PAPER • OPEN ACCESS

Short-term Time Series Data Prediction of Power Consumption Based on Deep Neural Network

To cite this article: Kang Xu *et al* 2019 *IOP Conf. Ser.: Mater. Sci. Eng.* **646** 012027

Short-term Time Series Data Prediction of Power Consumption Based on Deep Neural Network

Kang Xu[1,*], Ruichun Hou[1], Xiangqian Ding[1], Ye Tao[2] and Zhifang Xu[3]

[1]School of Information Science and Engineering,Ocean University of China, Qingdao,China
[2]School of Information Science and Technology, Qingdao University of Science and Technology, Qingdao,China
[3]Qingdao Haier Technology Co., Ltd. , Qingdao, China

[*]xukang@stu.ouc.edu.cn

Abstract. Under the background of the rapid development of Internet technology and the popularity of smart grids, the analysis and prediction of short-term time series data of users' power consumption has important guiding significance for grid planning, management decision of economic sector and optimization and allocation of power resources. Considering that the traditional statistical-based time series analysis method is weak in generality and can not handle the complex linear problem in prediction, the long-term dependence of the ordinary cyclic neural network model is insufficient, and the time series data has multidimensional problems, a deep neural network is proposed. The PCA-LSTM model is used for time series data prediction. The model firstly uses the PCA (principal component analysis) method to reduce the dimensionality of the electricity consumption time series data, optimizes the number of input variables, and inputs the data into the long- and short-term memory network LSTM for training prediction. The experimental results show that the LSTM network prediction based on PCA improves the accuracy of short-term time series data prediction, and also improves the convergence speed of LSTM network. It proves that the method has better prediction performance and versatility.

1. Introduction

In recent years, with the rapid development and wide application of monitoring, information communication management and Internet technology, the integrated energy system [1] with smart grid as the core has achieved relatively rapid development. The concept of “energy internet” with the goal of achieving multi-energy open interconnection, energy free transmission, energy greening, marketization and energy efficiency has emerged as the new international academic and industrial circles. And the development of emerging industries such as the Internet of Things, big data, and cloud computing [2] has also been greatly pushed to the automation process of power companies such as power grids. In the process of power supply automation, the power grid will generate corresponding big data according to the electricity consumption information collected by the sensors collected by the home appliance (air conditioner) Internet of Things. These data have typical time series data characteristics, which are generally defined as the time according to the occurrence time. In the order of the order, the ordered sets formed by the values of the same statistical indicators are arranged. Time series data also has important application value and is one of the more popular research topics in recent years [3].It is extremely important to predict the short-term time series data of the above-mentioned power consumption, not only as one of the key technologies of the energy Internet, but also to predict that we can know the user's power demand in a certain period of time in the future. To meet the needs of users, while improving energy utilization and optimizing and distributing power resources, it plays an important role in the energy Internet.

At present, the research methods for time series data are mainly divided into two categories. One is a statistical-based approach [4], including traditional autoregressive models, Autoregressive Moving Average (ARMA) models [5], and Autoregressive Integrated Moving Average (ARIMA) model [6]. The ARMA model is obtained by combining a moving average process and a linear variance equation for predicting data with linear stationary properties, and the ARIMA model is used to predict linear non-stationary properties. Although this type of method is simple, it cannot handle complex nonlinear problems in prediction, so it will produce higher prediction errors when the variables change greatly. The other is based on artificial intelligence, mainly based on support vector regression algorithm [7], BP neural network [8] and artificial depth neural network algorithm [9] as a representative of machine learning methods and suitable for complex nonlinear time sequence. Among them, the support vector regression algorithm adopts the idea of minimizing structural risk. The model structure is easy to determine and the performance is stable. BP neural network also has this advantage, but the prediction accuracy is limited. With the advent of the era of big data, the improvement of computing power and the substantial increase of training data have provided support for deep learning [10], and the deep neural network represented by Recurrent Neural Network (RNN) [11] The advantages of strong versatility and high prediction accuracy have gradually become the hot research direction of time series prediction. The purpose of the Recurrent Neural Network (RNN) is to process sequence data. In the traditional neural network model, from the input layer to the hidden layer and then to the output layer, the layers are fully connected, and the nodes between each layer are disconnected, so this common neural network There is nothing that can be done about many complicated problems. RNN is called a cyclic neural network, that is, the current output of a sequence is also associated with the previous output, that is, the nodes between the hidden layers are no longer connectionless but connected. In theory, RNN can process sequences of any length, but when the number of layers in the network increases, RNN will appear as gradient disappearance and gradient explosion. To solve the above problem, long-term memory neural network (Long Short-Term Memory Neural Network) is introduced. Network, LSTM) [12]. The emergence of long-and short-term memory neural network (LSTM) is actually a new deep learning network based on RNN, which can be used to solve the problem of gradient disappearance and gradient explosion that RNN can't solve, thus effectively making up for the defects of ordinary RNN. And controls the cumulative speed of information and demonstrates superior capabilities in predicting long-range dependent timing data.

Long and short term memory neural network (LSTM), which has certain information mining capabilities for long-distance time series data, is also widely used in speech recognition, machine translation, fault detection, and time series data prediction. For LSTM, the longer the input, the more information it contains [13].However, for the selected raw data, having multiple measurable influencing factors, that is, having multiple dimensions, greatly increases the computational complexity of the neural network, resulting in inefficiencies and the like. Aiming at the above problems, this paper studies the user's electricity consumption forecast based on the historical data of residential electricity consumption time series, and proposes a PCA-LSTM model based on deep neural network for the prediction of short-term time series data of user’s research work. The model can effectively avoid the excessive dependence of data hypothesis in statistical methods, and avoid the disadvantages of LSTM network operation complexity. The PCA (principal component analysis) [14] algorithm is used to reduce the dimensionality of the data and extract relevant influencing factors. The main component in the optimization of the number of input variables, so as to achieve the effect of time series data prediction to effectively solve related problems.

2. Establishment of PCA-LSTM model based on deep neural network

2.1. PCA dimension reduction

When using the deep neural network to predict the short-term time series data of power consumption, the neural network can continuously train and learn through a large amount of historical data of user power consumption to determine the relationship between the input data and its parameters and achieve the ideal prediction. The effect is therefore much better than the traditional method. However, the prediction of short-term time series data on electricity consumption is not an easy problem because there are many factors that affect electricity consumption, such as geographical factors, weather, temperature, holidays, weekends/working days. If a lot of factors are used to make predictions, although it can reflect the trend of power consumption, the training convergence time of deep neural networks becomes very slow, resulting in inefficiency; if the influence factors are not considered, the historical data information will be Not enough expression, so we must consider the influencing factors but not all of them. At this time, we use PCA (principal component analysis) to extract the main components of the influencing factors. Because the raw data of the electricity consumption selected in this paper has multiple dimensions, the PCA dimension reduction is used to reduce the dimensionality of the original data, and the number of input variables is optimized to reduce the computational complexity of the neural network. The PCA principal component analysis method replaces the original multidimensional data with a low-dimensional comprehensive index by adopting a linear transformation to preserve the original data information to the greatest extent. The PCA operation flow is to de-average, calculate the covariance matrix and simultaneously calculate the eigenvalues and eigenvectors of the matrix, and keep the largest eigenvectors from small to large for eigenvalues, and transform the data into a new space constructed by eigenvectors. . The algorithm flow is:

Assume that the original sample data is represented as a matrix:

$$x = \begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & \vdots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{np} \end{pmatrix} \tag{1}$$

The first step: the original historical data needs to be standardized, and the calculation formula is:

$$x_{ij} = \frac{x_{ij} - x_j}{\sqrt{\operatorname{var}(x_j)}}, (i = 1, 2, ..., n; j = 1, 2, ..., p) \tag{2}$$

among them:

$$x_j = \frac{1}{n}\sum_{i=1}^{n} x_{ij} \tag{3}$$

$$\operatorname{var}(x_j) = \frac{1}{n-1}\sum_{i=1}^{n}(x_{ij} - x_j)^2, (j = 1, 2, ..., p) \tag{4}$$

Step 2: Calculate a matrix of sample correlation coefficients for the electricity consumption data.

$$r = \begin{pmatrix} r_{11} & r_{12} & \cdots & r_{1p} \\ r_{21} & r_{22} & \cdots & r_{2p} \\ \vdots & \vdots & \vdots & \vdots \\ r_{n1} & r_{n2} & \cdots & r_{np} \end{pmatrix} \tag{5}$$

After the standardized raw data is still represented by x, the correlation coefficient of the standardized sample data is:

IOP Conf. Series: Materials Science and Engineering **646** (2019) 012027 doi:10.1088/1757-899X/646/1/012027

$$r_{ij} = \frac{1}{n-1}\sum_{i=1}^{n} x_{ij}x_{ij}(i, j = 1,2,..., p) \tag{6}$$

Step 3: Use Jacobbian to find the eigenvalues of the correlation coefficient matrix r and the corresponding feature vector α :

$$\lambda_i, i = 1, 2,..., p \tag{7}$$

$$\alpha_i = (\partial_{i1}, \partial_{i2},..., \partial_{ip}), (i = 1, 2,..., p) \tag{8}$$

The fourth step: select the main component that meets the requirements and find the principal component expression. The sample data obtained by principal component analysis generally can obtain p principal components, but the information content of the p principal components is often different, and there will be differences between them. If the principal component corresponds to a large variance, then the corresponding amount of information is large, and vice versa. When actually selecting, it is usually chosen to reflect the main components of the original data information to a large extent. The basis of the selection is to accumulate and sum according to the contribution rate of each principal component, that is, the proportion of each principal component in the sum of all variances. The mathematical expression of the contribution rate is:

$$a_i = \lambda_i / \sum_{i=1}^{p} \lambda_i \tag{9}$$

The number of principal components to be selected at the end. That is, the determination of the value of k in the $F_1, F_2, F_3, \ldots, F_k$ is determined by the cumulative contribution rate $G(k)$ of the principal component. among them,

$$G(k) = \sum_{i=1}^{k} a_i \tag{10}$$

Usually, as long as the sum of the contribution rates reaches a certain value or more, it can be considered that the comprehensive indicators obtained can fully reflect the vast majority of information contained in the original sample data.

Step 5: Calculate the principal component $F_i = \partial_{1i}x_1 + \partial_{2i}x_2 + \ldots + \partial_{pi}x_p, i = 1, 2,..., p$, the specific form can be as follows,

$$\begin{pmatrix} F_{11} & F_{12} & \cdots & F_{1k} \\ F_{21} & F_{22} & \cdots & F_{2k} \\ \vdots & \vdots & \vdots & \vdots \\ F_{n1} & F_{n2} & \cdots & F_{nk} \end{pmatrix} \tag{11}$$

Step 6: Calculate the principal component load, the formula is:

$$q_i = \sum_{i=1}^{p} \sqrt{\lambda_i \alpha_i}, i = 1, 2,..., p \tag{12}$$

2.2. Long-term and short-term memory network (LSTM) model

Circulating neural network (RNN) [15] is a popular deep learning model in recent years, and has been successfully applied in many fields such as language recognition and text categorization. However, it also faces the problem of gradient explosion and gradient disappearance. The problem cannot be solved very well. Thus, the neural network model after optimization of the cyclic neural network-LSTM is obtained. The main change is the introduction of three types of gates at each index index position when remembering information: input gates, forgetting gates, and output gates. Through these three kinds of gate structures, the state information of the unit is controlled and maintained, the cell state information is stored and modified, and the information is selectively added or the existing information is deleted according to the result, thereby solving the long-term gradient of the RNNs. The problem of explosion and disappearance has also realized long-term memory and proved its superiority in a large number of experiments, so it is widely used in time series problems of long-term dependence. For the prediction problem mentioned in this paper is a typical timing problem, so the model is selected for training and prediction.

The specific internal structure of LSTM is shown in Figure 1:

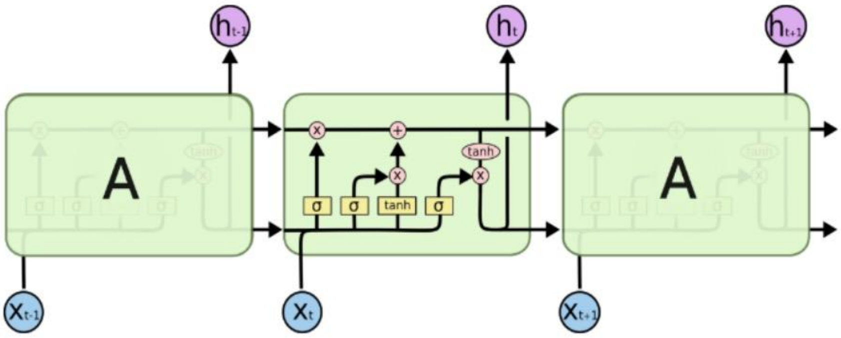

Figure 1. LSTM model structure

The steps for the LSTM to update the internal state during training are as follows:

In the first step, the LSTM needs to pass the forgetting gate to determine the information that the cell state needs to be discarded. The forgetting gate obtains a value between 0 and 1 according to the historical information of the previous state and the currently input information to represent the probability of forgetting the state of the upper layer of cells. That is to say, the forgetting gate is to control whether or not to forget, that is, to control whether the state of the upper layer hidden cells is forgotten with a certain probability. Using the activation function (generally the sigmoid function), The hidden state of the input previous sequence is denoted as h_{t-1} and the input of this sequence is recorded as x_t by which the output of the forgetting gate between [0.1] is recorded as $f(t)$; 0 represents all information of the discarded history state, and 1 represents all information about the historical status. Where x_t represents the input of t time, h_{t-1} represents the value of the hidden layer at $t-1$ time, W_f , U_f , and b_f represent the weight of the LSTM hidden layer in the forgetting gate, the weight of the current input, and the offset in the forgotten gate; $f(t)$ represents the output of the forgetting gate. The specific calculation is as follows:

$$f(t) = \sigma(W_f \cdot h_{t-1} + U_f \cdot x_t + b_f) \tag{13}$$

In the second step, the LSTM determines the update information through the input gate and places it in the cell. After the circulatory neural network has forgotten part of the previous state, it needs to supplement the latest memory from the current input, and the input gate is responsible for processing the input of the current sequence position to supplement the latest memory. The input gate is composed of two parts. The first part uses the sigmoid function to output as i_t , the second part uses the tanh function to output as a_t , and the two obtain the information updated to the cell state by the bitwise product, where W_i , U_i , W_a , U_a , b_i , b_a are the coefficients and offsets of the linear relationship. The specific calculation is as follows:

$$a_t = \tanh(W_a \cdot h_{t-1} + U_a \cdot x_t + b_a) \tag{14}$$

In the third step, the cell state is updated. After the first two "gates", the deletion and addition of the delivery information can be determined, that is, the update can be performed. Both the results of the Forgetting Gate and the results of the input gate will act on the Cell Status C_t and put the C_{t-1} update to the final cell state C_t , That is:

$$C_t = C_{t-1} * f_t + i_t * a_t \tag{15}$$

Finally, the result of the final cell output is determined by the output gate. First, the input information is determined by the sigmoid layer to determine the output of the input information, and then the cell state is processed by the tanh function. Finally, the final output is determined by bitwise multiplication with the previously obtained sigmoid value, and the specific calculation is as follows:

$$O_t = \sigma(W_o \cdot h_{t-1} + U_o \cdot x_t + b_o) \tag{16}$$

$$h_t = O_t * \tanh(C_t) \tag{17}$$

Long-term and short-term memory network (LSTM) models are effectively avoided by using very specialized structural design. The problems of gradient disappearance and explosion in the training of conventional cyclic neural network model (RNNs) can effectively avoid the use of historical sequence information. . In order to fully predict the power consumption short-term sequence data, the LSTM model is used to effectively predict the data based on the dimensionality reduction of the PCA.

3. PCA-LSTM model for time series data prediction

3.1. Data preprocessing

By observing the previous time series data prediction experiment, in the process of selecting this historical data and influencing factors, different types of influencing factors will be selected according to the principle of comprehensiveness. Because there are different dimensions in the data, the numerical difference is relatively large, and the input and output range of the nonlinear activation function in the neural network model is limited. In order to avoid linear representation of historical data, the data is compressed and the neural network is avoided. When the element appears saturated, it is necessary to standardize the data of the time series data so that the range of the range is in the same interval. The purpose of data standardization processing is to remove the influence of the data from the dimension and summarize the statistical distribution of the same sample.

In the short-term time series data of the user's electricity consumption, due to the different influencing factors between different user historical days, the value of the electricity consumption varies greatly. If the processing is directly processed, the prediction result will be inaccurate, so in order to improve the data prediction. The accuracy of the analysis requires normalization of the data for each dimension to reduce errors and improve accuracy. There are many methods of normalization, with min-max maximum and minimum normalization, that is, linear changes to the original data to map values to the [0-1] range; and Z-score standardization method. The normalization method used in this paper is the min-max method for data standardization. The formula is as follows:

$$X^* = \frac{X_i - x_{\min}}{x_{\max} - x_{\min}} \tag{18}$$

Among them, $x_{\min}$ is the smallest value in the original sample data sequence, $x_{\max}$ is the largest value in the original sample data, X^* is the data sequence that the neural network needs to input, and X_i is the original data sequence.

3.2. Using PCA to perform feature dimensionality reduction on training samples

As mentioned above, in the actual time series data prediction, the short-term time series data of power consumption often contains multiple influencing factors, resulting in a relatively large amount of information or even redundant information. If the training sample dimension of the deep neural network is too much, the corresponding network will become more complicated, which will lead to prolonged network training time, inefficient, and the prediction accuracy may also decrease accordingly. Subjective selection of the dimensions of the training samples will easily lead to the loss of important information, resulting in a decrease in the predictive performance of the neural network. If the original data samples are first subjected to dimensionality reduction, the primary response is selected with less loss information. The comprehensive dimension will reduce the computational complexity of the neural network, improve the training efficiency of the neural network, and greatly improve the prediction performance. In this paper, the PCA dimension reduction processing of the input short-term time series data is used to achieve data dimensionality reduction.

After reducing the dimensionality of the sample data PCA, we can use a small number of features to train and predict the deep neural network, which not only improves the accuracy of the experimental prediction but also improves the convergence speed of the neural network.

3.3. Training and testing of LSTM network model

The LSTM network model is used to build a deep learning framework. At the same time, the LSTM network model is built using the popular Tensorflow deep learning platform. There are many activation functions. The relu function is used in this paper.

First, some of the original power consumption sequence data samples are taken as the training set, and the remaining small power consumption time series data samples are used as the test set. At the end of the training, the weights and offsets of the LSTM network are saved, and the trained deep neural network model is used to test on the corresponding test set, to see how accurate the predictions is, and to compare the predictive effect fobtained by simply using the RNN and LSTM neural network models. effect.

The LSTM network model for short-term time series data prediction needs to determine the following parameters: Timestep time step, number of hidden layer layers, number of hidden layer neurons, number of iterations, learning rate, etc. In the process of training, the corresponding hyperparameters are corrected in real time according to the effect of the training, so as to achieve a perfect prediction effect.

4. Experiment and result analysis

4.1. Introduction to the data set

The data used in the short-term time series data forecasting of this electricity consumption is a user's electricity record from a total of 380 days from July 1, 2017 to July 15, 2018, and is mainly sampled by smart meters. The data taken include user ID, electricity date, total electricity consumption, “tip” time consumption, “peak” time consumption, “flat” time consumption, and “valley” time consumption. The total amount of electricity used by each user per day and the characteristics of the electricity consumption data during the “peak peak” period, as well as the influencing factors mainly include date, day of the week, holidays, temperature, weather, and so on.

Accurate analysis and prediction first need to process a large amount of raw data, which is different from the data preprocessing mentioned above. At this time, it mainly includes the combination of data, removal of invalid records, and removal of sparse records. The sample data are all from the real life data, all the acquisition process will be affected by external factors, there is "noise", and some of the user's part of the period of record data will be missing, resulting in a large error. In order to ensure the reliability and reality of the forecast, the users with complete data in the original data set and the data with total power consumption less than 50 are selected, because the survey found that the residential electricity demand does not exceed 50 kWh per day. Through filtering, a total of 4,456 users and 380 days of electricity usage data from July 1, 2017 to July 15, 2018 were selected. Among them, a part of the original data is selected as the training set for the training of the model, and the remaining data of the original data is used as a test set to verify the feasibility of the model.

4.2. Evaluation criteria for network models

In this paper, the evaluation criteria for the short-term time series data prediction results of users is the Root Mean Square Error, which reflects the performance of the model to control the absolute error. The calculation formula is:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}[\hat{p}(i) - p(i)]^2} \tag{19}$$

In the formula: $p(i)$ represents the actual value of the electricity consumption data, $\hat{p}(i)$ represents the predicted value of the user electricity consumption data; n represents the number of predicted verification data; i is the predicted point serial number. The smaller the $RMSE$, the better the prediction effect of the model.

4.3. Experimental results and comparative analysis

In this paper, the PCA-LSTM model based on deep neural network is used to predict the short-term time series data of the user's electricity consumption. In order to more intuitively express the prediction effect of the text model on the user's electricity consumption data, all users are selected from May to July 2018. A total of 60 days of electricity recorded data for prediction, and compared with SVM, RNN model, the comparison of the user's electricity consumption prediction effect and actual electricity consumption and the comparison between the predicted values and actual values of different models are shown in Figures 2, 3, and 4. Shown as follows:

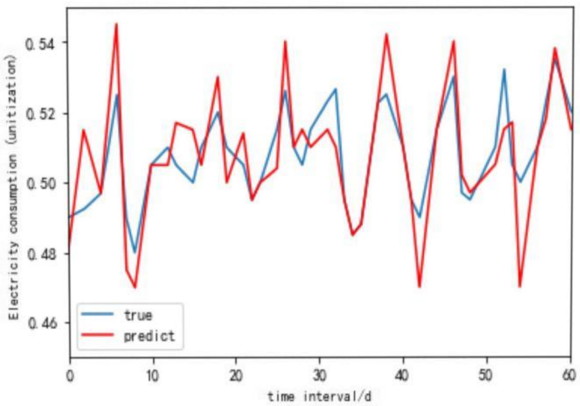

Figure 2. SVM model predicted value and actual value fit

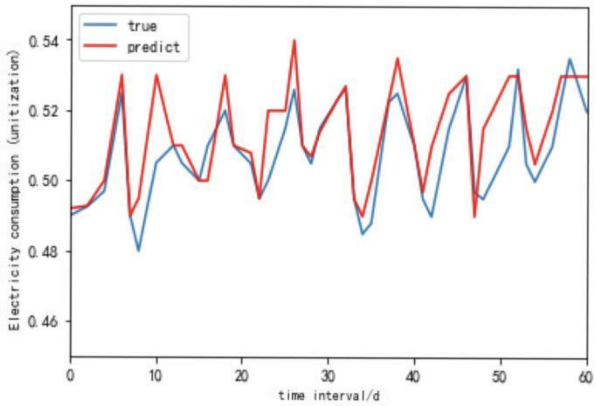

Figure 3. Comparison of predicted and actual values of RNN model

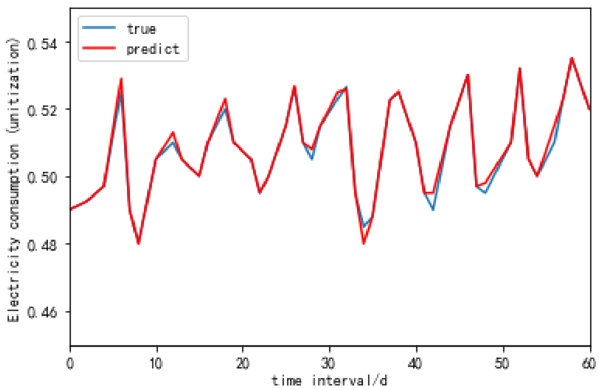

Figure 4. Comparison of predicted and actual values of PCA-LSTM model

In the above figure, we can visually see the prediction effects of different models. From the comparison between the predictions of the various time periods and the real values, we can see that the proposed model can make a good prediction for the user's electricity consumption time series data. The result also shows that the prediction method of this paper is effective, and the user short-term time series data can be well fitted.

Table 1. Comparison of experimental results between benchmark model and PCA-LSTM

Model	RMSE
HA	1.73
ARIMA	1.55
SVM	0.97
BP	0.72
RNN	0.64
PCA-LSTM	0.48

As shown in Table 1 above, the prediction effects of the deep neural network based PAC-LSTM model and the other five benchmark models used in this paper are given. By comparing the results of RMSE, the following conclusions can be drawn:

(1) Compared with other methods, the PCA-LSTM model based on deep neural network used in this paper achieves the best prediction effect and the best performance, so the model can better fit the user's electricity consumption time series data. ;

(2) The effect of the HA method is much lower than other prediction models, which is the worst method, because the method only uses the historical data average to predict the future, ignoring the characteristics of many data itself, making the error of the result larger;

(3) The prediction effect of the neural network model is better than the regression model as a whole, which also shows that the model used in this paper has a good prediction effect;

Through the comparative analysis of the above results, it is verified that the model proposed in this paper has high prediction accuracy and good prediction effect, and can better fit the time series data and power consumption behavior of user electricity consumption.

5. Conclusion

For the long-term and short-term memory network (LSTM) model, the characteristics of long-term and short-term dependence information of learning time series data and the excessive data dimension lead to the decline of neural network convergence speed. This paper proposes a PCA-LSTM model based on deep neural network for users. The short-term time series data of power consumption is used for prediction, and the prediction effect obtained by the model is simply compared with other models to highlight the feasibility of the model. By analyzing the problems of traditional neural network RNNs, and using the variant model LSTM based on the PCA dimension reduction, the required prediction effect is achieved. The experimental results show that the proposed prediction model not only improves the accuracy of time series data prediction, but also improves the convergence speed of LSTM network, and verifies the validity of the model.

In the future research, we will continue to explore the parameter selection problem of the network model, so how to conduct the adjustment is also the next research work.It is also necessary to consider expanding the data set and applying different data sets for training tests, thereby further improving the generalization ability of the model.

Acknowledgments

This work was financially supported by National Key Research and Development Plan 2017YFB400903.

References

[1] Wu J Z.The Driving and Current Situation of European Integrated Energy System Development[J].Automation of Electric Power Systems,2016,40(5):1-7.

[2] Luo J Z,Jin J H,Song A B,et al. Cloud computing: architecture and key technologies[J].Journal on Communications,2011,32(7):3-21.
[3] Weigend A S.Time series prediction: forecasting the future and understanding the past[M].Routledge,2018.
[4] BOX G E, P, JENKINS G M,REINSEL G C,et al.Time series analysis:Forecasting and control [M].New York:John Wiley & Sons,2015.
[5] Xiong Z B.Research on GDP Time Series Prediction Based on ARIMA and Neural Network Integration[J].Journal of Mathematical Statistics and Management,2011,30(2): 306-314.
[6] Langkvist M,Karlsson L,Loutfi A.A review of unsuper-vised feature learning and deep learning for time series modeling[J].Pattern Recognition Letters, 2014,42: 11-24.
[7] ZHANG Fan, DEB C,LEE S E, et al.Time series forecasting for building energy consumption using weighted Support Vector Regression with differential evolution optimization technique [J]. Energy and Buildings,2016,126:94-103.
[8] WONG F S .Time Series forecasting using backpropagation neural networks[J]. Neurocomputing,1991,2(4):147-159.
[9] WANG Z,SRINIVASAN R S.A review of artificial intelligence based building energy use prediction:contrasting the capabilities of single and ensemble prediction models[J].Renewable mand Sustainable Energy Reviews, 2017,75:796-808.
[10] Cheng X Q, Yan X L, Wang Y Z, et al.Overview of Big Data Systems and Analysis Techniques[J].Journal of Software,2014,25(9):1889-1908.
[11] CONNOR J,ATLAS L.Recurrent neural networks and time series prediction[C] //IJCNN-91-Seattle International Joint Conference.on Neural Networks. Seattle,WA,USA:IEEE,1991:301-306.
[12] HOCHREITER S, SCHMIDHUBER J. Long short-term memory [J].Neural computation, 1997,9(8):1735-1780.
[13] Lai G,Chang W C,Yang Y,et al. Modeling long-and short-term temporal patterns with deep neural networks[C]//The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval.ACM,2018:95 一 104.
[14] YANG Kairui,MENG Fanrong,LIANG Zhizhen. Adaptively weighted PCA algorithm [J].Computer Engineering and Applications,2012,48(3):189-191.
[15] Hochreiter S,Bengio Y,Frasconi P,et al.Gradient flow in recurrent nets the difficulty of learning long-term dependencies[C].A Field Guide to Dynamical Recurrent Networks,Wiley-IEEE Press,2001:237-243.