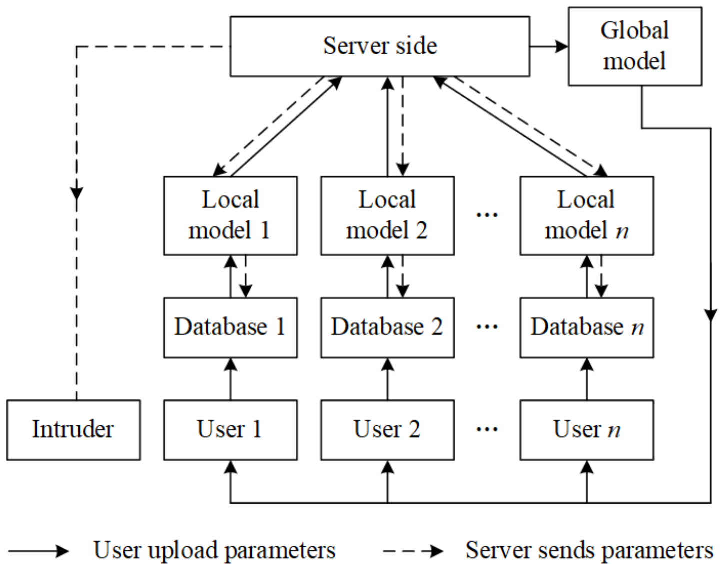


legal compliance, so it is designed in a form that does not require direct data exchange or collection. It not only protects user privacy but also solves the data island dilemma faced by the artificial intelligence industry.

FL aims to establish a FL model based on distributed datasets. Under the FL framework, there should be no less than two client nodes that will jointly train a global model with the cooperation of a central server, and during the global model training process, it is required that the original data does not go out of the edge. The working principle of FL is as follows.

In Fig. 2, the shared global model is trained by a federation composed of participating devices, which is hosted by a central server. This method enables edge nodes to collaboratively train models on the edge side without the need to share original training data. Here, the edge side model parameter update will replace the training data. The central server will aggregate these parameter updates and generate a new round of global models, repeating this process until the global model converges.

Figure 2. Operational principle of FL



The FL system consists of one server and K users. Let d_k represent the database of edge side user U_k, $k = 1,2,3,\ldots,K$. The goal of the server side is to learn the models constructed by each user based on their respective datasets. The client participating in edge side training needs to find a model parameter α to minimize the center loss function, as shown below:

$$\alpha = \sum_{k=1}^{K} p_k \alpha_k \tag{1}$$

In eq(1), α_k is the parameters trained by the k-th user, and α represents the parameters after server aggregation. The meaning of p_k is as follows:

$$p_k = \frac{|d_k|}{\sum_{k=1}^{K} |d_k|} \geq 0 \tag{2}$$

In eq(2), $\sum_{k=1}^{K} |d_k|$ represents the total number of data samples. Therefore, the optimization problem of FL can be represented as follows:

$$\alpha' = \underset{\alpha}{arg} \min \sum_{k=1}^{K} p_k L_k \left(\alpha, d_k\right) \tag{3}$$

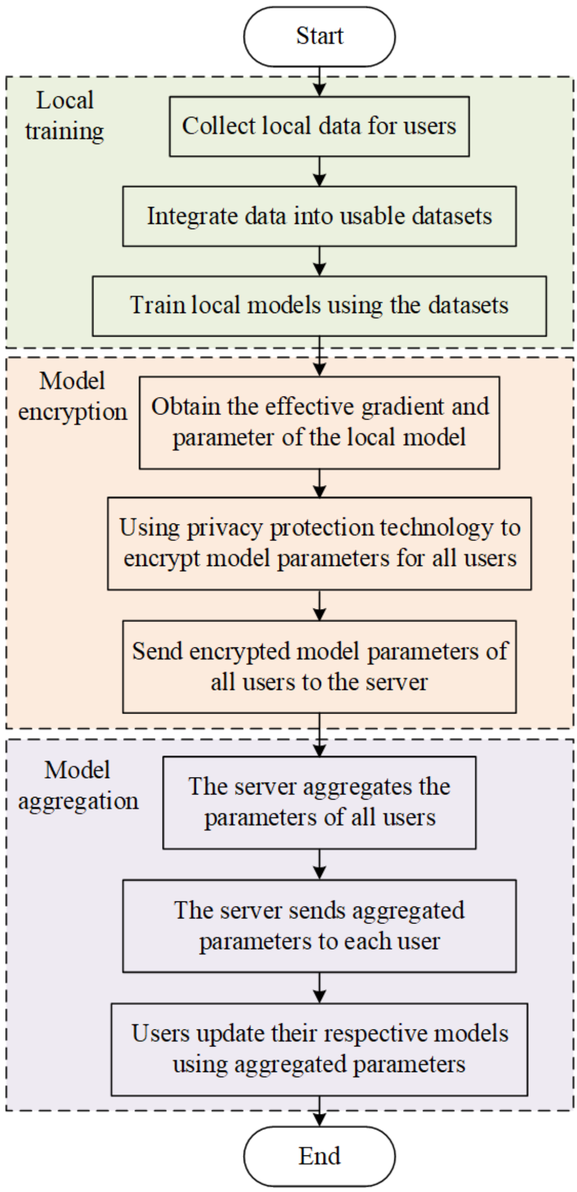
In eq(3), $L_k\left(\alpha, d_k\right)$ is the edge loss function of the k-th user.

The training process of a FL system is shown in Fig. 3.

3.2.2 EMD-LSTM Model

The prediction of REC data relies on scientific prediction theories and reasonable prediction models. There are two main directions, one is the mechanism driven method, which relies on analyzing the characteristics of REC data through physical and mechanism information. However, this method cannot fully mine historical data and has poor accuracy. The second is the model driven approach, which utilizes machine learning models, probability models, or time series models to mine historical data for training and apply it to unknown data. As the difficulty of data collection decreases and

Figure 3. The training process of FL system



computing power becomes more efficient, model driven methods are more suitable for the current transformer warning system.

LSTM is a RNN network oriented towards temporal data. Targeting structured and unstructured data, corresponding LSTM combination prediction models can be established for data types. In addition, the REC data series has nonlinear and non-stationary characteristics, and using EMD to decompose the series can effectively enhance the availability of data. EMD is an adaptive signal decomposition method that is particularly suitable. After decomposition, each sequence component is predicted, and then the prediction results of each sequence component are reconstructed to obtain a combined prediction result that meets the accuracy requirements.

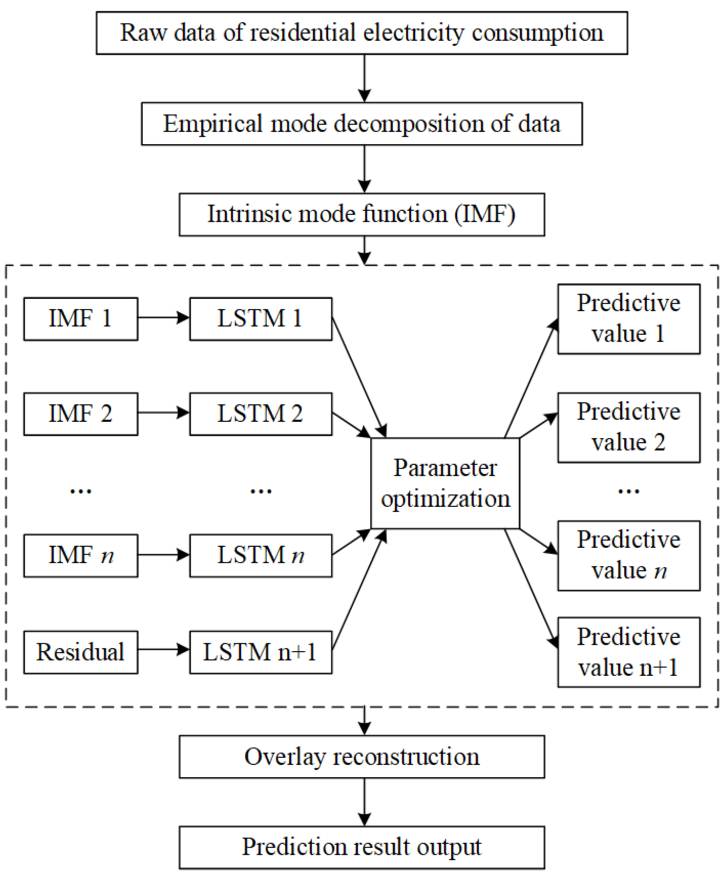
Based on LSTM and EMD, an edge side combination model EMD-LSTM for FL can be constructed, as shown in Fig. 4.

In Figure 4, IMF represents the natural mode function, which can be used to analyze the frequency spectrum, solve the vibration behavior of the system under a single mode, analyze the modal matching problem of the balance system, and study the process of object vibration and its attenuation. Here it is mainly used for adaptive signal decomposition. n represents the number of the LSTM model.

3.2.3 FL Deep Model

REC data belongs to non-independent identically distributed data, and traditional FL creates global models by aggregating these different model updates. However, this model can only obtain common

Figure 4. Edge side training model EMD-LSTM



features of customers participating in training, so it may perform poorly on specific user data. To address these issues, consider personalizing FL in specific issues and providing personalized models for each customer in the federation.

By combining EMD-LSTM with FL, an EMD-LSTM-FL model is constructed, which enables customers to train independent models using edge side data. Each client first trains the model on a common dataset to execute EMD-LSTM and then fine tunes the model based on edge side data before the FL training phase. This is equivalent to designing a structure of "basic layer + personalized layer." The basic layer shares data with the central server, while the personalization layer maintains privacy on the client side for edge side training.

Train edge side models of different users using EMD-LSTM in the EMD-LSTM-FL model. EMD-LSTM utilizes the ReLU function for activation. When aggregating models, weight ratios are used for weighted integration to obtain the aggregated model, as shown below:

$$\alpha_{FedL} = \omega\alpha + \left(1-\omega\right)\alpha_0 \tag{4}$$

In eq(4), α_{FedL} represents the parameters of the aggregation model. α and α_0 represent the model parameters distributed at different edge nodes, respectively. ω represents the weight ratio of different edge node model parameters. This method converges faster than traditional neural networks and has better convergence performance during parameter joint initialization.

3.4 REC Prediction Process

The REC prediction process based on DL and FL under the CEC architecture is as follows:

Step 1: Obtain historical raw data based on REC information on the user end and send this data to the nearest edge node.
Step 2: Build a model framework and parameter encryption transmission mechanism for edge side training based on FL.
Step 3: Select feature inputs based on REC data and decompose the historical raw data sequence of REC using EMD. Divide the modal decomposed data into training and testing sets.
Step 4: Apply the DL model LSTM to the edge side model in FL to form the FL depth model.
Step 5: Train the edge side model of each edge node using the training set and verify the trained model using the test set to determine whether the model meets the accuracy requirements. If the requirements are met, proceed to the next step. Conversely, readjust the model hyperparameters and repeat the training process.
Step 6: Obtain the available model parameters for each edge side model through edge side training.
Step 7: Encrypt all edge side model parameters and send the encrypted parameters to the central server in the cloud.
Step 8: The central server builds an aggregation model by continuously optimizing and updating the parameters of all user edge side models.
Step 9: Use an aggregation model for online prediction of REC.

The specific prediction process is shown in the following.

4. SIMULATION ANALYSIS

4.1 Experimental Setup

The experimental hardware settings are shown in Table 1.

Figure 5. Typical architecture of FL

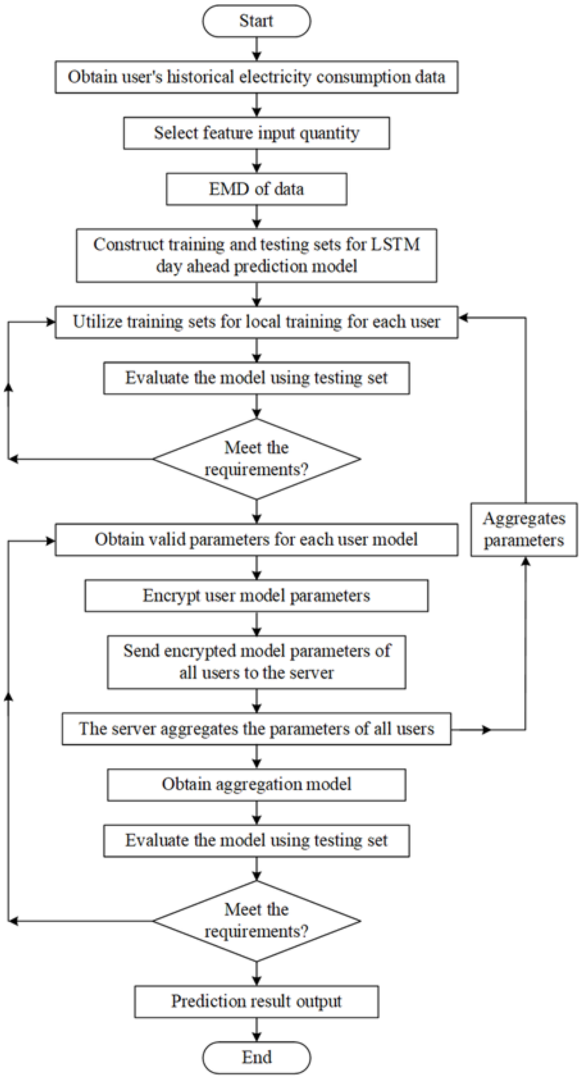


Table 1. Hardware configuration for the experiment

Hardware	Configuration
OS	Win 10
CPU	Intel Core i5
GPU	NVIDIA GTX 1660
Running memory	32 GB

Choosing the appropriate number of hidden layers is an important consideration in the design and training of deep neural networks based on the specific tasks and characteristics of the dataset. Generally speaking, using at least two hidden layers can effectively enhance the model's expressive power and feature learning effectiveness. In this experiment, we determined the optimal number of hidden layers to be 5 through experimental adjustments.

In order to adapt to unstable objective functions and maximize computational efficiency while reducing memory usage, Adam was selected as the optimizer.

The Adam optimizer can adaptively adjust the learning rate based on historical gradient information. In the early stages of training, the Adam optimizer uses a higher learning rate and can quickly converge. In the later stage of training, using a smaller learning rate can more accurately find the minimum value of the loss function. Therefore, we set the initial learning rate to 0.1 for the Adam optimizer to make adaptive adjustments.

The configuration of the experimental software is as follows.

4.2 Data Set

The dataset used in the experiment was actual REC data provided by a power supply company in Zhejiang Province, China. This dataset includes hourly electricity usage data, house feature data, and weather data for 132 users in a certain substation area over the past three years. The specific situation of the data is as follows:

(1) Energy usage data is sampled in hours. In this way, 24 samples can be taken per day, resulting in 24 sampling values. Hourly energy usage data is recorded in kilowatt hours.
(2) The housing feature data includes various building information, such as the type of housing feature (old-fashioned single-story residential buildings, modern high-rise apartments, etc.), housing orientation, surrounding rent situation (to determine consumption level), and HVAC type (gas fireplace, electric heating, fixed air conditioning, etc.).
(3) Weather data is collected from the nearest weather station to the residence, including hourly outdoor temperature, outdoor relative humidity, and a brief textual description of the day's weather (such as cloudy or sunny days).

The visualization display of REC data for a certain resident is shown below.

4.3 Evaluating Indicator

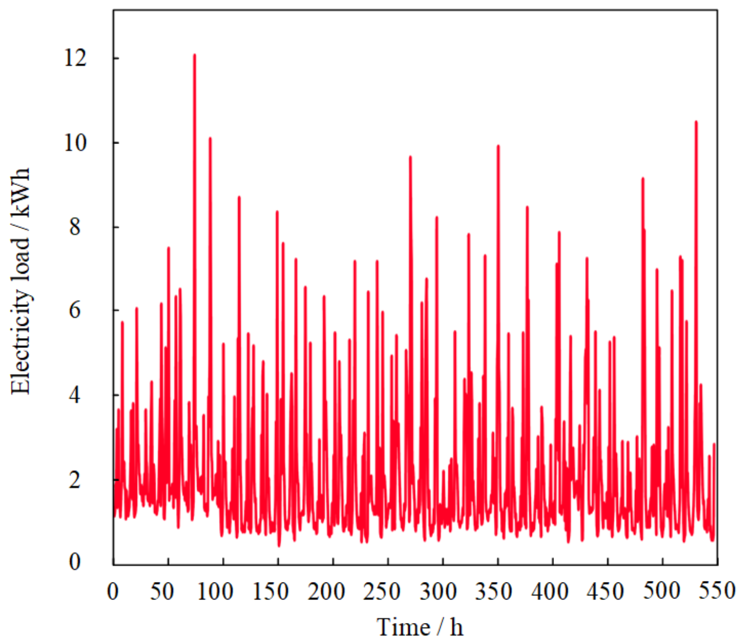
Select the average absolute error E_{MAE} and root mean square error E_{RMSE} as evaluation indicators in the REC prediction experiment. E_{MAE} and E_{RMSE} respectively reflect the error between the predicted and measured values, so the smaller the two, the better the prediction effect; that is, the higher the accuracy of the prediction. Their calculation method is as follows:

$$E_{MAE} = \frac{1}{N}\sum_{n=1}^{N}\left|e'\left(t\right) - e\left(t\right)\right| \tag{5}$$

Table 2. Software and parameter configuration for the experiment

Parameter	Value
Programming Language	Python3.7
Deep learning framework	PyTorch
Number of hidden layers	5
Optimizer	Adam
Learning rate	0.01

Figure 6. REC data of a certain user



$$E_{RMSE} = \sqrt{\frac{1}{N}\sum_{n=1}^{N}\left[e'\left(t\right) - e\left(t\right)\right]^2} \tag{6}$$

In eq(5) and (6), $e\left(t\right)$ is the measured value of photovoltaic output, $e'\left(t\right)$ is the predicted value of photovoltaic output before the day at time *t*, and *N* is the total number of time series points synthesized by all samples in the test set.

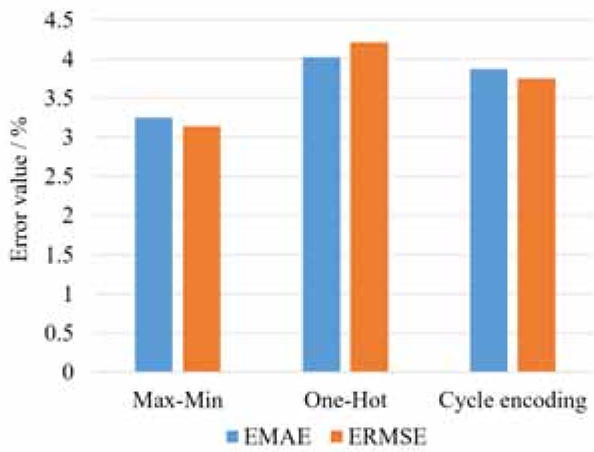
4.4 Model Training

After selecting appropriate data, it is necessary to preprocess the data before conducting training. The REC prediction model based on DL and FL requires real and complete load data input to learn the internal laws of load changes.

The following data preprocessing methods are used for the original time features, including maximum minimum normalization, One Hot encoding, and period encoding, to analyze the impact of feature preprocessing methods on prediction accuracy. This experiment predicts the load within a certain month in the HUE dataset, while the remaining data is used as the training set. E_{MAE} and E_{RMSE} are as follows.

In Fig. 7, the model with max min normalization of the original data features has the best prediction performance, followed by periodic encoding, and the error indicator with One Hot encoding is the worst. Therefore, in the following experiment, max min normalization was used to preprocess the raw data. If data of different dimensions are trained directly in the model without processing, it may affect the training speed and prediction accuracy. Therefore, it is necessary to normalize the data; that is, scale the dataset proportionally to the range of [0,1], which can effectively remove the influence of dimensionality, transform dimensional values into dimensionless pure values, and solve the problem of significant numerical differences between different features. At the same time, it can

Figure 7. The relationship between EMAE and load similarity

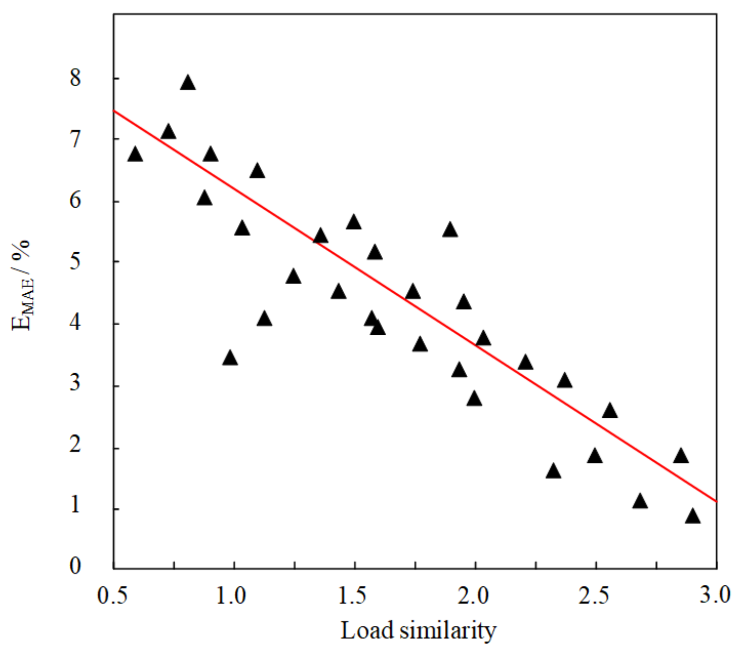


also reduce data noise, improve training speed, and prevent overfitting. The following is an analysis of the relationship between prediction error and load similarity. Load similarity refers to the degree of similarity between the input sample and the predicted result.

The variation of E_{MAE} with load similarity in the experiment is as follows.

In Fig. 8, there is a negative correlation between E_{MAE} and the load similarity of the predicted day. When the load similarity between similar days and predicted days increases, the corresponding prediction error will decrease. That is to say, the selection of similar days has a positive impact on building REC prediction.

Figure 8. The relationship between EMAE and load similarity



4.5 Comparative Analysis

In order to further verify the superiority of the REC prediction methods based on DL and FL under the proposed CEC architecture, comparative experiments will be conducted with the DMD-DL method in [25], the ALA-LSTM method in [30], and the DBTR method in [31], respectively. Among them, [25] is a study conducted on the volatility of electricity consumption among different individual residents. This method can be adaptively applied to individual residential electricity consumption prediction, and is representative in predicting individual residential electricity consumption. [30] achieves aggregation of residential electricity consumption patterns by identifying clustering structures, which can represent methods based on clustering structures. [31] considers the impact of the continuous popularization of electric vehicles on the composition of household electricity consumption data, which is representative in the application of new energy.

The number of model parameters and training time for different methods are as follows when conducting experiments using the same dataset and under the same experimental conditions.

The data in Table 3 indicates that the proposed method can achieve convergence at the fastest speed.

The REC prediction results of different methods are as follows.

The prediction accuracy obtained by using different methods for REC prediction is shown in Fig. 10.

In Fig. 9 and 10, the REC of proposed method are closest to the real curve and the final REC prediction accuracy is the highest, reaching 96.56%.

The following experimental analysis is conducted on the impact of different methods on the prediction time scale in the REC prediction process.

Table 3. Number of model parameters and training time for different methods

	No. of parameters	Single iteration time	Training time
Proposed method	5867	3.2 s	8.3 min
Ref. [25]	8543	4.7 s	10.2 min
Ref. [30]	7492	5.5 s	12.73 min
Ref. [31]	10538	6.8 s	15.43 min

Figure 9. Result of residential electricity consumption prediction using different methods

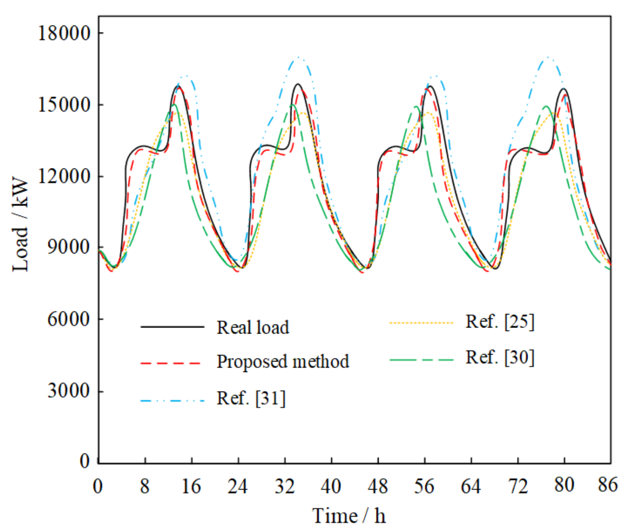
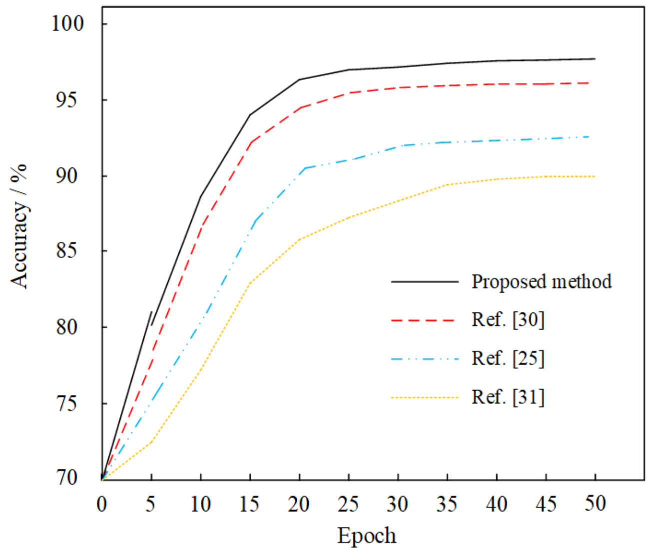


Figure 10. Accuracy of REC prediction using different methods



The curve in Fig. 11 verified that the proposed method still has better predictive ability over a longer time scale.

Using different methods to predict the REC situation, the overall E_{MAE} and E_{RMSE} comparison results are shown in Fig. 12, 13, and Table 4, respectively.

The above experimental results provide the REC prediction results of four methods under the same experimental conditions using the same dataset. The results indicate that the experimental results of the proposed method are superior to other methods.

Figure 11. The degree to which different methods are affected by the predicted time scale

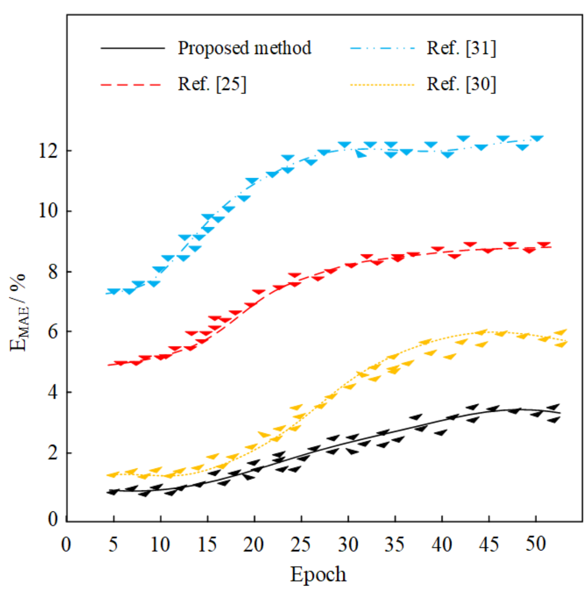


Figure 12. EMAE for REC prediction using different methods

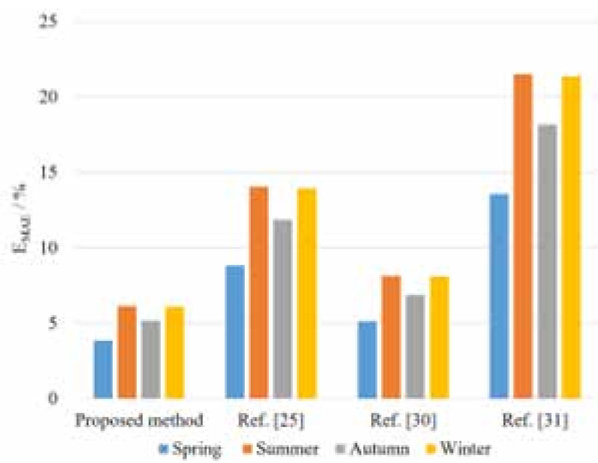


Figure 13. ERMSE for REC prediction using different methods

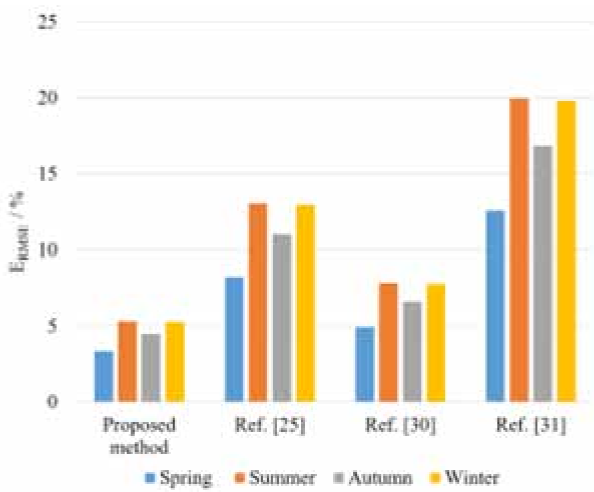


Table 4. Comparison of experimental results for different methods

	E_{MAE}	E_{RMSE}
Proposed method	3.54%	3.48%
Ref. [25]	8.45%	8.23%
Ref. [30]	12.49%	12.18%
Ref. [31]	4.57%	4.72%

4.6. Discussion

We will discuss the above experimental results here:

(1) In Table 3, the proposed method has the least number of parameters compared to the other three comparative methods. Therefore, its single iteration time and model training time are both minimized, and it can complete convergence at the fastest speed.
(2) In Fig. 9 and 10, the REC of proposed method are closest to the real curve, and they tend to converge after the 20th iteration. Compared with the other three comparison methods, the proposed method converges the fastest and the final REC prediction accuracy is the highest, reaching 96.56%.
(3) In Fig. 11, the accuracy of REC prediction decreases with increase of time scale. This is because during the rolling prediction process, some predicted values are used as inputs to the neural network, causing errors to gradually accumulate and increase. In the process of predicting changes in the time scale, the proposed method consistently outperforms the other three comparative methods in terms of prediction accuracy, with the highest indicator E_{MAE} not exceeding 4%. Verified that the proposed method still has better predictive ability over a longer time scale.
(4) The above experimental results provide the REC prediction results of four methods under the same experimental conditions using the same dataset. The REC prediction of the proposed method has the smallest E_{MAE} and E_{RMSE}, divided into 3.54% and 3.48%. The E_{MAE} of the proposed method decreased by 4.91%, 8.95%, and 1.03% compared to the other three comparative methods, while E_{RMSE} decreased by 4.75%, 8.70%, and 1.24%, respectively. This is due to the introduction of EMD-LSTM in the traditional FL architecture for constructing edge side training models in FL. By personalizing FL, it improves its poor performance on specific user data, and overall reduces and improves the accuracy of REC prediction.

5. CONCLUSION

A REC prediction method based on DL and FL under the CEC architecture is proposed to address the issues of difficult user privacy assurance, poor convergence speed, and prediction accuracy in traditional REC prediction methods. The experimental verification results show that the CEC architecture can break the traditional centralized data processing process and achieve rapid processing of massive data by providing nearest services. Building a REC prediction model based on FL does not require direct data exchange or collection, and can effectively safeguard data privacy without affecting data exchange and sharing. By constructing an edge side training model in FL based on EMD-LSTM to personalize FL, the problem of poor performance of traditional FL on specific user data can be improved, greatly reducing the error of REC prediction. At the same time, we need to realize that the proposed method did not conduct relevant analysis on the impact of different performance terminal devices on prediction results in actual scenarios. Therefore, the next phase of research will focus on studying the problem of different performance terminal devices being unable to synchronously upload their edge side model parameters in practical scenarios by introducing the idea of asynchronous FL.

CONFLICT OF INTEREST

The authors of this publication declare there are no competing interests.

ACKNOWLEDGMENT

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors. Funding for this research was covered by the author(s) of the article.

REFERENCES

Ahmed, R., Sreeram, V., Mishra, Y., & Arif, M. D. (2020). A review and evaluation of the state-of-the-art in PV solar power forecasting: Techniques and optimization. *Renewable & Sustainable Energy Reviews*, *12*(6), 124–133. doi:10.1016/j.rser.2020.109792

Alfieri, L., & De Falco, P. (2020). Wavelet-Based decompositions in probabilistic load forecasting. *IEEE Transactions on Smart Grid*, *11*(2), 1367–1376. doi:10.1109/TSG.2019.2937072

Barman, M., & Dev Choudhury, N. B. (2022). Season specific approach for short-term load forecasting based on hybrid FA-SVM and similarity concept. *Energy*, *174*(4), 886–896.

Basiri, M. E., Nemati, S., Abdar, M., Cambria, E., & Acharya, U. R. (2021). ABCDM: An Attention-based Bidirectional CNN-RNN Deep Model for sentiment analysis. *Future Generation Computer Systems*, *115*(5), 279–294. doi:10.1016/j.future.2020.08.005

Bento, P. M. R., Pombo, J. A. N., Calado, M. R. A., & Mariano, S. J. P. S. (2021). Stacking ensemble methodology using deep learning and ARIMA models for short-term load forecasting. *Energies*, *14*(21), 21–30. doi:10.3390/en14217378

Chen, K. J., Chen, K. L., Wang, Q., He, Z., Hu, J., & He, J. (2022). Short-Term load forecasting with deep residual networks. *IEEE Transactions on Smart Grid*, *10*(4), 3943–3952. doi:10.1109/TSG.2018.2844307

Chodakowska, E., Nazarko, J., & Nazarko, L. (2021). ARIMA models in electrical load forecasting and their robustness to noise. *Energies*, *14*(23), 22–30. doi:10.3390/en14237952

El-Hendawi, M., & Wang, Z. L. (2020). An ensemble method of full wavelet packet transform and neural network for short term electrical load forecasting. *Electric Power Systems Research*, *182*(3), 192–200. doi:10.1016/j.epsr.2020.106265

Forootani, A., Rastegar, M., & Sami, A. (2022). Short-term individual residential load forecasting using an enhanced machine learning-based approach based on a feature engineering framework: A comparative study with deep learning methods. *Electric Power Systems Research*, *210*(3), 218–226. doi:10.1016/j.epsr.2022.108119

Gao, X., Li, X. B., & Zhao, B. (2022). Short-Term electricity load forecasting model based on EMD-GRU with feature selection. *Energies*, *12*(6), 18–27.

Gonzalez, R., Ahmed, S., & Alamaniotis, M. (2023). Implementing very-short-term forecasting of residential load demand using a deep neural network architecture. *Energies*, *16*(9), 63–72. doi:10.3390/en16093636

Han, F. J., & Wang, X. H. (2023). Adaptive individual residential load forecasting based on deep learning and dynamic mirror descent. *Frontiers in Energy Research*, *10*(5), 58–67. doi:10.3389/fenrg.2022.986146

Hong, T., Pinson, P., Wang, Y., Weron, R., Yang, D., & Zareipour, H. (2020). Energy forecasting: A review and outlook. *IEEE Open Access Journal of Power and Energy*, *7*(3), 376–388. doi:10.1109/OAJPE.2020.3029979

Hou, T. T., Fang, R. C., Tang, J. R., Ge, G., Yang, D., Liu, J., & Zhang, W. (2021). A novel short-term residential electric load forecasting method based on adaptive load aggregation and deep learning algorithms. *Energies*, *14*(22), 135–143. doi:10.3390/en14227820

Huang, C. J., Shen, Y. M., Chen, Y. H., & Chen, H.-C. (2021). A novel hybrid deep neural network model for short-term electricity price forecasting. *International Journal of Energy Research*, *45*(2), 2511–2532. doi:10.1002/er.5945

Kong, W., Dong, Z. Y., Jia, Y., Hill, D. J., Xu, Y., & Zhang, Y. (2019). Short-Term residential load forecasting based on LSTM recurrent neural network. *IEEE Transactions on Smart Grid*, *10*(1), 841–851. doi:10.1109/TSG.2017.2753802

Li, Y., Zeng, J. B., Shan, S. G., & Chen, X. (2022). Occlusion aware facial expression recognition using CNN with attention mechanism. *IEEE Transactions on Image Processing*, *28*(5), 2439–2450. doi:10.1109/TIP.2018.2886767 PMID:30571627

Mustaqeem, K., & Kwon, S. (2021). MLT-DNet: Speech emotion recognition using 1D dilated CNN based on multi-learning trick approach. *Expert Systems with Applications*, *167*(3), 12–20. doi:10.1016/j.eswa.2020.114177

Park, K. J., & Son, S. Y. (2023). Residential load forecasting using modified federated learning algorithm. *IEEE Access : Practical Innovations, Open Solutions*, *11*(4), 40675–40691. doi:10.1109/ACCESS.2023.3268530

Phyo, P. P., & Jeenanunta, C. (2021). Daily load forecasting based on a combination of classification and regression tree and deep belief network. *IEEE Access : Practical Innovations, Open Solutions*, *9*(2), 152226–152242. doi:10.1109/ACCESS.2021.3127211

Rendon-Sanchez, J. F., & De Menezes, L. M. (2022). Structural combination of seasonal exponential smoothing forecasts applied to load forecasting. *European Journal of Operational Research*, *275*(3), 916–924. doi:10.1016/j.ejor.2018.12.013

Shabbir, N., Kutt, L., & Raja, H. A. (2021). *Machine learning and deep learning techniques for residential load forecasting: A comparative analysis.* 2021 IEEE 62nd International Scientific Conference of the Power-and-Electrical-Engineering of Riga-Technical-University, Location Riga, Latvia.

Shahapure, K. R., & Nicholas, C. (2020). *Cluster quality analysis using silhouette score*. Proceedings of the 7th IEEE International Conference on Data Science and Advanced Analytics (DSAA), Univ Technol Sydney. doi:10.1109/DSAA49011.2020.00096

Shahid, F., Zameer, A., & Muneeb, M. (2021). A novel genetic LSTM model for wind power forecast. *Energy*, *223*(3), 11–20. doi:10.1016/j.energy.2021.120069

Sheng, Z. Y., Wang, H. W., Chen, G., Zhou, B., & Sun, J. (2021). Convolutional residual network to short-term load forecasting. *Applied Intelligence*, *51*(4), 2485–2499. doi:10.1007/s10489-020-01932-9

Tran, T. T. K., Bateni, S. M., Ki, S. J., & Vosoughifar, H. (2021). A review of neural networks for air temperature forecasting. *Water (Basel)*, *13*(9), 15–23. doi:10.3390/w13091294

Wang, F., Xuan, Z. M., Zhen, Z., Li, K., Wang, T., & Shi, M. (2020). A day-ahead PV power forecasting method based on LSTM-RNN model and time correlation modification under partial daily pattern prediction framework. *Energy Conversion and Management*, *212*(5), 14–22. doi:10.1016/j.enconman.2020.112766

Wang, Y., Chen, Q. X., Hong, T., & Kang, C. (2022). Review of smart meter data analytics: Applications, methodologies, and challenges. *IEEE Transactions on Smart Grid*, *10*(3), 3125–3148. doi:10.1109/TSG.2018.2818167

Wu, D., Xu, Y. T., & Jenkin, M. (2022). Short-Term load forecasting with deep boosting transfer regression. *IEEE International Conference on Communications (ICC)*, Location Seoul, South Korea. doi:10.1109/ICC45855.2022.9838983

Xie, K., Yi, H., Hu, G. Y., Li, L., & Fan, Z. (2020). Short-term power load forecasting based on Elman neural network with particle swarm optimization. *Neurocomputing*, *416*(3), 136–142. doi:10.1016/j.neucom.2019.02.063

Zhang, C., & Li, R. (2021). A novel CloseDLoop clustering algorithm for hierarchical load forecasting. *IEEE Transactions on Smart Grid*, *12*(1), 432–441. doi:10.1109/TSG.2020.3015000