

PAPER • OPEN ACCESS

Application of LSTM-LightGBM Nonlinear Combined Model to Power Load Forecasting

To cite this article: You Zhou *et al* 2022 *J. Phys.: Conf. Ser.* **2294** 012035

View the [article online](#) for updates and enhancements.

You may also like

- [The Application of LightGBM in Microsoft Malware Detection](#)
Qiangjian Pan, Weiliang Tang and Siyue Yao
- [Fast prediction of reservoir permeability based on embedded feature selection and LightGBM using direct logging data](#)
Kaibo Zhou, Yangxiang Hu, Hao Pan et al.
- [Disruption prediction and model analysis using LightGBM on J-TEXT and HL-2A](#)
Y Zhong, W Zheng, Z Y Chen et al.

Application of LSTM-LightGBM Nonlinear Combined Model to Power Load Forecasting

You Zhou^{1a*}, Qing Lin^{1b*}, Di Xiao^{1c}

¹ College of Electrical Engineering and Control Science, Nanjing Tech University, Nanjing, 211816, China

^{a*}202061106013@njtech.edu.cn, ^{b*}1251246918@qq.com, ^cxiaodi_12@sina.com

Abstract—The accurate prediction of power system load is extremely important for the operation of the power market and the safe operation of the power grid. In order to improve the accuracy of short-term load forecasting of power systems, a combination model based on long short term memory network (LSTM) and light gradient boosting machine (LightGBM) is proposed. The experiment first decomposes historical load data by EMD, uses historical weather data and load data decomposed by EMD to establish LSTM prediction model and LightGBM prediction model respectively, and then these two predicted values are linearly combined to obtain the final predicted value. The electrical load data of the 2016 Electrician Mathematical Contest in Modeling is used as an example to verify. The experimental results show that the LSTM-LightGBM combined model has higher forecasting accuracy and application prospects for power load forecasting than traditional load forecasting methods and standard LSTM and LightGBM load forecasting methods.

1. Introduction

Electricity load forecasting is one of the basic tasks in the power industry and has a very important role in the stable operation of the power system as well as in normal production. Accurate prediction of power load is also beneficial to the dispatch of power system, which constantly provides high quality electricity for users, reduces environmental pollution and guarantees air quality.

The traditional methods for power load forecasting analysis mainly include Autoregressive Moving Average Model (ARAM), Vector Autoregression model (VAR), Autoregressive Integrated Moving Average model, (ARIMA) and their variants. In the literature [1], wavelet decomposition and single-branch reconstruction were first applied to the load data, and then the reconstructed sequences were predicted by matching BP and ARIMA models, and finally the prediction results were obtained, and the experimental results showed that the prediction accuracy could be improved. The literature [2] combines the density clustering algorithm with ARIMA, and the experimental results show that the error of the prediction results is within a reasonable range. The traditional method mainly relies on historical electric load data to predict the future electric load values, and does not consider the influence of external factors, such as meteorological conditions, on the predicted objects, so the deviation of the results may be large when the external conditions change too much.

Since the 1990s, machine learning theory has been developed, and many researchers have started to use the theory of machine learning for power load forecasting. Machine learning methods for the analysis of electric loads mainly include single machine learning models such as support vector regression models, random forests, decision trees, BP artificial neural networks, and LSTM networks for forecasting electric loads. The literature [3] used similar data after clustering as a training set



sample to construct a decision tree for random forest regression prediction, and finally the results were corrected using rough set theory, and the average absolute error percentage of the results with the actual load was 2.09%. The literature [4] combined the attention mechanism with LSTM neural network to achieve short-term electricity load forecasting.

Microsoft proposed the Light Gradient Boosting Machine (LightGBM) model framework in 2017, which is an improved model of Gradient Boosting Decision Tree (GBDT) with fast running speed, low memory usage, high accuracy rate, and supports parallel training to handle large-scale data, and is widely used in classification and regression tasks.

Most of the aforementioned methods consider only the linear or nonlinear components of the electric load, ignoring one of the remaining components. Empirical Mode Decomposition (EMD) can decompose the load signal into a residual signal and a finite number of relatively smooth signals, which reduces the non-smoothness of the original electric load data. In this paper, we propose a combined EMD-based electric load prediction model, which decomposes the electric load data by EMD, predicts them by LSTM network and LightGBM algorithm respectively, and then combines them by the error inverse method to determine the weights occupied by the two models. Experiments with real data prove that LSTM-LightGBM has better forecasting effect than other models.

2. Related works

2.1. Long Short Term Memory network

Recurrent Neural Network (RNN) has been widely used in current time series forecasting^[5]. In the conventional BP neural network model, only the layers are fully connected to each other, while the nodes within the layers are not connected to each other. RNNs consist of ordinary BP neural networks with power connections implemented on the implicit layers. The processing of time series by RNN is improved because the implicit layer at the next moment can receive information from the previous moment. The LSTM solves the problem of gradient explosion and gradient disappearance by introducing a gating mechanism^[6].

2.2. LightGBM Model

The decision trees used in the GBDT algorithm can only be regression trees, because each tree of the algorithm learns the results and residuals of all the previous trees^[7]. The GBDT algorithm builds a regression tree of residuals in the direction of the gradient of decreasing residuals, and then adds the results of the individual decision trees together as the final prediction output. LightGBM, an improved model of the GBDT algorithm, improves the speed and accuracy of the GBDT algorithm and reduces memory consumption.

Traditional GBDT algorithms tend to consume computational time in the construction of decision trees, which require finding the optimal segmentation points. The general approach is to sort the feature values and then enumerate all possible feature points. This approach wastes time and requires a lot of memory. LightGBM uses an improved histogram algorithm, which divides the continuous feature values into k intervals and selects the partition points among the k values. Therefore, it has better speed and space efficiency than GBDT algorithm. Meanwhile, decision tree is a weak classifier, and using histogram algorithm will have a regularization effect, which can effectively prevent overfitting.

In terms of reducing training time and memory consumption, the LightGBM algorithm uses a depth-limited leaf-wise growth strategy, where each time from the current leaf node, the leaf with the largest gain is de-split, followed by a cycle until the depth limit is reached^[8]. Compared with traditional methods such as layer-by-layer growth, the depth-limited leaf-wise growth strategy reduces computation and memory consumption and prevents overfitting.

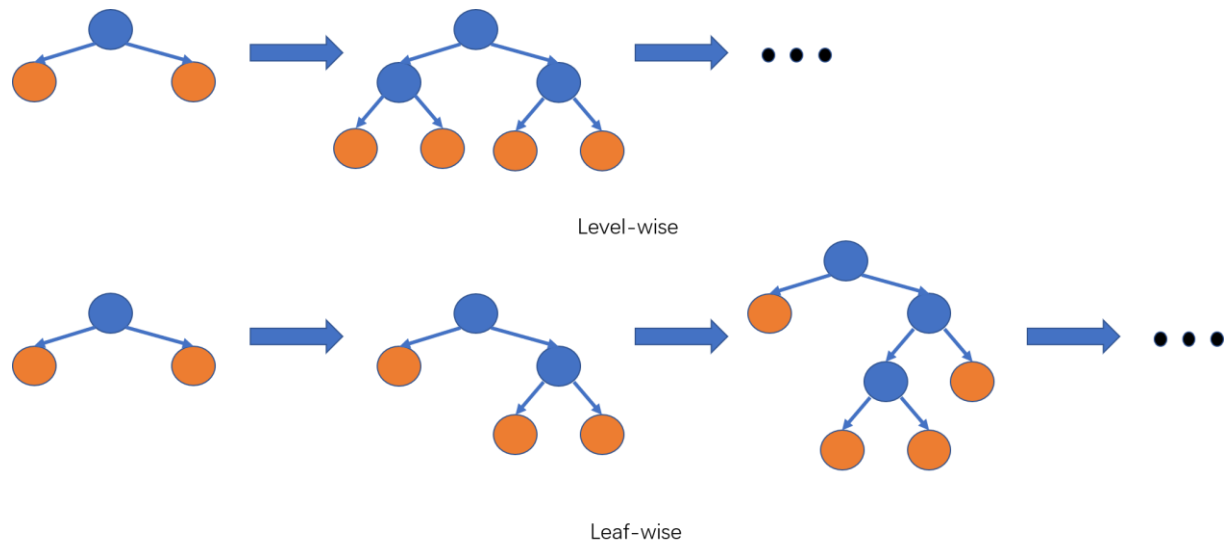


Fig.1 Comparison between Level-wise and Leaf-wise

2.3. Empirical Mode Decomposition

The EMD method can decompose the nonlinear power load data into a finite number of Intrinsic Mode Functions (IMFs) with multiple IMF components and a residual term^[9]. The extreme value points of the IMF components as well as the excess zero points differ by at most one, and its extreme value envelope determined by the local extreme value points, and the minimal value envelope determined by the local minimal value points envelopes are zero. Each IMF component represents fluctuations at different time scales.

2.4. Construction of the Combination Model

Combining a neural network model with a tree model for time series prediction can overcome the shortcomings of a single model and make more accurate predictions. The combined model proposed in this paper first performs EMD decomposition of the original data, and then lets the LSTM model and LightGBM model predict the decomposed IMF components separately to obtain the results. Finally, the error inverse method is used to combine the two prediction results.

Assume that the predicted value of the LSTM prediction model is $P(t)$, the predicted values of the LightGBM prediction model are $L(t)$, then the combined prediction model is:

$$Y_t = \omega_1 P(t) + \omega_2 L(t), t = 1, 2, \dots, n \quad (1)$$

$$\omega_1 = \frac{\varepsilon_L}{\varepsilon_P + \varepsilon_L} \quad (2)$$

$$\omega_2 = \frac{\varepsilon_P}{\varepsilon_P + \varepsilon_L} \quad (3)$$

where, $\varepsilon_P, \varepsilon_L$ are the average relative errors of LSTM and LightGBM, respectively, and Y_t is the prediction value at time t .

3. Experimental Results

3.1. Experimental Data Description

The data selected in this paper were obtained from the standard data set provided by the 9th "China Electrical Engineering Society Cup" National Student Electrical and Mathematical Modeling Competition. The data set contains the electric load values of a region from January 1, 2012 to January 10, 2015 and the weather conditions of that day (the data are counted by day). A total of 974 days of

data from 2012/1/1 to 2014/8/31 was selected as the training set and 132 days of data from 2014/9/1 to 2015/1/10 was selected as the test set.

3.2. Performance Evaluation Metrics

To verify the effectiveness of the combined LSTM-LightGBM model, the root-mean-square error (RMSE), the mean absolute percentage error (MAPE), and the coefficient of determination (R^2) between the predicted and true values are calculated. RMSE and MAPE reflect the error between the predicted and true values, the smaller the better. R^2 reflects the proportion of the explainable part of the predicted value, that is, the explanatory power of the prediction model, the larger the better.

3.3. Predicted results

The IMF function and the residuals obtained by EMD decomposition are used to train the LSTM neural network and the LightGBM model, respectively. the parameters of the LSTM are set as the number of hidden layers is 2, the input layer and both hidden layers are composed of LSTM cells, and the number of neurons in both hidden layers is 128. the output layer is a fully connected layer with 128 inputs and 1 output. the learning rate of the LightGBM is 0.1 and the leaf depth is limited to 20. The learning rate of LightGBM is 0.1 and the depth limit of the leaf is 20. The predicted results are shown in Fig.2 and Table 1.

Table 1 Prediction Results

	RMSE	MAPE	R^2
ARIMA	79143.5	0.086	0.427
LSTM	70162.5	0.064	0.500
LightGBM	65566.5	0.059	0.585
LSTM-LightGBM	64849.5	0.057	0.615

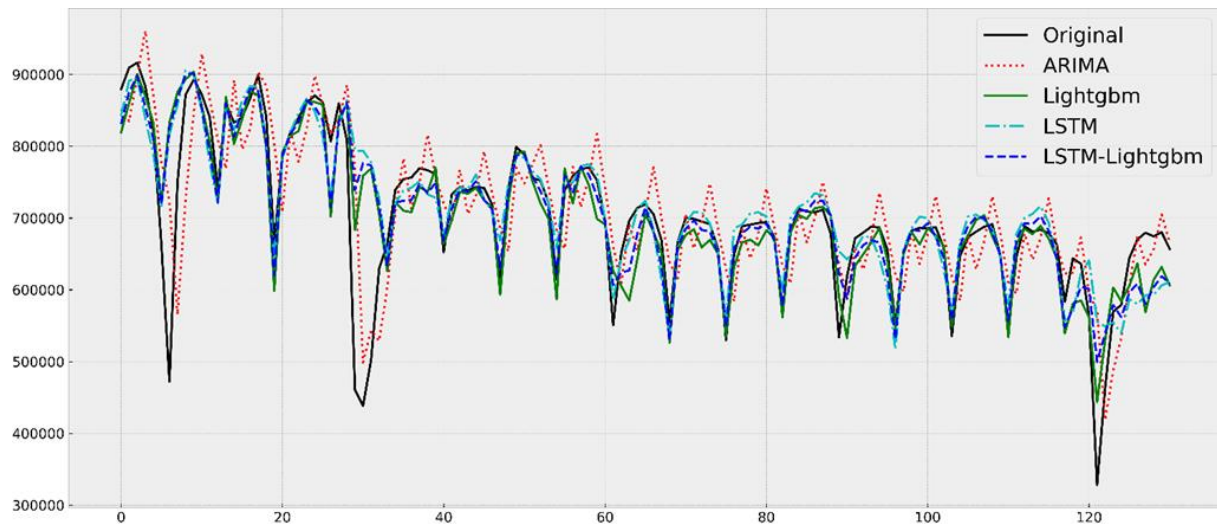


Fig.2 The prediction results of LSTM-LightGBM

4. Conclusion

Because of the strong nonlinearity of the electric load data, it is difficult to make predictions. In this paper, a combined LSTM-LightGBM model is proposed to predict the electric load, and the following conclusions are obtained:

- (1) EMD can decompose the electric load into a set of smoother components, which can improve the prediction accuracy.
- (2) The combination of LSTM and LightGBM using the error inverse method can take into account the advantages of both models and achieve better prediction results.

- (3) Prediction accuracy can be improved by combining specific usage scenarios.

Acknowledgments

This work was supported by Postgraduate Research & Practice Innovation Program of Jiangsu Province (No. KYCX21_1137).

References

- [1] Zhu, L., Wang, X., Ma, J., Chen, Q., & Qi, X. (2020). Short-term load forecast of integrated energy system based on wavelet packet decomposition and recurrent neural network. *Electr. Power Constr*, 41, 131-138.
- [2] Chen J Y, Ding J Y, Tian S M, et al. (2018). An improved density peaks clustering algorithm for power load profiles cluster-ing analysis. *Power System Protection and Control*, 46(20): 85-93.
- [3] Guo, W., Che, L., Shahidehpour, M., & Wan, X. (2021). Machine-Learning based methods in short-term load forecasting. *The Electricity Journal*, 34(1): 106884.
- [4] Wei J, Zhao H T, Liu D N, et al. (2021). Short-term power load forecasting method by attention-based CNN-LSTM. *Journal of North China Electric Power University(Natural Science Edition)*, 48(01), 42-47.
- [5] Liu B, Wang M L, Li Y, et al. (2021). Deep learning for spa-tio-temporal sequence forecasting[J]. *Journal of Beijing University of Technology*, 47(08): 925-941.
- [6] He, K., Ji, L., Wu, C. W. D., & Tso, K. F. G. (2021). Using SARIMA–CNN–LSTM approach to forecast daily tourism demand. *Journal of Hospitality and Tourism Management*, 49, 25-33.
- [7] Zhang X, Ding Y F, Wang G. (2021) Electric load forecasting based on pre-training GRU-LightGBM. *Computer Systems & Applications*, 30(08): 288-292.
- [8] Zhang, Y., Zhu, C., & Wang, Q. (2020). LightGBM-based model for metro passenger volume forecasting. *IET Intelligent Transport Systems*, 14(13), 1815-1823.
- [9] Yao H G, Mu N G. (2021). Prediction of Financial Time Series by EMD-LSTM Model. *Computer Engineering and Appli-cations*, 57(05): 239-244.
- [10] Zhang X, Ding Y F, Wang G. (2021) Electric load forecasting based on pre-training GRU-LightGBM. *Computer Systems & Applications*, 30(08): 288-292.