Particle Swarm Algorithm to Optimize LSTM Short-Term Load Forecasting

Xifeng Guo[1], Yichun Zhang[1], Nan Chen[1], Jiuyi Lv[1], Yilin Wang[2]

1. Shenyang Jianzhu University, Shenyang 110168, China.

2. The Fourth Construction Engineering Co., Ltd., China Construction Second Engineering Bureau, Tianjin 300457, China.

Abstract: Accurate load forecasting is of great significance for national and grid planning and management. In order to improve the accuracy of short-term load forecasting, an LSTM prediction model based on particle swarm optimization (PSO)algorithm is proposed. LSTM has the characteristics of avoiding gradient disappearance and gradient explosion, but there is a problem that parameters are difficult to select. Therefore, particle swarm optimization algorithm is used to help it select parameters. The experimental results show that the optimized LSTM has higher prediction accuracy.

Keywords: LSTM; PSO; Short-Term Power Load Forecasting

1. Introduction

When Google's AlphaGo beat the Korean Go world champion Lee Sedol, when we analyze the reasons for the failure, we will find that in Go we can only improve by playing against our opponents again and again, such a game takes an hour or even for us. Longer, and artificial intelligence can use algorithms to play games with itself all the time to improve its own disadvantages. In one hour of our chess game, he may have played 10 games, and in a week, artificial intelligence may surpass us in ten years. The number of chess games, and Google also used the process of various previous Go games to let the artificial intelligence learn. From this point of view, Lee Sedol's failure was not accidental. Artificial intelligence is not only applied to Go. In fact, artificial intelligence has long been connected with our lives. Through machine learning, computer systems have been able to perform speech recognition, image classification, text recognition, etc. The test results of the game of Go further prove that Computers can handle more complex algorithms, which also marks the gradual entry of human beings into the era of artificial intelligence. The digitalization trend of the energy industry has created new opportunities for the application of artificial intelligence and machine learning technologies in the power industry [1-3]. How to reduce the waste of electric energy in the power system is the most critical problem encountered at present. Due to the difficulty in storing a large amount of electric energy and the characteristics of power demand changing all the time, it is required that the power generation of the system should achieve dynamic balance with the change of load, and improving the accuracy of load forecasting is conducive to improving the utilization rate of power generation equipment and the effectiveness of economic dispatch [4]. The accuracy of predictions is increasingly important. However, at present, there is a problem of insufficient holiday forecasting accuracy in this field. Since it is difficult to balance power generation and load changes during holidays, it is easy to cause overloaded operation of the power grid, it is difficult to quickly complete power dispatching, and it is easy to cause damage to related working devices.

At present, the research on load forecasting at home and abroad is mainly divided into two categories. The first category is the time series analysis method, in which the time series is divided into univariate time series and multivariate time series. Among them, multivariate time series requires multiple features for each time node to build a model. The main algorithms include differential integrated moving average autoregressive model (ARIMA), random forest, XGBOOST, etc. Reference [4] compares and analyzes XGBOOST with multiple algorithms , indicating that the constructed power load XGBOOST prediction model has advantages over random forest, Bayesian and KNN methods in terms of computational speed and prediction accuracy. The basic idea of the above algorithm is to predict the future load value from the past load value and the current load value of the random time series. The class is a machine learning method such as backpropagation (BP) neural network is used to predict in literature [5-7], gray projection and random forest algorithm are used in literature [8], and deep belief network is used in literature [9] to predict , the literature [10] uses the multi-core support vector machine algorithm for regression prediction, and the expert system method for prediction [11] and so on. Although the LSTM network model can fully reflect the long-term historical process in the input time series data, it cannot mine the effective information and potential relationships contained in non-continuous data. A common problem of these algorithms is the lack of consideration of time series, which requires researchers to add time features to ensure the accuracy of prediction.The above algorithms have problems such as insufficient accuracy and slow running speed when dealing with a large amount of nonlinear relational data. The LSTM algorithm has the advantages of the above algorithms and has fast running speed and high accuracy. However, LSTM has the problem that the parameters are difficult to determine, and different parameters have a huge impact on the model. Therefore, the PSO algorithm is proposed to optimize the parameters of the LSTM algorithm. The experimental results show that it has higher prediction accuracy and prediction speed.

2. Lstm

2.1 LSTM Principle

LSTM network is an improved time recurrent neural network(RNN). It has also been modified since it was proposed, adding additional Forgotten door. The improved LSTM network solves the problem of "gradient disappearance" in model training, and can learn long- and short-term dependency information of time series. It is the most successful RNN architecture at present and has been applied in many scenarios.

3.PSO

3.1 PSO Principle

The particle swarm algorithm simulates birds in a flock by designing a massless particle. The particle has only two properties: speed and position. The speed represents the speed of movement, and the position represents the direction of movement. Each particle searches for the optimal solution independently in the search space, and records it as the current individual extremum, and shares the individual extremum with other particles in the entire particle swarm, and finds the optimal individual extremum as the entire particle The current global optimal solution of the swarm, all particles in the particle swarm adjust their speed and position according to the current individual extreme value found by themselves and the current global optimal solution shared by the entire particle swarm. The following animation vividly shows the process of the PSO algorithm:

PSO is initialized as a group of random particles (random solution). Then iteratively find the optimal solution. In each iteration, the particle updates itself by tracking two "extremes" (pbest, gbest). After finding these two optimal values, the particle updates its velocity and position by the following formula.

$$V_{i,t+1} = \omega \times V_{i,t} + c_1 \times rand \times (pbest - X_{i,t}) + c_2 \times rand \times (gbest - X_{i,t}) \quad (1)$$

$$X_{i,t+1} = X_{i,t} + \lambda \times V_{i,t+1} \quad (2)$$

The first part of formula (1) is called [memory item], which represents the influence of the size and direction of the previous speed; the second part of formula (1) is called [self-recognition item], which is the best way to point from the current point to the particle itself. A vector of points, indicating that the particle's action comes from its own experience; the third part of formula (1) is called [group cognition term], which is a vector from the current point to the best point of the population, reflecting the relationship between particles collaboration and knowledge sharing. The particle decides the next movement through its own experience and the best experience of its peers. Based on the above two formulas, the standard form of PSO is formed.

4.Experimental Simulation

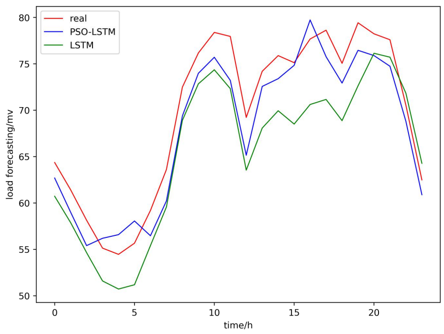
4.1 Data Introduction

This article selects the load value of a certain area in Jiangsu. The data set includes average temperature, maximum temperature, minimum temperature, relative humidity, week type, and load at each time, a total of 29 features.

4.2 IPSO Performs Parameter Optimization

Due to the large number of LSTM parameters, this article intends to use the PSO algorithm to optimize the hyperparameters of LSTM. The optimization parameters include learning rate, number of iterations, and number of hidden layer neurons.

The parameters of the LSTM are set to the data obtained from the PSO iteration and compared with the manually set LSTM parameters. The evaluation standard is RMSE.The prediction result is as shown in Figure (1)



Fig(1)forecast results

Table1. RMSE

Algorithm	RMSE
LSTM	2.655
PSO-LSTM	2.013

The prediction accuracy of the LSTM model optimized by particle swarm optimization is higher than that of the LSTM model with manually set parameters, which verifies the effectiveness and accuracy of the improved LSTM algorithm using PSO-LSTM for short-term power load prediction

5. Conclusion

The PSO optimized LSTM algorithm proposed in this paper has been proved by practical application that it has faster running speed and higher accuracy, and it can be applied to more situations where prediction is required to solve the problem of insufficient short-term prediction accuracy.

References

[1] Kang, CQ., Xia Q., Liu, M., Power system load forecasting [M]. Beijing:China Electric Power Press. 2007.

[2] Shi, JQ., Zhang, JH., Ultra short-term photovoltaic refined forecasting model based on deep learning [J]. Electric Power Construction, 2017, 38(6): 28-35.

[3] Wei, D., Gong, QW., Lai, WQ., et al. Research on internal and external fault diagnosis and fault-selection of transmission line based on convolutional neural network[J]. Proceedings of the CSEE, 2016, 36(S1):21-28.
[4] Li, GY., Li, W., Tian, XL., et al. Short-term electricity load forecasting based on the XGBoost algorithm [J]. Smart Grid, 2017, 7(4):274-285.
[5] Liu, Y., Xu LX., High-performance back propagation neural network algorithm for classification of mass load data [J]. Automation of Electric Power Systems, 2018, 42 (21): 96-103. DOI:10.7500/ AEPS20171215005.
[6] Su, XN., Liu, TQ., Cao, HQ., et al. Amultiple distributed BP neural networks approach for short-term load Forecasting based on Hadoop framework [J].Proceedings of the CSEE, 2017, 37(17): 4966-4973.
[7] Zhu, HB., Cui Y., Xiong, H., Load forecasting based on improved BP neural network [J]. Modern Electronics Technique, 2016,39(20): 64-66.
[8] Wu, XY., He JH., Zhang, Pei, et al. Power system short-term load forecasting based on improved random forest with grey relation projection [J]. Automation of Electric Power Systems. 2015, 39(12): 50-55.
[9] Kong, XY., Zheng, F., E, ZJ., et al. Short-term load forecasting based on deep belief network[J]. Automation of Electric Power Systems, 2018, 42 (5):133-139. DOI: 10.7500/AEPS 20170826002.
[10] Wu, QH., Gao, Jun, Hou, GS., et al. Short-term load forecasting support vector machine algorithm based on multi-source heterogeneous fusion of load factors[J]. Automation of Electric Power Systems, 2016, 40(15): 67-72.
[11] Shen, BX., Using expert system to predict the power load in the area[J]. Electronics Demand Side Management, 2005(2): 49-50.
Project: 1. Shenyang Science and Technology Plan（214030）
2. Open Fund of State Key Laboratory of Building Safety and Environment,
BSBE2021-0