The Research of Short-term Electric Load Forecasting based on Machine Learning Algorithm

Yaoying Wang*, Shudong Sun, Zhiqiang Cai

School of Mechanical Engineering Northwestern Polytechnical University Xi'an, China
e-mail: wangyaoying@mail.nwpu.edu.cn*

Abstract. The ability of electric load forecasting become the key to measuring electric planning and dispatching, and improving the accuracy of electric load forecasting has become one hot spot topic of scholars in recent years. The traditional electric load forecasting algorithm mainly include statistical learning algorithm. The electric load forecasting algorithm based on traditional forecasting algorithm, machine learning algorithm and neural network algorithm greatly improve the convergence approximation effect and accuracy. In recent years, the electric load forecasting algorithms based on the deep learning algorithm has greatly reduced the error and improved the measurement accuracy and robustness, such as RNN, GRU, LSTM, DBN, TCN, and so on, as well as the combination algorithm based on deep learning. This paper introduces the classical algorithm of deep learning in electric load forecasting. The main purpose is to explore the error and fitting degree of algorithms established by different algorithms, and provide reference for the selection of forecasting algorithms for various types of electric load forecasting.

Keywords: electric load forecasting;machine learning; deep learning; forecasting accuracy.

1. Introduction

The electric load forecasting is based on the historical electric load data in a electric supply region to forecast the electric demanding and supplying for a period of time in the future[1,2], which is used to support electric planning and dispatching operations, reduce grid line loss, improve electric quality, and meet electric supply demand. With the continuous development of new electric systems, the requirements for the accuracy of electric load forecasting are gradually improved[1]. Relevant research shows that the annual operation cost of the electric system will increase by tens of millions of yuan for every 1% increase in the error of electric load forecasting[2]. Therefore, it is urgent to improve the accuracy and robustness of electric load forecasting.

This paper reviews the electric load forecasting algorithm based on statistical learning algorithm, machine learning algorithm, neural network algorithm and deep learning algorithm. This paper focuses on the deep learning algorithm which greatly reduces the error, improves the measurement accuracy and robustness, and provides a reference for the selection of forecasting algorithms for difference types of electric load forecasting.

2. short-term electric load forecasting

2.1 The classification of electric load forecasting

The electric load forecasting is generally divided into four types which are long-term electric load forecasting, medium-term electric load forecasting, short-term electric load forecasting and ultra short-term electric load forecasting[3]. Long-term electric load forecasting generally forecasts the electric load in the several tens of years. Medium-term electric load forecasting mainly refers to forecasting the electric load in the coming years. Short-term electric load forecasting mainly forecasts electric load in seasons, weeks, days, or even shorter periods. Ultra-short-term electric load forecasting is generally used to output the load changes in the few minutes to hours. It considers short forecasting time, fast forecasting speed, and high forecasting accuracy, so as to achieve reasonable electric generation capacity dispatching to meet the operation requirements and minimize electric generation costs.
The electric load forecasting can be generally divided into urban civil electric load, commercial electric load, rural electric load, industrial electric load and other electric load forecasting. Among them, urban civil electric load forecasting mainly refers to the household load forecasting of urban residents; Business electric load forecasting and industrial electric load forecasting refer to the forecasting of serving business and industry respectively; Rural electric load forecasting refers to the forecasting of all rural loads (including rural civil electricity, production and irrigation electricity and commercial electricity); Other electric load forecasts include municipal electric, such as street lighting, public utilities, government offices, railways and trams, military and other loads.

2.2 The characteristics of short-term electric load forecasting

(1) **Cyclicity**: it mainly reflected in the same seasonal trend of daily load change[4]. From 0am to 23am, the load fluctuation law is similar. The general rule is that the electricity consumption in holidays and weekends will be slightly higher than that in working days.

(2) **Seasonality**: With the change of seasons, the electric demand and electric consumption mode of users change, resulting in the difference of electric load consumption[5]. The increase in the use of refrigeration equipment in summer and the increase in the demand for heating equipment in winter will make the daily variation trend of electric load different in different seasons.

(3) **Randomness**: The electric consumption will increase due to such uncertain factors as emergencies and extreme weather, so the daily electric load may have random disturbances beyond the normal range[6].

2.3 The algorithms of short-term electric load forecasting

The enthusiasm of domestic and foreign scholars for the load forecasting research of electric system originated from the 1970s[7]. After long-term practice, the scholars have proposed many new algorithms and technologies. They can be roughly generally divided into three categories: traditional forecasting algorithm, machine learning algorithms and deep learning algorithm[8]. The traditional forecasting algorithm include time series algorithm and regression analysis algorithm. Machine learning algorithm mainly include artificial neural network, decision tree, wavelet analysis, random forest, support vector machine, etc. The deep learning algorithm mainly include RNN, LSTM, DBN, TCN, etc[9]. They have made great progress in dealing with nonlinear problems and received extensive attention.

3. Traditional and modern forecasting algorithms

3.1 The time series algorithm

is essentially a regression forecasting. Its calculation speed is fast. By using less historical load data, it can accurately forecast the regularity of historical load changes. The typical time series algorithm is Box Jenkins algorithm forecasting algorithm[10]. Box. Jenkins algorithm forecasting algorithm only considers the time series input of historical data.

3.2 Regression analysis algorithm

The regression analysis algorithm is a commonly used algorithm in early load forecasting, including unitary and multivariate regression analysis. The unitary linear regression algorithm only considers the impact of a single load factor; the multiple linear regression is a linear combination with multiple factors, but the relationship of multiple factors is nonlinear, so the accuracy of short-term load forecasting is not high for those with many influencing factors. ARMA is the most commonly used model forecasting algorithm [11].
3.3 Trend extrapolation algorithm

The trend extrapolation algorithm[12] is a algorithm to forecast electric load based on the internal change law of historical electric load data. The trend extrapolation algorithm often neglecting the relationship between environmental variables and electric consumption data. When the changes of the electric load curve are more complex, the trend extrapolation algorithm will have large forecasting errors at the inflection point [13].

3.4 Grey mathematical theory

The grey mathematical forecasting model was first proposed by Professor Deng Julong [14]. It transforms the load sequence into a regular generating sequence through generating transformation, and then models and completes the forecasting. Hsu[15] compared the grey forecasting model with the forecasting algorithms of time series and exponential smoothing, and concluded that grey models (GM) are more suitable for STL[16]. The biggest characteristic of the grey mathematical theory is that it does not need a lot of data, while the grey mathematical model changes in an exponential form, and its change form is single, so it is one-sided.

3.5 Expert system algorithm

The expert system algorithm complete the forecasting by analyzing the historical electric consumption data and weather data, combining the professional analysis of experts in related fields to extract the forecasting rules[17]. The expert forecasting algorithm can combine the professional theoretical knowledge and experience of experts to a certain extent, there are still some difficulties in the process of transforming some experiences into applications, but which limits its own development and the combination with machine learning algorithms.

4. Shallow machine learning algorithm

4.1 Neural network algorithm

1) The artificial neural network(ANN)

Artificial neural network is one kinds of nonlinear complex network formed by many processing units connected with each other, and can carry out duplicated logical calculation and realize nonlinear relationship system. Its four basic characteristics are non linearity, non limitation, non quality and non convexity. It has strong learning and generalization ability, and is very suitable for electric load forecasting. At present, the multilayer feedforward neural network proposed by Minsley and Papert is the most commonly used network structure. Artificial neural network is widely used mainly because it can solve the process and system which is difficult to describe with analytic rules, and is good at simulating complex nonlinear mapping relations [18].

2) The radial basis function (RBF) neural network

The RBF neural network algorithm has good approximation effect on the nonlinear algorithm. Because the electric load forecasting is affected by the real-time price, it has good function approximation function and performs well in function regression, with fast approaching speed and good forecasting progress. the literature[19] proposed an improved RBF network with PSO algorithm. Through comparison, this algorithm effectively solves the problem that PSO algorithm which can be falled into local optimal solution. Compared with the traditional RBF forecasting algorithm, it has faster convergence speed, and the average percentage error can be controlled within 1.2%.

3) The back propagation (BP) neural network

The BP neural network algorithm proposed by Rumelhart et al. [20] The BP neural network is error back propagation, and its rule is the difference back propagation process between the actual value and the error value. The actual value within the normal range is obtained by continuously
correcting the error threshold [20], but the BP algorithm has problems such as long iteration period and slow convergence speed. The traditional BP algorithm converges slowly, resulting in low network training efficiency. However, the PSO algorithm is a good solution algorithm to improve the generalization level of BP neural network. The accuracy of forecasting is high, and the short-term electric load forecasting can be well forecasted.

4.2 The random forest algorithm

The random forest algorithm is a classifier that includes multiple decision trees. The random forest algorithm combines multiple weak classifiers by voting or averaging, which makes the last results have high generalization and accuracy, and solves the shortcomings of the weak generalization ability of the decision tree. The random forest algorithm is often used in load forecasting for its good generalization insensitivity and performance to noise. The random forest algorithm has many advantages. In fact, the performance of random forest algorithm is very practical and strong, so it is highly welcomed and concerned in the electric industry, and its application future is very far-reaching and broad [21].

4.3 The support vector machines(SVM) algorithm

The support vector machines(SVM) algorithm is essentially one generalized linear classifier, which classifies the input data by the supervised machine learning algorithm. It has some advantages in solving the scence with small samples, nonlinearity. The support vector machine algorithm can effectively train sample and find the global optimal solution with quadratic programming, and avoid the local optimal solution. The SVM algorithm has the advantages of fast convergence in the short-term electric load forecasting, and has some controllable factors. Nowadays, The SVM algorithm is effective to solve engineering problems by weak linearity and by related smooth data curves. However, it is still difficult to solve the problem of strong nonlinear characteristics such as load forecasting[22].

5. Swarm intelligence algorithm

The swarm intelligence algorithms are usually inspired by nature. The key swarm intelligence algorithms include particle swarm optimization algorithms, ant colony algorithm, bee colony algorithm, and Lvyan homing algorithm. Algorithms generally have many controllable interface variables. When the algorithm is selected properly and the parameters are accurate, it can solve nonlinear problems in multi-dimensional space, To solve the problem of black box algorithm, it has the ability of global search, can complete multi-dimensional and multi-objective optimization, and even can achieve wonders on some algorithms. In application, we can directly apply swarm intelligence algorithm to select and optimize parameters of other algorithms such as neural networks.

5.1 Particle swarm optimization(PSO) algorithm

According to the rule of birds' activities and foraging behavior, scholars proposed particle swarm optimization (PSO) algorithm. PSO algorithm has the advantages of fast convergence and easy programming, and is widely used in many areas. In recent years, although scholars use PSO algorithm to solve short-term load forecasting more effectively than genetic algorithm, the proposed algorithm has some shortcomings, such as having slow convergence and fall into the local optimization. The reason is that the diversity is not maintained and the updating algorithm is too simple, which leads to the stop of particle evolution and the simplification of particles in the population[23].

5.2 Ant colony algorithm (ACA)

Ant Colony Algorithm (ACA) has a remarkable effect on such high dimension, non convex, discrete nonlinear combination optimization problems as medium and long-term load forecasting.
5.3 Bee colony algorithm (BCA)

Bee colony algorithm (BCA) improve the efficiency and accuracy of short-term electric load forecasting. The bee colony algorithm dynamically calculates the fitness of the bee colony in the process of iteration, sets the fitness threshold, divides the whole bee colony into two subgroups of hired bees and non-hired bees, and approaches the global optimum through hired bees. The scouting bees can steadily search for the global optimum, the bee colony algorithm has both local and global search capabilities in the iterative process, improves the convergence speed, accuracy and stability of the network[25].

6. Deep learning algorithm

With the increase of data and the demand for nonlinear algorithms in algorithming, the depth of neural networks is deepening. The successful application of deep learning algorithm in image recognition, natural language processing and other fields makes deep learning become the current research hotspot. The deep learning network constructs the connection layer between the input and the output layer with the hidden layer. In general, the deep learning algorithm is used in fault diagnosis, load and forecasting, operation regulation, etc. of electric system [26-27]. It can be mainly divided into the following [28]:

6.1 Automatic encoder algorithm

Automatic encoder algorithm is a representation learning algorithm through neural network. The encoding part is used to learn the features of hidden layer in the input signal, and the decoding part reconstructs the original input data with the new features learned. The depth self encoder was proposed by Professor Hinton and others in 2006. Later, other scholars also successively proposed sparse automatic encoder, noise automatic encoder, stack automatic encoder, shrink automatic encoder [29], etc.

Self encoder with deep learning algorithm is usually used to forecast short-term load. A study has proposed a stack denoising self encoder for short-term load forecasting. The author used two algorithms to conduct online forecasting and one day ahead forecasting respectively. The results forecasting of the previous day has a better effect[30].

6.2 Convolutional neural network (CNN) algorithm

Convolutional neural network (CNN) algorithm[31] is one kind of feedforward neural network algorithm that uses convolution computation with deep structure. The CNN algorithm is a partially connected network with only some adjacent nodes. The CNN algorithm compared with full connection neural network has the advantages of sparse connection and sharing weight. The CNN structure greatly reduces the complexity of algorithm, and reduces the parameters to prevent the computing speed from slowing down for having too many parameters in network. The key of CNN algorithm is convolution layer and pooling layer.

6.3 Restricted Boltzmann machine algorithm

The restricted Boltzmann machine (RBM) algorithm[32] is a random neural network algorithm. Boltzmann machine has no connection constraint. There is a connection between hidden nodes and visible nodes of the RBM algorithm, but there is no connection between two hidden nodes and visible nodes, that is, there is full connection between these layers, and no connection within these
layers. Obviously, the RBM algorithm corresponds to a bipartite graph. Visible variables and hidden variables are binary variables. The RBM algorithm is mainly used in deep learning of short-term electric load forecasting.

6.4 Recurrent neural network algorithm

1) Recurrent neural network (RNN)

The recurrent neural network (RNN) algorithm connection mode is different from that of the feedforward neural network. The RNN algorithm not only receive the data output of the previous neuron nodes, but also receive the data output of the previous neuron nodes of the same network layer. The difference layers are connected by the weights. It helps us to solve some related problems in short-term load forecasting with more reasonable structure with the RNN, including short-term memory network, gated cyclic unit network[33], etc.

Because short-term electric load forecasting is affected by many factors, such as weather, temperature, electricity price, holidays, etc., these factors are difficult to be quantifying, resulting in the low forecasting accuracy of traditional forecasting algorithms. It can be seen that the closer selection of short-term electric load forecasting algorithms to generate actual results, the better the forecasting effect. In 2002, Vermaak and Botha[34] used RNN algorithm for short-term load forecasting. Due to the dynamic characteristics of RNN, this algorithm can better capture the feature of input data.

With the development of RNN forecasting algorithm and the more comprehensive factors affecting short-term electric load forecasting, there are increasingly high requirements for forecasting accuracy. Scholars are committed to improve the efficiency and accuracy of the algorithm. It optimized the initial weight of RNN forecasting algorithm through ant colony algorithm and built an improved RNN forecasting algorithm. The results show that the algorithm converges rapidly, the forecasting accuracy is higher than that of the traditional RNN forecasting algorithm[35].

2) Long and short term memory neural network (LSTM)

One of the most widely used algorithms in the field of short-term electric load forecasting is LSTM. Literature[36] proposed a multivariate short-term electric load forecasting algorithm based on two-way short-term memory neural network and multivariate linear regression for user level integrated algorithm with more randomness and volatility. The bidirectional long short memory network is a common LSTM form. Its advantage lies in the ability to extract information from time series with strong correlation in both directions. Detailed practical examples show that the algorithm has higher forecasting accuracy.

Literature[37] proposed a CNN-LSTM short-term electric load forecasting algorithm with the attention mechanism. Through the feature extraction of the load data through CNN, LSTM learned the time series feature of the short-term electric load, and the attention mechanism was used to allocate different weights to the LSTM hidden layer, which effectively solved the problem of LSTM losing important information on long order data, which proves that the algorithm proposed has certain application potential.

Literature[38] proposed one ultra short-term electric load forecasting combination algorithm with LSTM and XGBoost. After the establishment of LSTM forecasting algorithm and XGBoost forecasting algorithm, the paper used the reciprocal error algorithm to combine LSTM and XGBoost for forecasting [39]. It is obviously higher than the comparison algorithm GRU, LSTM, XGBoost, GRU-XGB and other algorithm algorithms listed in the article.

Some scholars put forward a short-term load forecasting algorithm based on combination algorithm with Prophet and LSTM[40], established Prophet forecasting algorithm and LSTM forecasting algorithm respectively, and then used the least square algorithm to take different weight combinations to obtain new algorithms and forecast, and compared the single algorithm LSTM, Prophet, ARIMA, random forest and Prophet LSTM combined algorithm, Results were significantly better than other single algorithms.

3) Gate recurrent unit neural network (GRU)
The Gate recurrent unit (GRU) network reconstructed from it greatly improves the training speed of LSTM and reduces the dependence of machine learning on high-performance equipment on the premise of almost no loss of accuracy[41]. At present, the existing LSTM algorithm has the shortcomings of many training parameters and slow convergence speed. GRU's gated recurrent neural network is an improvement on LSTM. Because it combines LSTM's input gate and forgetting gate into update gate, it has the characteristics of few training parameters and fast convergence speed by maintaining the forecasting accuracy of LSTM algorithm.

The algorithm has significant advantages over LSTM and GRU algorithms in terms of forecasting accuracy and forecasting efficiency in short-term forecasting of electric load in a certain area. The Attention mechanism highlights the input characteristics that play a key role in forecasting. Due to the acceleration of GRU, the algorithm is faster and more accurate than LSTM in forecasting. Although the forecasting accuracy of GRU has been improved compared with the traditional forecasting algorithms, in the face of such high noise, high interference and discrete time series data as electric load, a single machine learning algorithm is usually difficult to eliminate interference and establish a algorithm with practical and accurate forecasting capability[42].

### 6.5 Deep Belief Network (DBN)

Deep belief network (DBN) is a variant of deep belief model, which is based on RBM's hidden layer output as the input of the next RBM's visible layer. Unlike DBM, DBN is undirected except that the lowest visible layer is a directed link. Literature [43] based on the historical short-term electric load forecasting data of the substation and the external environment data in location, avoids the problem of selecting features such as similar days by using the deep belief network algorithm with strong learning ability, and uses the Nadam momentum optimization algorithm to train the DBN to obtain the best parameters, which constitutes a learning framework for substation short-term electric load forecasting, and automatically adjusts the DBN structure based on the Keras deep learning framework, The optimal forecasting result is reached [44].

### 6.6 Temporal Convolutional Network (TCN)

In recent years, in the process of building composite algorithms, CNN performs well in extracting and fusing long-term sequence features, but performs poorly when used solely for time series modeling, and most of the deep learning algorithms will have the problem that the learning ability of the network degrades with the deepening of network layers. Based on the characteristics of CNN and dedicated to solving the degradation problem of deep network, TCN network algorithm has emerged, which can extract features from data with long-term dependency and reduce the degradation problem of deep network by adding residual blocks[45].

### 7. Integrated Learning algorithm

At present, Bagging, AdaBoost, XGBoost, GBDT and other integrated learning algorithms[46] are widely used. Bagging adopts parallel mode, and the training speed is fast, but some data sets can not be obtained due to back sampling, and each learner has the same weight, without considering its importance AdaBoost runs in serial mode and can adjust the weight according to the algorithm training results. The algorithm training accuracy is higher, but the number of weak learners is not easy to set and the training is time-consuming To solve this problem, literature[47] used the decision tree as a weak learner to build a random forest (RF) algorithm for load forecasting, and used rough set, gray projection technology, and Drosophila optimization algorithm to optimize the algorithm, all of which achieved good results.

XGBoost algorithm uses the second derivative, which improves the calculation efficiency and effectively reduces the over fitting problem. However, each iteration run needs to traverse the entire data set, which is not suitable for processing large data sets, and also has high requirements for memory In order to solve this problem, literature[48] uses the stacking technology of integrated
learning to combine XGBoost, light gradient boosting machine (LGBM), and multi-layer perceptron MLP algorithms for load forecasting, showing good performance advantages[49].

In addition to the above integrated learning algorithm, some experts also integrate SVM, neural network, integrated learning and other algorithms for combined forecasting. Literature[50] combined CNN and RF to effectively reduce MAPE To solve the problem of low forecasting accuracy caused by few data features, used CNN to extract the local trend of user electric load and combined LSTM algorithm for forecasting. The MAPE reciprocal weight (MAPE-RW) algorithm[51] to select the optimal weight of the algorithm combination, weighted the time series data forecasted by XGBoost and LSTM algorithms, and effectively reduced the forecasting error of a single algorithm[52] used CNN to extract data features and Seq2Seq to forecast user energy load. At the same time, attention mechanism and multi task learning algorithm were used to improve the accuracy of short-term electric load forecasting. The data preprocessing, multi-objective optimization, algorithm forecasting and algorithm evaluation are combined into a new load forecasting system. The EMD is used to decompose the load data, and the RBF, ELM and GRNN forecasting algorithms are combined to forecast, and the results are weighted to form the final forecast value[53-55].

Conclusion

From traditional algorithm to machine learning algorithm and then to deep learning algorithm, the construction of short-term electric load forecasting algorithm is constantly approaching the actual situation. From the existing forecasting algorithm, the forecasting algorithm based on LSTM, CNN, DBN, RNN and other deep learning algorithms is obviously better than the algorithm built by traditional algorithm and machine learning algorithm. However, the forecasting algorithm that combines intelligent optimization algorithm and other domain algorithms with deep learning network algorithms has stronger learning effect on external factors that are difficult to quantify, and has good scalability.

References


