Forest Chemicals Review
www.forestchemicalsreview.com
ISSN: 1520-0191
July-August 2022 Page No. 620-634
Article History: Received: 30 March 2022, Revised: 8 April 2022, Accepted: 15 April 2022, Publication: 30 April 2022

Short-Term Load Forecasting Based on LightGBM Parallel Ensemble Method

Chenyang Feng*
College of Electrical Engineering & New Energy, China Three Gorges University, Yichang, Hubei, China
*Corresponding Author.

Abstract:

High-accuracy short-term load forecasting can effectively guide the start-up and shutdown plans of coal generating units and reduce the waste of coal resources. To improve the short-term load forecasting accuracy, a LightGBM parallel Ensemble paradigm-based short-term load forecasting method for electric power is proposed. Firstly, the relevant input features of the load to be predicted are screened using the Spearman coefficients; then LightGBM, an efficient serial Ensemble method, is selected as the base learner, its parameters are optimized using the Ant lion optimizer, and cross-validation is used to ensure the diversity of different LightGBM base learners. Finally, the Bagging ensemble method is used as a parallel ensemble paradigm to achieve forecasting with the optimal weighting method combined with the base learners. Experimental analysis using Australian electricity consumption data as an arithmetic example shows that the LightGBM-based parallel ensemble method combines the advantages of serial and parallel ensemble methods to simultaneously reduce forecast bias and variance and improve the quality and speed of short-term electricity load forecasting.

Keywords*: LightGBM, Bagging, Ensemble method, Spearman coefficient, Electricity short-term load forecasting.*

I. INTRODUCTION

Coal occupies an important energy position as the main source of electricity supply worldwide. High-precision electric short-term load forecasting can guide the economic start-up and shutdown plan of coal units and reduce the waste of coal resources. In the context of a large number of new energy sources connected to the power system, higher requirements are placed on the accuracy and stability of electrical load forecasting. Short-term electric load forecasting refers to forecasting the load for the next day to several days to guide short-term system scheduling and day-to-day unit start-up and shutdown plans [1]. The short-term electrical load is characterized by high randomness and many influencing factors. Machine learning is a technique to find effective knowledge from big data, and its use for electrical load forecasting mainly includes two types: single learning and ensemble method. Single learning theory is mature, simple and has many schools of thought, and several methods including symbolism, connectionism and statistical learning have been studied in the field of short-term load forecasting for electricity [2-5]. Single learning mechanistically carries out analysis and prediction only from a single perspective, while Ensemble

Forest Chemicals Review
www.forestchemicalsreview.com
ISSN: 1520-0191
July-August 2022 Page No. 620-634
Article History: Received: 30 March 2022, Revised: 8 April 2022, Accepted: 15 April 2022, Publication: 30 April 2022

learning makes full use of the characteristics of multiple base learners to complement each other's strengths [6]. Ensemble learning can be classified into serial and parallel ensemble paradigms according to the base learner generation method. The basic idea of the serial ensemble method is to use the correlation between the base learners, and the literature [7] proposes to construct a prediction model based on two-layer Extreme Gradient Boosting(XGBoost), with the first layer screening the effective feature set and the second layer used for load prediction, which can effectively improve the prediction accuracy. Literature [8] combined the CNN-LSTM network and XGBoost algorithm in a weighted form to further improve the prediction accuracy; literature [9] used XGBoost combined with K-mean clustering to effectively avoid the loss of accuracy due to feature redundancy; literature [10] combined serially Ensemble XGBoost and LSTM with the error inverse method to obtain better prediction accuracy. The basic idea of the parallel ensemble learning method is to use the independence between the base learners, and the literature [11] Ensemble the Gate Recurrent Unit (GRU) in parallel with Bagging to improve the model resistance to interference, which is more stable than the general model; the literature [12] used the random forest parallel ensemble algorithm to predict the low-frequency components to achieve multi-frequency domain prediction to improve the prediction accuracy. The literature [13] uses Bagging Ensemble Gradient Boosting Decision Tree (GBDT) for the prediction of fuzzy processed load data with improved generalization performance and accuracy. Serial ensemble and parallel ensemble, when combined organically, will bring out the advantages of both and help to further improve the prediction quality. In this paper, we propose a LightGBM parallel ensemble learning method based on the combination of serial and parallel ensembles to improve the prediction accuracy and stability of the model, and conduct a comparative validation analysis using Australian load data.

II. MATERIALS AND METHODS

The electricity load data used for the experiment is the statewide residential electricity load public dataset of Olmer Grid, New South Wales, Australia, from 2006 to 2010, with a load sampling period of 0.5 hours and 48 points collected daily, combined with the real-time weather temperature (℃), climate humidity (hPa), real-time electricity price (AUD) corresponding to the sampling points, and historical load values for the previous 48 moments, totaling 36433 samples [14]. Among them, data from January 1, 2008 to December 31, 2009 were used for training and validation of the load prediction model, and January 1, 2010 to January 7, 2010 were used to test the validity of the prediction model. An overview of the experimental data is shown in TABLE I.

TABLE I. Overview of experimental data

Example	**Data for New South Wales, Australia**	
Sample	Training set + validation set	Test set
Sampling time	2008/1/1~2009/12/31	2010/1/1~2010/1/7 Daily data
Sample size	0.5 hours	

Forest Chemicals Review
www.forestchemicalsreview.com
ISSN: 1520-0191
July-August 2022 Page No. 620-634
Article History: Received: 30 March 2022, Revised: 8 April 2022, Accepted: 15 April 2022, Publication: 30 April 2022

Amount of data	35089	1344
Included Features	Weather temperature (C), climate humidity (hPa), real-time electricity price (AUD), historical load (MW)	

2.1 Ensemble Learning

Ensemble learning improves algorithm generalization by combining several identical or different basic learners with some strategy to obtain better learning results than a single learner [15]. Ensemble learning can be classified into serial and parallel ensembles based on the generation form of the base learners. In a serial ensemble, each base learner is formed sequentially, and there is strong dependency between base learners, represented by AdaBoost (Adaptive Boosting) algorithm; in the parallel ensemble, each base learner is formed in parallel, and the difference between base learners is emphasized, represented by Bagging algorithm. Load prediction is a regression problem, and the following is an example of the regression problem.

2.2 Bagging Algorithm

The Bagging algorithm belongs to the parallel ensemble paradigm of Ensemble learning, which combines several base learners to form a strong learner. The Bagging algorithm forms several subsets of training data by autonomously sampling (Bootstrap Sampling) the original training set with arbitrary put-backs. Based on these different training datasets several basic learners are trained and then some strategy is used to combine these base learners to obtain the output [16]. The principle of Bagging algorithm is shown in Fig 1.

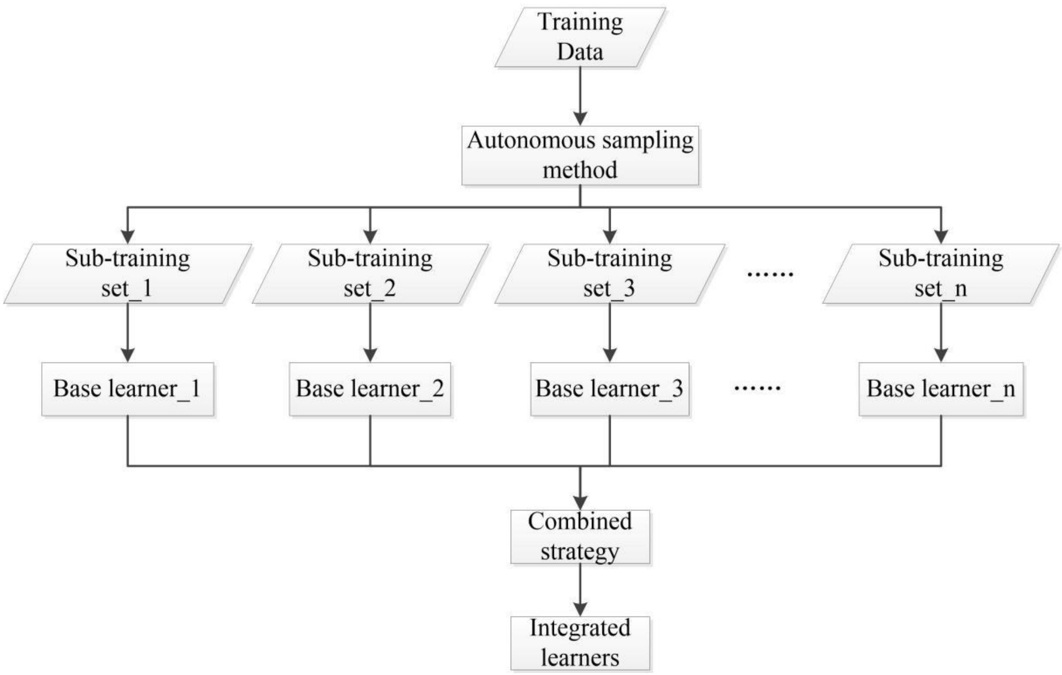

Fig 1: the principle of Bagging algorithm

Forest Chemicals Review
www.forestchemicalsreview.com
ISSN: 1520-0191
July-August 2022 Page No. 620-634
Article History: Received: 30 March 2022, Revised: 8 April 2022, Accepted: 15 April 2022, Publication: 30 April 2022

2.3 LightGBM Algorithm

The AdaBoost algorithm is the basic algorithm of serial Ensemble learning, which improves the learning effect by adaptive adjustment of the base learner weights, but abnormal samples given higher learning weights to predict the problem for example will show a decrease in prediction accuracy. The GBDT (Gradient Boosting Decision Tree) algorithm uses a decision tree as the base learner based on the AdaBoost algorithm and optimizes the model parameters by the gradient descent method to improve the learning stability, but its base learner is single and only uses the first-order derivative optimization model. XGBoost (Extreme Gradient Boosting Machine) improves the GBDT algorithm by adding a regular term to prevent overfitting and using second-order derivative optimization for faster gradient descent. To improve the computational time consumption and accuracy of large-scale data processing, Microsoft introduced the LightGBM (Light Gradient Boosting Machine) algorithm in 2017, which optimizes the training sample input based on the XGBoost algorithm and improves the leaf node splitting strategy to enhance the prediction accuracy [17].

The base learners of the LightGBM algorithm are decision trees, and each base learner is generated serially and combines T decision trees according to the form of Eq. (1), which represents the set space of all decision trees. The goal of forming a new decision tree is to minimize the loss function including regression error and complexity.

The main improvements of the LightGBM algorithm based on the XGBoost algorithm are gradient one-sided sampling, leaf growth strategy with depth restriction, and mutually exclusive feature bundling. Gradient one-sided sampling means that when forming the decision tree training set, samples with large gradient amplitude are retained and samples with small gradient amplitude are sampled by a prescribed ratio based on the gradient value of the loss function of the samples. The leaf growth strategy with depth restriction differs from the layer-by-layer decision tree growth strategy in that only the leaves with the largest splitting gain are split each time, reducing the complexity of decision tree generation.

$$f_T(X) = \sum_{t}^{T} a_t f_t(X), f_t \epsilon F \tag{1}$$

The principle of the LightGBM algorithm is shown in Figure 2. The training set of each decision tree is formed by gradient one-sided sampling, and the decision trees are grown using a leaf growth strategy with depth restrictions. Each decision tree is generated serially and sequentially, and finally, the final learner is formed using the combination strategy of the weighting method.

Forest Chemicals Review
www.forestchemicalsreview.com
ISSN: 1520-0191
July-August 2022 Page No. 620-634
Article History: Received: 30 March 2022, Revised: 8 April 2022, Accepted: 15 April 2022, Publication: 30 April 2022

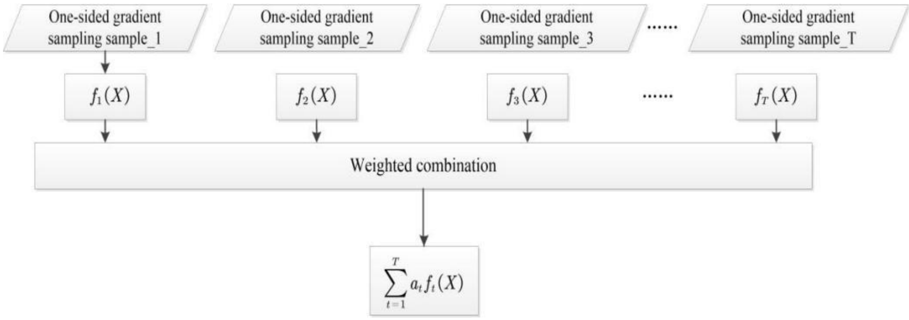

Fig 2: the principle of LightGBM algorithm

2.4 LightGBM Parallel Ensemble Learning for Electricity Short-Term Load Forecasting Process

In this paper, we choose the Bagging-LightGBM model based on a parallel-style ensemble for short-term electric load forecasting. The process is shown in Figure 3, and the implementation steps are as follows.

Step 1: Analyze the factors related to the load to be forecasted using the Spearman coefficient method to determine the input characteristic quantities for short-term load forecasting of electricity. The main factors related to the electricity consumption of electricity.

Forest Chemicals Review
www.forestchemicalsreview.com
ISSN: 1520-0191
July-August 2022 Page No. 620-634
Article History: Received: 30 March 2022, Revised: 8 April 2022, Accepted: 15 April 2022, Publication: 30 April 2022

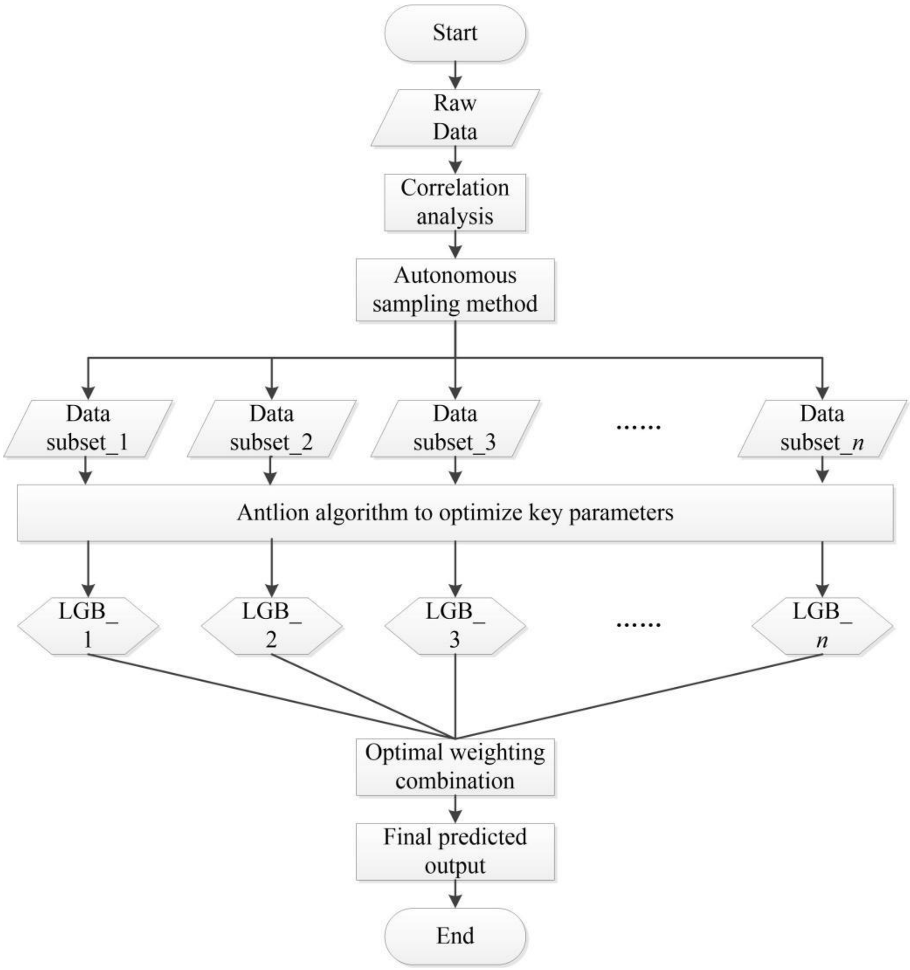

Fig3: LightGBM Parallel Ensemble Learning for Electricity Short Term Load Forecasting Process

Consumers include temperature, humidity, and historical load, which are the basis for carrying out such regression tasks as forecasting.

Step 2: The initial training set is sampled with put-back to obtain a subset of training data and generate a base learner based on the LightGBM algorithm.

To ensure the diversity of such base learners of the LightGBM algorithm, the training set of the prediction model is formed by sampling with put-back sampling for the data set containing load as well as weather.

Step 3: Optimize the LightGBM key hyperparameters using the Antlion algorithm and cross-validate them with the validation set [18].

Step 4: The different base learners are combined on the test set using the Bagging parallel ensemble method to output the final prediction results according to the optimal weighted combination method

Forest Chemicals Review
www.forestchemicalsreview.com
ISSN: 1520-0191
July-August 2022 Page No. 620-634
Article History: Received: 30 March 2022, Revised: 8 April 2022, Accepted: 15 April 2022, Publication: 30 April 2022

described in the literature [19].

The steps given above are the steps of single-point forecasting of the short-term load of electric power, while multi-point forecasting of the future load is generally required. In this paper, we adopt an iterative forecasting strategy, i.e., we first use historical data to obtain the relevant characteristic variables of the forecasted load, and use the multi-point forecasts obtained from previous forecasts as historical values when forecasting the nth point in the future, so as to obtain multi-point forecasts iteratively.

In this paper, the core parameters of the LightGBM algorithm are optimized using the ant-lion algorithm, which achieves global optimization by simulating a colony of ant lions hunting ants [18]. The initial training set is again divided into a new training set and a validation set, and the optimized hyperparameter set is obtained by the ant-lion algorithm in a cross-validation manner as shown in TABLE II. In this paper, the number of LightGBM-based learners is chosen to be 10 for the example comparison.

TABLE II. LightGBM parameter settings

Parameter Name	Parameter Role	Setting Value
Learning rate	Obtain a more stable model	0.05
Maximum depth	Control overfitting	6
Maximum number of leaf nodes	An increase can improve accuracy but also overfitting	240
Random sampling sample proportion	Improve generalization ability	0.8
Feature sampling rate	Accelerated training to control overfitting	1.0
Regularization parameter L1	Prevent overfitting	0.5
Regularization parameter	Prevent overfitting	0.5

2.5 LightGBM Parallel Ensemble Learning Electricity Short-Term Load Forecast Evaluation Method

In this paper, we use the evaluation indexes commonly used for short-term load forecasting: Root Mean Square Error, Mean Absolute Error and Mean Relative Error, where Root Mean Square Error and Mean Absolute Error are used for comprehensive error analysis and Mean Relative Error reflects whether there is systematic bias in the model, both of which are used to measure the accuracy of load forecasting, as defined by Eq.

$$E_{RMSE}=\sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i-\hat{y}_i)^2} \tag{2}$$

Forest Chemicals Review
www.forestchemicalsreview.com
ISSN: 1520-0191
July-August 2022 Page No. 620-634
Article History: Received: 30 March 2022, Revised: 8 April 2022, Accepted: 15 April 2022, Publication: 30 April 2022

$$E_{MRE}=\frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i-\hat{y}_i}{\hat{y}_i}\right|\times 100\% \tag{3}$$

$$E_{MAE}=\frac{1}{n}\sum_{i=1}^{n}|\hat{y}_i-y_i| \tag{4}$$

Where, n represents the number of test points, the true value of the load, and the model prediction value.

In addition, the stability of the prediction method is tested by examining the degree of error discretization, i.e., standard deviation, at each prediction point with the following equation.

$$\sigma_{RMSE}=\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2} \tag{5}$$

Where, x_i represents the deviation value of the model at the ith test point and $\bar{x}$ represents the mean of the Root Mean Square Error of the n test points.

2.6 Historical Load Characteristics Screening

The short-term load has periodic characteristics between adjacent time periods, and the periodic characteristics of load change can be effectively explored by analyzing the correlation between the current load and the load of adjacent time periods. The input historical load characteristics selection is divided into preliminary screening and final screening, and the preliminary screening is firstly completed by using the load change rate, and then the final screening is completed by using the Spearman coefficient.

The historical load is used for the load data within 3 days before the moment to be predicted, and the initial screening of the load change rate is used. If $Q_i(t)$ represents the load at time t of day i, the historical load with a smaller rate of change between the historical load sequence and the current load is $Q_i(t-1)$, $Q_i(t-2)$, $Q_{i-1}(t+1)$, $Q_{i-1}(t+2)$, and the maximum advance prediction time is 24 hours, as obtained by statistics.

The Spearman coefficient method of analyzing correlation can avoid the interference of data quality, has low requirements for data distribution, and is suitable for the practical needs of short-term load forecasting for electricity. To further analyze the historical load correlation, the Spearman coefficients of the forecast load and the preliminary screening of the historical load are calculated as in TABLE III.

Forest Chemicals Review
www.forestchemicalsreview.com
ISSN: 1520-0191
July-August 2022 Page No. 620-634
Article History: Received: 30 March 2022, Revised: 8 April 2022, Accepted: 15 April 2022, Publication: 30 April 2022

TABLE III. Spearman's coefficient for historical and forecast loads

Historical load Forecast load	$Q_{i-1}(t+1)$	$Q_{i-1}(t+2)$	$Q_i(t-2)$	$Q_i(t-1)$
$Q_i(t)$	0.86	0.79	0.94	0.97

Experimentally comparing the effects of different historical load inputs on the prediction accuracy, Model I load input feature contains only $Q_i(t-1)$, Model II load input feature contains $Q_i(t-1)$, $Q_i(t-2)$, Model III load input feature contains $Q_i(t-1)$,$Q_i(t-2)$, $Q_{i-1}(t+2)$, and Model IV load input feature contains $Q_i(t-1)$,$Q_i(t-2)$ $Q_{i-1}(t+2)$,$Q_{i-1}(t+1)$. Short-term load forecasting was carried out for 2010, and the prediction comparison results are shown in TABLE IV.

As can be seen from TABLE IV, by selecting $Q_i(t-1)$, $Q_i(t-2)$,$Q_{i-1}(t+2)$ as the load input feature, the model can fully learn the valid information and implied laws in the data, and the prediction accuracy indicators e_{RMSE} and e_{MRE} are substantially improved. Given that meteorological information such as temperature and humidity are closely related to load, the input variables for the prediction regression task in this paper include six variables: $Q_i(t-1)$, $Q_i(t-2)$,$Q_{i-1}(t+2)$, temperature at the prediction moment, humidity at the prediction moment, and real-time electricity price.

TABLE IV. Prediction errors for different combinations of load characteristics

Models	The Year 2010	
	E_{RMSE}/MW	E_{MAE}/MW
Model I	568.99	448.65
Model II	192.17	135.21
Model III	146.13	90.24
Model IV	217.82	144.47

III. RESULTS AND DISCUSSION

3.1 Comparison Between Working Day and Rest Day Forecasts

The similarity between the normal daily electricity load curves in the experimental area is high, and the load curves of different rest days are also close, but the differences between working days and rest days are more distinct. To further illustrate the influence of short-term load forecasting accuracy by day type,

Forest Chemicals Review
www.forestchemicalsreview.com
ISSN: 1520-0191
July-August 2022 Page No. 620-634
Article History: Received: 30 March 2022, Revised: 8 April 2022, Accepted: 15 April 2022, Publication: 30 April 2022

TABLE V shows the load forecasting of LightGBM parallel Ensemble learning for seven consecutive days of a week in 2010. It can be seen that the forecasting accuracy of LightGBM parallel Ensemble learning from Monday to Friday is higher than that of Saturday and Sunday, mainly because there are fewer data on rest days and the forecasting model does not fit the data on rest days as well as weekdays. The prediction process is smoother for both working days and rest days, and the proposed method meets the requirements of practical applications.

TABLE V. Comparison of working day and rest day forecasts

Date		LightGBM-based parallel Ensemble learning load prediction model	
		E_{RMSE}/MW	E_{MAE}/MW
Work Day	Monday	111.52	91.89
	Tuesday	126.98	99.42
	Wednesday	134.26	103.45
	Thursday	141.12	105.84
	Friday	139.25	104.89
Rest Day	Saturday	186.99	133.34
	Sunday	198.05	141.59

3.2 Comparison of Prediction with Multiple Learners

In order to test the prediction effect of the model in this paper, it is compared with the statistical learning-based SVR (Support Vector Regression) algorithm, the symbolism-based KNN (K-Nearest Neighbor) algorithm, the connectionism-based BP (Back Propagation) algorithm and the serial Ensemble XGBoost algorithm, parallel Ensemble RF (Random Forest) algorithm, and string-parallel Ensemble Bagging-XGBoost algorithm are compared for accuracy evaluation metrics and model stability metrics. In order to test the learning effect of the model and increase the credibility of the comparison between different models, the training set was set as samples from 2008 to 2009, and the test set was selected from one week of data in January, April, July and October, a representative month of the four seasons in 2010.

3.2.1 Prediction comparison with single learner

Single learners have a simple structure and tend to be less effective in prediction than Ensemble learners. The prediction accuracy of the KNN in TABLE VI is much better for weekdays in April and July

Forest Chemicals Review
www.forestchemicalsreview.com
ISSN: 1520-0191
July-August 2022 Page No. 620-634
Article History: Received: 30 March 2022, Revised: 8 April 2022, Accepted: 15 April 2022, Publication: 30 April 2022

compared to January and October, and its prediction performance in different test sets varies widely, showing a sawtooth variation in Fig 4, indicating that the method is poorly fitted to large data sets. The BP neural network still shows large ups and downs on accuracy tests over multiple months. The variance of the other single learners is also much higher than the model in this paper. From Fig 4, it can be seen that the larger error of the single learner occurs in the midday period, which is due to the rapid temperature change in the midday period in Australia, while the LightGBM parallel Ensemble learning fits significantly better than the single learner at this stage.

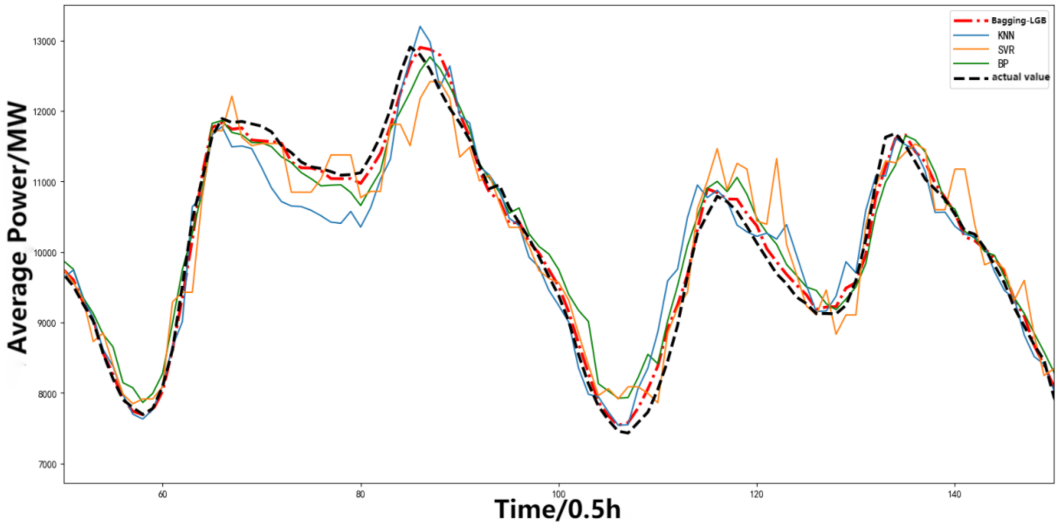

Fig 4: Comparison with single learner load prediction curve

3.2.2 Comparison of prediction with Ensemble learners

In order to investigate the advantages of this model over other Ensemble learners, LightGBM, XGBoost, RF (Random Forest) and parallel Ensemble learning with LightGBM are selected as reference comparisons. Again, the test experiments were done on the working day data of January, April, July and October 2010 to compare the prediction accuracy, model robustness and prediction elapsed time, and the error comparison results are shown in TABLE VI.

Forest Chemicals Review
www.forestchemicalsreview.com
ISSN: 1520-0191
July-August 2022 Page No. 620-634
Article History: Received: 30 March 2022, Revised: 8 April 2022, Accepted: 15 April 2022, Publication: 30 April 2022

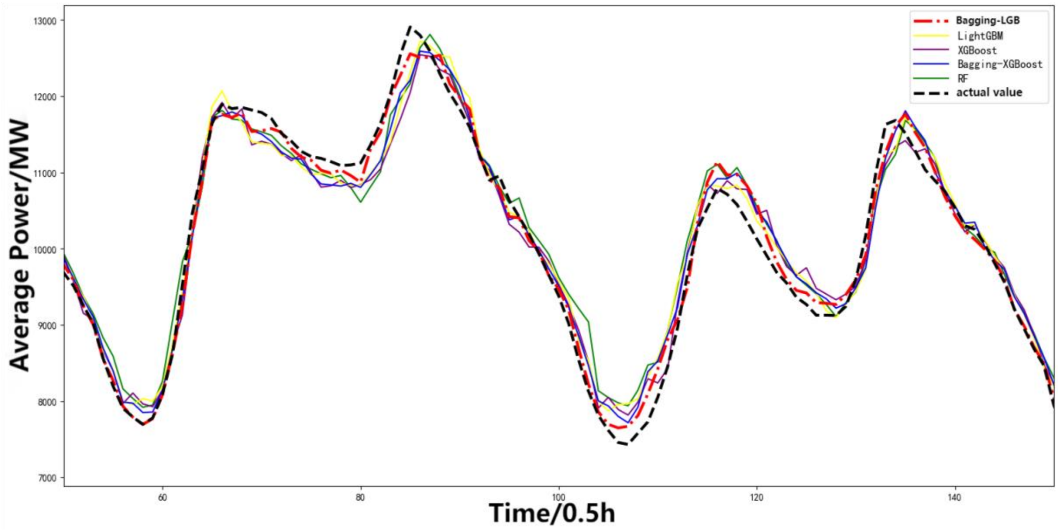

Fig 5: Comparison of load prediction curve with Ensemble learner

Ensemble learners can sufficiently reduce the prediction bias and improve the model generalization ability compared to single learners. However, the joint serial-parallel ensemble is clearly more advantageous than the serial ensemble or parallel ensemble. Linking TABLE VI and Fig 5, it can be seen that LightGBM and XGBoost, like gradient boosting class algorithms and Ensemble in serial paradigm, significantly improve the prediction accuracy compared to RF in parallel ensemble paradigm in January and April time periods, but slightly less than RF in model stability metrics, and if LightGBM does not use parallel optimization computation, its own stability measures with RF will further pull apart. Therefore, LightGBM parallel ensemble learning combines the advantages of serial ensemble and parallel ensemble, and improves the prediction speed by nearly 6 times and the prediction accuracy by about 2% compared with Bagging-XGBoost of the same structure, fully demonstrating the advantages of LightGBM as a parallel load prediction model-based learner.

TABLE VI. Comparison of predictions with multiple learners

Time/evaluation metrics		**Models in this paper**	**Single learner**			**Ensemble Learner**			
			KNN	SVR	BP	LightGBM	XGBoost	Bagging--XGBoost	RF
January	E_{RMSE}/MW	**138.13**	382.36	193.45	352.71	165.69	173.82	161.65	197.31
	$E_{MRE}/\%$	**1.13**	2.21	1.34	2.18	1.17	1.28	1.16	1.35
	σ_{RMSE}/MW	**87.26**	335.45	193.45	352.71	90.46	121.75	88.46	84.45

Forest Chemicals Review
www.forestchemicalsreview.com
ISSN: 1520-0191
July-August 2022 Page No. 620-634
Article History: Received: 30 March 2022, Revised: 8 April 2022, Accepted: 15 April 2022, Publication: 30 April 2022

	T/s	**0.72**	0.45	10.90	5.11	0.12	0.83	5.61	5.73
April	E_{RMSE}/MW	**140.35**	350.01	189.41	345.96	165.27	180.23	154.25	164.16
	$E_{MRE}/\%$	**1.14**	2.17	1.33	2.16	1.17	1.29	1.15	1.16
	σ_{RMSE}/MW	**60.47**	304.67	198.84	324.26	80.46	134.5	65.84	78.59
	T/s	**0.75**	0.47	11.5	5.14	0.47	11.5	5.74	5.14
July	E_{RMSE}/MW	**109.39**	326.34	177.64	366.07	150.99	151.90	150.81	168.99
	$E_{MRE}/\%$	**0.98**	2.11	1.29	2.18	1.16	1.17	1.16	1.18
	σ_{RMSE}/MW	**89.34**	310.45	184.24	312.17	135.87	100.28	91.25	115.67
	T/s	**0.71**	0.51	12.3	5.34	0.51	12.3	5.71	5.34
October	E_{RMSE}/MW	**133.29**	364.05	180.64	360.28	166.17	174.34	168.59	196.30
	$E_{MRE}/\%$	**1.13**	2.17	1.32	2.16	1.17	1.28	1.18	1.35
	σ_{RMSE}/MW	**57.28**	350.46	196.51	289.89	153.46	120.34	89.46	143.71
	T/s	**0.77**	0.49	11.2	5.21	0.15	0.85	5.76	5.80

IV. CONCLUSION

In this paper, the LightGBM parallel Ensemble learning model is used for short-term load forecasting of electricity, and the effective historical load input features are screened a priori using the Spearman correlation coefficient, and the Antlion algorithm globally optimizes the search for LightGBM key parameters for differentiated forecasting analysis of weekdays and rest days, comparing and interpreting the forecasting performance of this paper's model with different models including common single learners and classical Ensemble learners in four seasons, and the results show that.

In terms of model evaluation metrics, the serial, parallel joint ensemble framework based on LightGBM parallel-style ensemble learning for power short-term load forecasting outperforms a single learner in terms of accuracy and stability. In addition, it is more stable than a simple serial Ensemble learner and has higher prediction accuracy than a simple parallel Ensemble learner, combining the advantages of high accuracy of serial Ensemble learning and high stability of parallel Ensemble learning.

In terms of prediction speed, although the LightGBM-based parallel-style Ensemble learning model is more time-consuming than KNN and LightGBM, the model in this paper still has superiority considering that the primary goal of power load forecasting applications is to meet higher prediction accuracy.

It is shown that LightGBM is more suitable as a base learner for parallel ensemble applications for the prediction of massive load data in terms of accuracy, stability and time consumption metrics by comparing the performance of Bagging-XGBoost, which has the same construction as the LightGBM parallel-style Ensemble learning model.

Considering the higher accuracy requirement for short-term load forecasting, future research will be conducted around the combination strategy of LightGBM parallel-type ensemble to further improve the forecasting accuracy.

ACKNOWLEDGEMENTS

This research was supported by the National Natural Science Foundation of China&Youth Program (Grant No. 51807109).

REFERENCES

[1] Kang C Q, Xia Q, Liu M. Power System Load Forecasting (2nd edition). Beijing: Electric Power Publishing House, 2017:17-19

[2] Li X, Li C B, Cao Y J, et al. A New Decoupled Decision Tree Algorithm for Short-Term Load Forecasting Proceedings of the CSU-EPSA, 2013, 25(03): 13-19.

[3] Peng W, Wang J R, Yin Q S. A Short-term Load Forecasting Model Based on Attention-LSTM in the Electricity Market. Power System Technology. 2019, 43(05): 1745-1751.

[4] Yang Y Y, Wang J, Peng H, et al Application of Improved Extreme Learning Machine Based on Membrane Clustering in Short-term Load Forecasting Water Resources and Power, 37(12): 181-184.

[5] Zhao P, Dai Y M. SVM Power Load Forecasting Based on Real-time Electricity Price and Weighted Grey Relational Projection Power System Technology. 2020, 44(04): 1325-1332.

[6] Zhou, C.-H. Machine learning. Beijing: Tsinghua University Press, 2016: 33-35, 171-173.

[7] Sun C, Lv Q, Zhu S T et al. Ultra-short-term Power Load Forecasting Considering Multi-feature Effects Based on Double-layer XGBoost algorithm High Voltage Engineering, 2021,47(08):2885-2898.

[8] Zhuang J Y, Yang G H.et al. CNN-LSTM-XGBoost Short-term Power Load Forecasting Method Based on Multi-model fusion. Electric Power, 2021, 54(05): 46-55.

[9] Jiang J, Liu H Q, Li H.et al. Prediction Method of Peak Load of Distribution Network Line Based on XGBoost Power System Protection and Control: 2021, 49(16): 119-127.

[10] Chen Z Y, Liu J B, Li C et al. Ultra-short-term power load forecasting based on combined model of LSTM and XGBoost Power System Technology, 2020,44(02):614-620.

Forest Chemicals Review
www.forestchemicalsreview.com
ISSN: 1520-0191
July-August 2022 Page No. 620-634
Article History: Received: 30 March 2022, Revised: 8 April 2022, Accepted: 15 April 2022, Publication: 30 April 2022

[11] Wang K, Zhang Z S, Han A Y, et al. Short-term load forecasting of bidirectional GRU ensemble neural network based on bagging Proceedings of the CSU-EPSA1-8[2021-10-04]. https://doi.org/10.19635/j.cnki.csu-epsa.000591.
[12] Ma Y, Zhang Q, Li G Let al. Research and Analysis of Short-Term Load Forecasting Based on Frequency Domain Decomposition Electric Power, 2020, 53(04):114-121.
[13] Bi Y F, Han A Y, Zhang Z S.et al Research on Short-Term Load Forecasting Model Based on Fuzzy Bagging-GBDT. Proceedings of the CSU-EPSA, 2019, 31(07):51-56.
[14] Australian Energy Market Operator.AEMO's first national gas forecasting report. [2014-11-17]. http://www.aemo.com.au.
[15] Zhou C.-H. Ensemble learning fundamentals and algorithms. Beijing: Electronic Industry Press, 2020:25-26
[16] Breiman, Leo. Bagging predictors. Machine learning, 24. 2 (1996): 123-140.
[17] Ke G, Meng Q, Finley T, et al. Lightgbm: A highly efficient gradient boosting decision tree. Advances in neural information processing systems, 2017, 30: 3146-3154.
[18] Mirjalili S. The ant lion optimizer. Advances in engineering software, 2015, 83: 80-98.
[19] Chen W N, Hu Z J, Yue J P et al Short-term load forecasting based on long short-term memory network and LightGBM combined model Automation of Electric Power Systems,2021,45(04):91-97.