Multi-layer Long Short-term Memory based Condenser Vacuum Degree Prediction Model on Power Plant

Kuan Lu[1]*, Song Gao[1], Pang Xiangkun[1], Zhu lingkai[1], Xiangrong Meng[1], Wenxue Sun[2]

[1] State Grid Shandong Electric Power Research Institute, Jinan, Shandong, 250021, China
[2] State Grid Zhangqiu Power Supply Company, Jinan, Shandong, 250001, China

Abstract: A multi-layer LSTM (Long short-term memory) model is proposed for condenser vacuum degree prediction of power plants. Firstly, Min-max normalization is used to pre-process the input data. Then, the model proposes the two-layer LSTM architecture to identify the time series pattern effectively. ADAM（Adaptive moment）optimizer is selected to find the optimum parameters for the model during training. Under the proposed forecasting framework, experiments illustrates that the two-layer LSTM model can give a more accurate forecast to the condenser vacuum degree compared with other simple RNN (Recurrent Neural Network) and one-layer LSTM model.

1 Introduction

The condenser is the cold source of thermodynamic cycles in power plants, and its performance directly affects the unit's peak load regulation capability, operation safety and thermal efficiency. Condenser vacuum degree is an index that comprehensively reflects the condenser's running state. Due to complexity of the model of condenser vacuum, an effective method is needed to predict condenser vacuum degree so as to realize an optimal control.

In recent years, varieties of machine learning algorithms have been applied to prediction of condenser vacuum and achieved certain results. Ge Xiaoxia et al. [1] proposed a prediction model of condenser vacuum based on Drosophila algorithm optimization generalized regression neural network. Li Jianqiang et al. [2] constructed a steam condenser vacuum prediction model based on the PSO-SVR model. Zhang Liping et al. [3] used the particle swarm optimization algorithm to optimize Elman network parameters to predict the steam condenser vacuum on the high- and low-pressure sides. Wang Jianguo et al. [4] used particle swarm optimization to optimize radial basis function neural network (RBF) parameters and established a condenser vacuum prediction model. Zhang Hai et al. [5] put forward an Elman neural network condenser vacuum model based on particle swarm optimization. However, with the increase of data volume and input dimensions, machine learning models show obvious weaknesses in the calculation speed and forecast accuracy, which are mainly reflected by over-fit during training, slow training speed and poor generalization ability of the models. On the other hand, existing studies have not considered the relationship of temporal sequence in condenser vacuum, thus lacking the modeling research to improve the time series effect.

LSTM [6] is a deep learning model with a recurrent network structure. It has made a series of achievements in natural language processing [7], image recognition [8] and other fields, and its prediction accuracy and reliability have been significantly improved. Daniel M, Kasun A and Milos M [9] compare the short-term power load forecasting accuracy of standard LSTM model and Seq2seq (Sequence to sequence) network based on LSTM. Zheng H T, Yuan J B and Chen L [10] give a general framework, which combines similar day selection, empirical mode decomposition and LSTM model to predict short-term load.

In this paper, a multi-layer LSTM based condenser vacuum degree prediction model for power plants is proposed. Firstly, the real-time weather and vacuum history information is normalized using Min-max method. Secondly, a two-layer LSTM architecture is designed as the forecast model. In the process of training, the model adopts ADAM method to find the optimum parameters. Last, this model is applied a power plant in Shandong Province from July 1 to 26, 2017 for vacuum degree prediction. Compared with the baseline results from simple RNN network model and single-layer LSTM model, the prediction effect of this method has been significantly improved.

2 Long short-term Memory

LSTM is a kind of gate-controlled recurrent neural network, which uses input gate, forgetting gate and output gate to deal with the gradient vanishing problem due to the long-time dependency in inputs. The structure of LSTM is shown in Fig. 1:

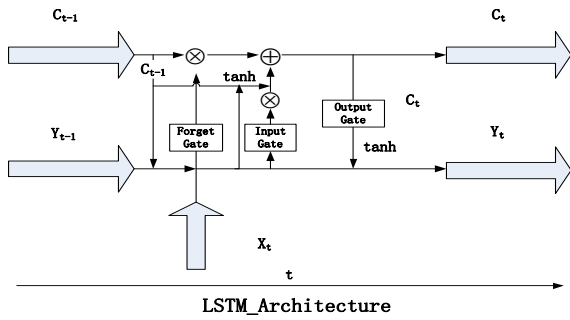*Corresponding author's e-mail: lk83@163.com

Figure1. LSTM Cell Structure

In the LSTM structure described above, the input sequence is $= [x_1, x_2, x_3, ..., x_T]$, where T is the length of the temporal sequence. The output sequence is $y = [y_1, y_2, y_3, ..., y_T]$ wherein, y_t is represented as the output value at t. The cell state sequence is $c = [c_1, c_2, c_3, ..., c_T]$, where c_t is the state value. In addition, i_t, f_t and o_t denote output gate, forget gate and output gate.

$$i_t = \sigma_{sig}(W_i c_{t-1} + U_{ix} x_t + U_{iy} y_{t-1} + b_i) \quad (1)$$

$$f_t = \sigma_{sig}(W_f c_{t-1} + U_{fx} x_t + U_{fy} y_{t-1} + b_f) \quad (2)$$

$$o_t = \sigma_{sig}(W_o c_{t-1} + U_{ox} x_t + U_{oy} y_{t-1} + b_o) \quad (3)$$

$$c_t = f_t \otimes c_{t-1} + i_t \otimes \phi_{tanh}(W_c c_{t-1} + U_{cx} x_t + U_{cy} y_{t-1} + b_c) \quad (4)$$

$$y_t = o_t \otimes \phi_{tanh}(c_t) \quad (5)$$

where $\otimes$ denotes the Hadamard product; W, U and b are weight matrices which need to be learned during training; ϕ_{tanh} and σ_{sig} denote the Tanh activation function [11] and Sigmoid activation function.

3 Condenser Vacuum Degree Prediction Model

3.1 Data Preprocessing

In order to prevent neurons from over-saturation during training, input data need to be normalized. In addition, it can also help to prevent the objective function from being affected by excessive deviation of a certain dimension due to its unit's effect. The formula of normalized data here is:

$$x_{scaled} = \frac{x - x_{min}}{x_{max} - x_{min}} \quad (6)$$

where x_{max} and x_{min} are the maximum and minimum vectors of input.

3.2 Model Construction

Condenser vacuum degree is affected by inlet temperature of cooling water, temperature rise of cooling water and end difference of condenser [13], of which the

inlet temperature of cooling water is mainly affected by ambient temperature, and external atmospheric pressure. Therefore, the model takes air temperature, humidity, wind speed, air pressure and historical vacuum values as inputs. The vacuum degree of the next day was taken as the output variable. The model adopts a two-layer LSTM structure as shown in Fig. 2.
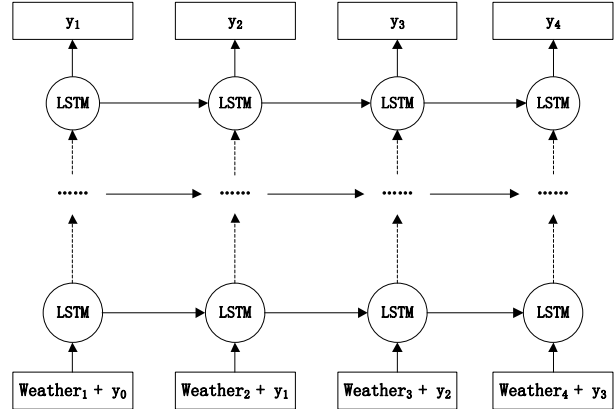

Figure2. LSTM based Condenser Vacuum Degree Prediction Model

The inputs of the model are weather data in the forecast period as well as vacuum value in the historical period. Output is the vacuum value in the forecast period. Here, the time step is set to 144, since the predicting vacuum value for the next 24 hour is required a 10 minutes time interval. Mean square error for the target loss function is as follows:

$$Loss = \frac{1}{M} \sum_{t=1}^{M} (y_t - \hat{y}_t)^2 \quad (7)$$

ADAM [14] is selected as the gradient descent optimization algorithm.

3.3 Forecasting Process

Using multi-layer LSTM network model to predict condenser vacuum degree mainly follows the steps as:

1. Data collection: select temperature, air pressure, wind speed, humidity data and condenser vacuum degree historical data, where the time interval should be 10 minutes for one day.

2. Data normalization: divide the collected data into training set data and test set data, and use a Min-max method to normalize the data to [0,1].

3. Model training and prediction: train the model using a two-layer LSTM architecture and the normalized test set data are brought into the trained model for prediction. The whole process is shown in Fig. 3.
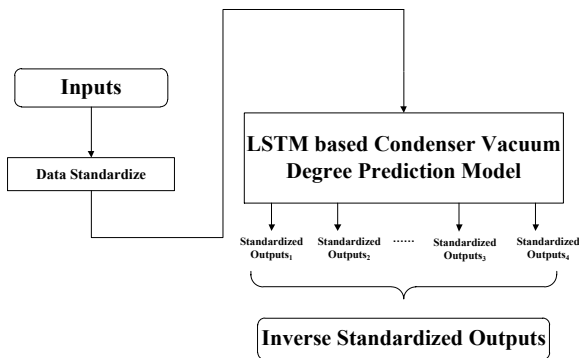
Figure3. Model Prediction Flowchart

4 Dataset and experimental results

The proposed model is applied to a 600MW power plant in Shandong Province for experiments. The training data ranges from June 1 to July 26, 2016. We use Pytorch to build a deep learning model and train the model using GPUs. Here, the batch size for training data is set to 32, and batch size for test data is 64. The initial learning rate is 0.01 and decays exponentially starting from the 100th training epochs. The error measurement here is R_{MSRE} (Root Mean Square Error, RMSE):

$$R_{MSRE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\frac{\hat{y}_i - y_i}{y_i}\right)^2} \qquad (8)$$

in which, the actual value is expressed as y_i and the predicted value is $\hat{y}_i$.

In the experiment, we select the simple RNN network and single-layer LSTM model as comparison. Fig. 4 shows the results for the three models. It can be observed that after 500 epochs of training, the error level of the simple RNN network model reaches the top of the three. This is due to the inability of the Simple RNN network in learning a temporal relationship, which results in an error level of 2.22%. Meanwhile, due to the only one layer in the model, the capacity of the single-layer LSTM model (with a prediction error of 0.95%) is smaller than that of the multi-layer LSTM model (with a prediction error of 0.91%), so the error level is higher than that of the latter.
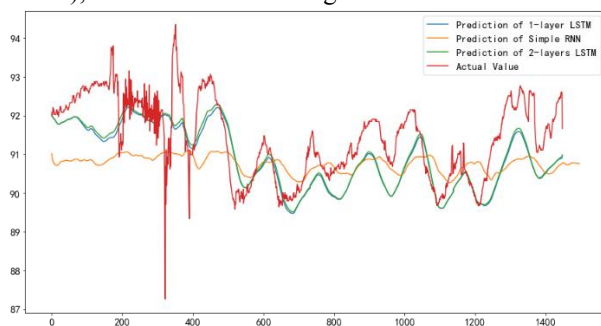

Figure4. Actual Values vs. Predictions

5 Conclusion

The goal of the present work is to investigate condenser vacuum degree prediction forecast performance using a multi-layer LSTM model. We show comparisons of this multi-layer model with a simple RNN and one-layer LSTM model that using a more complex architecture can increase the model capacity thereby lowering the test error. Using the proposed model in this paper, power plants could improve steam efficiency and grid companies can make more effective arrangement on the power distribution according to power plants' generation capacity.

References

1. Xiaoxia Ge, Hongchuang Xiao, et al. (2018) Vacuum prediction of condenser based on generalized regression neural network optimized by fruit fly algorithm. Turbine Technology, 60: 208-2012.

2. Jianqiang Li, Xingxu Chen, et al. Target value prediction of condenser vacuum degree based on ps0-svr model. Electric Power Science and Engineering, 2017(2):66-72.

3. Liping Zhang, Haotian Chen, Fengwei Wang, et al. (2015) Application of PSO-algorithm modified Elman neural network in vacuum prediction for dual pressure condensers. Thermal Power Generation, 3:53-57.

4. Jianguo Wang, Leping Lin, et al. (2015) A vacuum value prediction method for steam condensers using RBF neural network optimized by particle swarm algorithm. Thermal Power Generation., 10:72-76.

5. Hai Zhang, Jian Pu, et al. (2010) A model for predicting vacuum in the condenser based on Elman neural network by using particle swarm optimization algorithm. Thermal Power Generation, 4: 7-11.

6. Hochreiter, Sepp, Jürgen s. (1997) Long short-term memory Neural Computation., 9: 1735-1780.

7. Graves, Alex, Navdeep J, et al. (2013) Hybrid speech recognition with deep bidirectional LSTM. In Automatic Speech Recognition and Understanding (ASRU), 273-278.

8. Fan Xu, Hua Cheng, Yiquan Fang. (2017) A Gait pattern classification method based on CLSTM [J]. Journal of East China University of Science and Technology (Natural Science Edition), 43: 553-8.

9. Daniel L M, Kasun A, Milos M. (2016) Building energy load forecasting using deep neural networks. In: Industrial Electronics Society, IECON 2016-42nd Annual Conference of the IEEE, 7046-7051.

10. Huiting Zheng, et al. (2017) Short-term load forecasting using EMD-LSTM neural networks with a Xgboost algorithm for feature importance evaluation. Energies, 10:1168.

11. Goodfellow Ian, Bengio Yoshua, Courville Aaron. Deep learning. (2016) Cambridge, Massachusetts: MIT Press.

12. Han J, Moraga C. (1995) The influence of the sigmoid function parameters on the speed of backpropagation learning. In: International Workshop on Artificial Neural Networks, 195-201.

13. Chang Yu. (2014) Analyzes of the condenser vacuum degree prediction based on a 660MW power plant in China. Huazhong University of Science & Technology.

14. Kingma, Diederik, and Jimmy Ba. (2014) ADAM: A method for stochastic optimization. In: International Conference on Learning Representations (ICLR).