Short-term Load Forecasting of Residential Area Based on LSTM

Feng He[1], Menglin Wu[2, a] and Yinlong Zhao[2]

[1]School of Optoelectronic Engineering, Chongqing University of Posts and Telecommunications, Chongqing 400065, China;

[2]School of Communication and Information Engineering, Chongqing University of Posts and Telecommunications, Chongqing 400065, China.

[a]136489838@qq.com

Abstract

Driven by the new concept of smart grid, users and power control centers in the power system have realized the transformation of two-way information intelligence and flexible interaction. Whether it can accurately predict the short-term power load of residential communities will play an increasingly important role in future power regulation planning. Aiming at the randomness and volatility of short-term load in residential areas, a short-term load forecasting model based on long-short-term memory-cycle neural network is proposed. Experiments show that the model can accurately reflect the power demand trend of residential quarters in hours, and the prediction accuracy is better than the traditional recurrent neural network RNN and the fully connected back propagation neural network BP.

Keywords

Short-term load forecasting, LSTM, deep learning.

1. Introduction

With the rapid development of power system scale, the safety and stability of power systems have gradually been valued by people. As an important basis for power system analysis, load forecasting has also ushered in rapid development. Accurately predicting the short-term power load of residential communities, you can use the peak supply and demand relationship to reach equilibrium, reduce the waste of power resources, and have a positive impact on the protection of the ecological environment. Therefore, short-term power load forecasting of residential communities is of great practical significance[1-3].

Regarding the method of predicting electric load, many scholars have studied it in recent years. In [4], three load prediction models based on K nearest neighbor algorithm are established. The three independent model prediction results are combined by adjusting the weighting factor as the final output. Literature [5] proposed a multivariate K-nearest neighbor prediction model to predict the electrical load for multiple time periods in the future. In [6], a hybrid prediction model is built by wavelet transform and Bayesian neural network, and an input selection scheme is proposed. This method obtains accurate power load prediction accuracy in system-level load forecasting.In recent years, with the continuous development of deep learning technology, some deep learning models have been gradually applied to the study of time series data. Among them, the recurrent neural network is a neural network with a self-circulating structure, which allows the persistence of information flow of time series data on the network layer, and is theoretically suitable for processing time series data. However, the network structure of ordinary rnn is simple, which leads to the disappearance or explosion of the RNN network gradient, which affects its ability to learn the long-term time series relationship. In order to overcome these problems, Hochreiter et al. first introduced the LSTM architecture and

subsequently became widely used.Aiming at the problem of short-term electric load data of residential quarters, such as single dimension, strong randomness and strong volatility, this paper proposes a prediction model based on LSTM to predict the power load data in hours.

2. LSTM Recurrent Neural Network

LSTM is a long-term and short-term memory network model, which is a deformed structure of a recurrent neural network. The LSTM unit replaces the hidden unit common in ordinary RNNs and provides a unique unit state for the constantly repeating neural network module. The state of the cell is propagated along this path, and the derivative of the state value does not have a gradient disappearance or explosion problem. The LSTM unit allows the neural network to accumulate or update information.

2.1. Activation Function

The most basic unit of control information flow is the two functions sigmoid and tanh. The function curve of the function sigmoid and the function tanh is shown in Figure. 1 and Figure. 2.

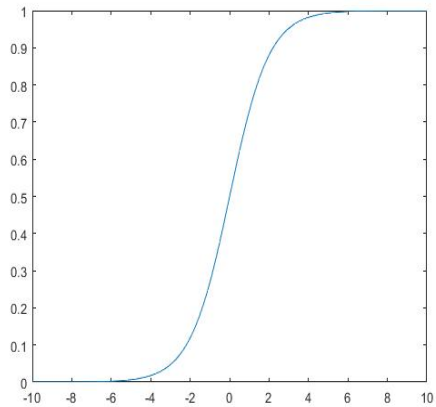

Fig 1. Sigmoid Function

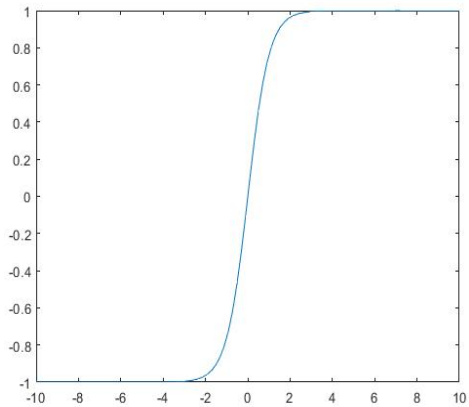

Fig 2. Tanh Function

It can be found from the figure that the function sigmoid value range is (0, 1), which is used in the LSTM model to control whether to select the information flow, indicating the importance of the information flow. The value of the tanh function is (-1, 1), which is used to control the input size of the selected stream. When the longer information is useless, the LSTM unit can also let the network forget the old unit state, set the sigmoid selection value to 0, and set the tanh value according to the importance of the new information, and inject to update the cell state. Therefore, when adjusting the weight, regardless of how far the state on the time series is from the current time node, it can affect the current time node state. This increases the sensitivity of the recurrent neural network to long-term delay states and captures information about the temporal state of any distance sequence. Relevant scholars pointed out that the inherent daily life law of residential power users is one of the important factors affecting the energy consumption in the later period. Therefore, the time series characteristics of LSTM are very suitable for short-term power load forecasting

2.2. The Propagation Process of LSTM Model

The LSTM forward propagation process is described in accordance with Figure 3 and Figure 4. The single-time LSTM cell structure is shown in Figure 3. The connection of the unexpanded LSTM cell structure at subsequent time points is shown in Figure 4.

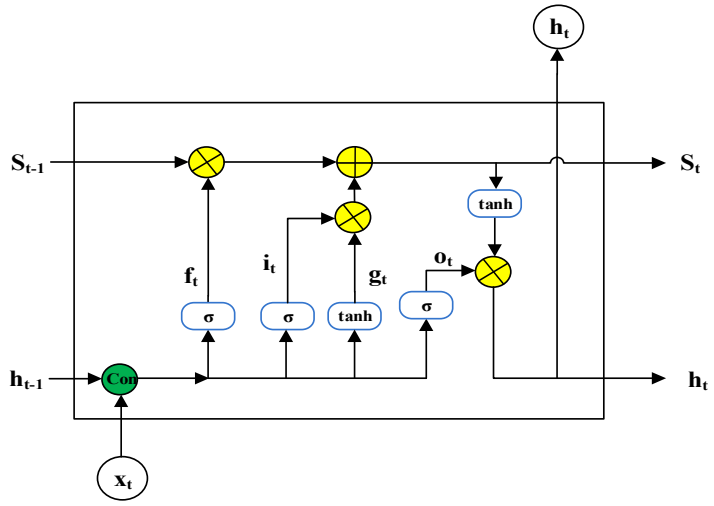

Fig 3. LSTM Internal Unit Structure

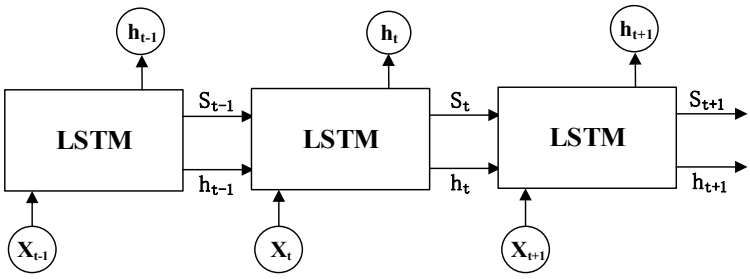

Fig 4. Multi-time Step Network Connection Diagram of Unexpanded LSTM Unit Structure

Combine h_{t-1} and x_t together to assign s_t, f_t, i_t, g_t, and o_t represent the output of the forgetting gate, input gate, input node, and output gate, respectively. W_f, W_i, W_g, W_o are the weight matrix corresponding to the input of the network activation function, and these weights are repeatedly shared in any sequence of time steps. σ stands for the activation function and tanh stands for the activation function. σ represents the *sigmoid* activation function and *tanh* represents the *tanh* activation function.

$$s_t = Con(h_{t-1,} x_t) \tag{1}$$

$$f_t = \sigma[W_f \cdot s_t + b_f] \tag{2}$$

$$i_t = \sigma[W_i \cdot s_t + b_i] \tag{3}$$

$$g_t = \tanh[W_g \cdot s_t + b_g] \tag{4}$$

$$o_t = \sigma[W_o \cdot s_t + b_o] \tag{5}$$

$$S_t = f_t \cdot S_{t-1} + i_t \cdot g_t \tag{6}$$

$$h_t = o_t \cdot \tanh(S_t) \tag{7}$$

LSTM performs network node weight training through BackPropagationTroughTime, inversely calculates the error term of each LSTM neuron output value and real value, calculates the gradient of each weight according to the corresponding error term, and applies the gradient. Finally, the weights are updated along the gradient direction.

3. Short-term Load Forecasting Model Based on LSTM Network

Predicting short-term power load of residential communities through LSTM. First, input the residential electricity consumption data collected by the smart meter, determine the parameters of the time series LSTM prediction model, perform weight training on the neural network and output the load forecasting result.

3.1. Short-term Power Load Data Preprocessing

Due to the occasional failure of the smart metering terminal equipment in the smart grid, the storage device of the power control center does not back up the updated data in time, and the communication equipment is abnormal when the user interacts with the control center, and the channel noise is high, which may result in the loss of the collected power load data. There are problems such as noise anomalies. Therefore, it is necessary to perform data cleaning operation on the collected original power data. The vacancy value in the original power data, the abnormal value containing noise, and the repeated value are processed to make the original power data more perfect.

It can be known from the internal structure of the LSTM network that since the activation functions sigmoid and tanh within the LSTM recurrent neural network have a small value range, they are very sensitive to the size of the data range. Therefore, according to the nature of the short-term power load data itself, the input power data is reduced to the range of (0, 1) by the minimum-maximum normalization method of equation (8), so that the LSTM predicts the subsequent work of the model.Let the set $T_A = (ta_1, ta_2, \cdots, ta_t)$ be the data set composed of the total load at the t time of the residential area, and normalize the load data of the residential area.

$$ta_t^* = \frac{ta_t - \min(T_A)}{\max(T_A) - \min(T_A)} \tag{8}$$

When predicting the power consumption of the power load at time L+1, it is necessary to input the power consumption at the time of the previous L in the prediction model. The input sequence length is L. The power data of the set T_A is processed to obtain a form as shown in the equation (9). The original set is a one-dimensional vector, and after the form is transformed, the set T_A becomes a set of $(t-L)\times L$.

$$T_A = \begin{pmatrix} ta_1 & \cdots & ta_L \\ \vdots & \ddots & \vdots \\ ta_{t-L} & \cdots & ta_{t-1} \end{pmatrix} \tag{9}$$

3.2. Building The LSTM Prediction Model

The pre-processed power data is divided into a training set and a verification set as input of the prediction model. When the power load sequence x corresponding to each row of the input matrix is sent to the LSTM unit corresponding to the hidden layer of the LSTM network, the LSTM hidden layer can be arbitrarily set. The number of hidden layers is formed when the number of hidden layers is large, and the number of hidden layers is set to 2. As shown in Figure. 5, the LSTM cell units are connected to each other, and their own LSTM unit state S and unit output value H are transmitted to the corresponding LSTM unit at the next moment. After all the LSTM unit information streams of the hidden layer are transmitted, the information is transmitted uniformly to the next layer of the hidden layer, and the connection manner of the two hidden layers is also a full connection mode.

After continuous learning by the network, the network uses the batch gradient descent algorithm to adjust the parameters according to the actual power consumption and the error of the predicted value, so that the mapping relationship of the prediction model approximates the nonlinear relationship between the real input and output.

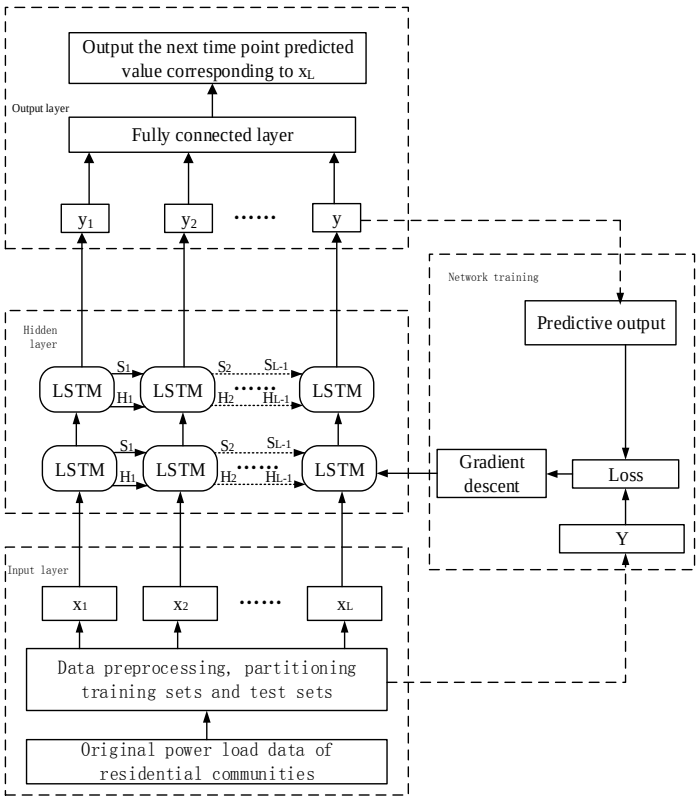

Fig 5. The LSTM Prediction Model

4. Analysis of Experimental Results

4.1. Evaluation Index for Short-term Power Load Forecasting

Through the LSTM prediction model, short-term power load forecasting is performed on various users in the community, and relevant evaluation indicators are needed to measure the accuracy of the prediction results. The evaluation indicators used in this paper are root mean square error (RMSE) and mean absolute percentage error (MAPE). The smaller the value of these two parameters, the more accurate the prediction effect. The calculation formula for the two parameters is as follows.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(Y_i - y_i)^2} \tag{10}$$

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{Y_i - y_i}{Y_i}\right| \tag{11}$$

4.2. Experimental Data

The simulation data for this chapter comes from the actual electrical load data of 90 households in the residential community, as well as the total power load data of the residential community. The electricity consumption of each household and residential community is recorded every hour. Through the analysis of historical electricity consumption data of residential residential quarters and households, the short-term power load value in the future is predicted.

4.3. Experimental Results

After the hyperparameters of the LSMT prediction model are determined, the total power load is predicted. The evaluation index uses RMSE and MAPE, and compared with the traditional cyclic neural network RNN and BP neural network. In the comparison experiment process and the unified hyperparameter scheme, the RNN neural network is the same as the LSTM, and the number of input nodes of BP is set to 20; the hidden layer parameters are consistent with the LSTM, the number of layers is set to 2, and the number of nodes is 10; The number of nodes is set to 1, and the value is weighted with reference to the power load at the next time in the input time series. The prediction results of the various methods are shown in Figure 6. The LSTM fits the actual power load curve best.

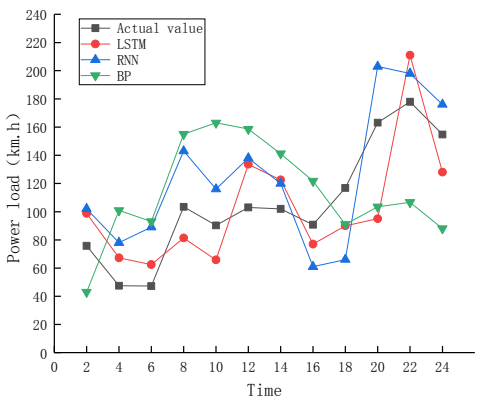

Fig 6. Three Neural Networks for Predicting Power Load Forecast of Residential Areas

The evaluation index results is own in Table 1. The MAPE of the LSTM prediction result is 26.52%, which is reduced by 26.25% and 51.16%, respectively, relative to RNN and BP. The value of RMSE was 30.21, which was 9.14% and 42.96% lower than RNN and BP, respectively. Therefore, the LSTM prediction model has improved the prediction accuracy of the total load of the residential area.

Table 1. Neural Network Evaluation Index Results

Network Model	MAPEme 1	RMSE Scheme 2
LSTM	26.52%	30.21
RNN	35.96%	32.97
BP	54.3%	52.71

5. Conclusion

This paper proposes a short-term power load forecasting model based on LSTM, and uses the power data of a certain area to train, and predicts the power demand of the community within 1h. The model overcomes the problem of low dimensionality and randomness of the data itself, and provides theoretical support for the accurate and timely prediction of the power demand of the smart grid at the residential area level, and promotes the widespread popularization of individualized power consumption schemes and reduces energy waste. An important influence. However, the model takes a little longer to train the model, and the subsequent work is prepared to seek a more efficient weight training method to improve the operational efficiency of the prediction model.

References

[1] Yun Z, Quan Z, Caixin S, et al. RBF Neural Network and ANFIS-Based Short-Term Load Forecasting Approach in Real-Time Price Environment[J]. IEEE Transactions on Power Systems, 2008, 23(3):853-858.

[2] Li H, Zhao Y, Zhang Z, et al. Short-term load forecasting based on the grid method and the time series fuzzy load forecasting method[C]// International Conference on Renewable Power Generation. IEEE, 2015:1-6.

[3] Zhang R, Dong Z Y, Xu Y, et al. Short-term load forecasting of Australian National Electricity Market by an ensemble model of extreme learning machine[J]. IET Generation Transmission and Distribution, 2013, 7(4):391-397.

[4] Zhang R, Xu Y, Dong Z Y, et al. A composite k-nearest neighbor model for day-ahead load forecasting with limited temperature forecasts[C]// Power and Energy Society General Meeting. IEEE, 2016:1-5.

[5] Al-Qahtani F H, Crone S F. Multivariate k-nearest neighbour regression for time series data-A novel algorithm for forecasting UK electricity demand[C]// Neural Network. IEEE, 2013:1-8.

[6] Ghayekhloo M, Menhaj M B, Ghofrani M. A hybrid short-term load forecasting with a new data preprocessing framework[J]. Electric Power Systems Research, 2015, 119(119):138-148.