Research on load loss prediction of distribution network outage based on hybrid neural network

Jiamin Liao*
CHINA SOUTHERN POWER GRID CO., LTD.
*Corresponding author’s e-mail: s2120334@siswa.um.edu.my

Abstract—In order to fully mine the relationship between temporal characteristics in load data and improve the accuracy of load forecasting, a load forecasting method based on convolutional neural networks (CNN) and gated recurrent unit (Gru) hybrid neural network is proposed. Taking date factors, climate factors and similar daily load factors as input, the sample data sets in the region are grouped by K-means clustering method; Then cnn network is used to extract the relationship between features and load in high-dimensional space, construct the high-dimensional feature vector of time series, and input the results into Gru network; Finally, the parameters of Gru network model in each group are trained and the load forecasting value is output. The results show that the proposed load forecasting method has significant advantages in forecasting accuracy and efficiency compared with long short term memory (LSTM) network model, Gru network model, cnn-lstm network model, support vector machine regression model and decision tree model.

1. INTRODUCTION
The stable operation of power system requires real-time dynamic balance between power generation and load change. However, the randomness of load fluctuation is strong, the nonlinear characteristics are obvious, and the influencing factors are diversified [1]. Therefore, it is necessary to deeply explore the law of load change and improve the accuracy of load forecasting, so as to reasonably formulate power production plan and reduce power production cost, Effect of maintaining the balance between supply and demand of power resources [2-3].

At present, load forecasting methods at home and abroad are mainly divided into three categories: statistical method, time series forecasting method and artificial intelligence method. The statistical method is represented by Kalman filter method [4]; The time series method is divided into regression analysis method [5], multiple linear regression method [6], Fourier expansion method [7], etc..

2. CONSTRUCTION OF LOAD FORECASTING FEATURE SET
Reasonable construction of prediction feature set is very important to the convergence and prediction accuracy of the model.

1) Considering the date factor, the industrial load accounts for a large proportion in the urban power load, while the date factors such as weekends and holidays have a great impact on the industrial load. Based on this, a 6-D date feature including season, month, day, week date, whether it is a working day and whether it is a holiday is constructed. For \{ {l} \} \_ \{ Text \{r\} \ (T) \} for modeling.
2) Considering the climate factors, the changes of temperature, weather type and other climate factors will affect the residential power load, this paper constructs a six-dimensional climate feature including weather type, wind force, maximum temperature, minimum temperature, maximum humidity and minimum humidity \( w(T) \) for modeling.

3) Considering the similar daily load factors, because the load has time series characteristics, the model can learn the recent change law of load according to the daily load data close to the date to be measured, so as to enrich the prior information of the prediction model. Based on this, this paper selects the load data of four similar characteristic days before the forecast day to form a four-dimensional vector pair \( \langle l \rangle \) for modeling.

To sum up, a 16 dimensional feature vector including date factors, climate factors and similar daily load factors is constructed as the input of the prediction model, as shown in Table 1.

3. CNN-GRU HYBRID NEURAL NETWORK MODEL

In order to verify the scientificity and accuracy of the proposed load forecasting method, 731 days of power load data in a region of Zhejiang Province from January 1, 2016 to December 31, 2017 are selected as the data set, which is divided into training set and test set according to the ratio of 8:2, including 585 days of training set and 146 days of test set. The data is collected once a day, and the input feature dimension is 16 dimensions, as described in Table 1.

The cnn-gru hybrid neural network is used to train on the data set. The prediction results are compared with LSTM network, Gru network, cnn-lstm network, SVR regression model and decision tree model under the same conditions in terms of prediction efficiency and prediction accuracy. It can be known that the model has good prediction accuracy while maintaining high-efficiency training.

![Fig. 1 Variation trend of sum of squared errors and contour coefficient](image)

3.1 experimental environment configuration

The experimental environment of this paper adopts Intel i7-8750h processor and nvidia-gtx-1060 graphics
card. The algorithm model adopts Python 3.6 as the programming language, the software architecture is based on tensorflow framework, python framework and scikit learn algorithm framework, and the drawing tool adopts Matplotlib drawing module.

3.2 load cluster analysis

Carry out K-means clustering on the load data, and draw according to formula (7) (8). When the clustering category \( K \) increases, the variation trend of the sum of squares of errors and contour coefficient is shown in Figure 4. As can be seen from Fig. 4, according to the cluster evaluation elbow method [25], when the cluster category \( K \) is set to 2 or 3, the absolute value of the second derivative of the sum of squares of errors is large, which is suitable for selecting the cluster category; For the contour coefficient, when \( K = 3 \), the contour coefficient \( s(I) \) gets the maximum value of 0.48. Based on this, the clustering category \( K = 3 \) is set to 3, and the 585 day training set is divided into three categories: A, B and C. The capacity of the three categories of samples is 196, 135 and 254 respectively. The load clustering results from January 1 to December 31, 2016 are shown in Table 1 in the appendix. After training the three kinds of samples respectively, the final prediction model is integrated. The coordinate system is established based on the normalized maximum temperature, maximum humidity and weekly date, and the load clustering distribution is obtained, as shown in Figure 5. It can be seen from Figure 5 that the sample has a clear clustering effect in high-dimensional space.

![Fig. 2 Distribution of load clustering features](image-url)
3.3 determination of network architecture super parameters

This paper considers the super parameters of cnn-gru hybrid neural network training, including the number of layers of Gru network \(D\) and L2 regularization weight coefficient \(\mu\) in CNN training. Using the control variable method, based on 3000 iterations, the number of layers of Gru network is deepened step by step without setting the regularization term, so as to find the optimal network layer \(\{D\}^*\) to minimize the empirical risk of the model. On this basis, the L2 regularization weight coefficient \(\mu^*\) is adjusted to determine the L2 regularization weight coefficient \(\{\{\mu\}^*\}^*\) to minimize the structural risk of the model.

The sum of the electric power taken from all operating power equipment of the power customer to the power system at the same time is called the user's electric load at that time. Then the sum of the power loads supplied by a power grid to all users of the power grid at the same time is called the load of the power grid. The maximum load borne by the power grid in the same natural year is called the maximum load. The common division of load types is the power consumption of the whole society, which is divided into four categories: the first, the second, the third industry and residential power consumption. It can also be subdivided into four categories according to agriculture, forestry, animal husbandry and fishery, industry, geological survey and exploration, construction, transportation, posts and telecommunications, commercial food supply, marketing and storage, residential power consumption and others.

types of power load forecasting

Load forecasting is divided into system load forecasting and bus load forecasting. System load forecasting is mainly used for load control, safety monitoring, preparation of power grid operation plan and arrangement of power grid maintenance plan. Bus load forecasting is to convert the load forecasting value of power system into the calculated value of active load and reactive load of each bus in the system at a specified time.

(1) Divided by forecast time. According to the length of load forecasting time, power load forecasting can be divided into long-term load forecasting, medium-term load forecasting and short-term load forecasting.

(2) According to the characteristics of predicted load, it is divided into maximum load forecasting, minimum load forecasting, average load forecasting, peak valley difference forecasting, average peak load forecasting, average underestimated load forecasting, bus load forecasting, load rate forecasting, etc.

(3) It can be divided into urban civil load forecasting or commercial load forecasting, rural load forecasting, industrial load forecasting, etc.

characteristics of power load forecasting

Because the power system load forecasting has a series of characteristics such as inaccuracy, conditionality, timeliness and multi scheme, in order to accurately obtain the change of future load, on the one hand, we should fully grasp and make use of its characteristics, at the same time, we should also consider the influence of various factors and select the corresponding forecasting method according to the specific factors. Finally, the load forecasting model is established combined with the forecasting method. The characteristics of load forecasting are as follows:

(1) Inaccuracy. In the process of power system load forecasting, it is easy to be affected by many factors (such as temperature and season), and the related influencing factors are also developing and changing. In addition, it may be affected by emergencies, so the prediction results are inaccurate.

(2) Conditionality. Conditionality means that the influence of various factors must be comprehensively considered when forecasting the future power load data. The said conditions can be divided into the inevitable conditions that can reliably and directly affect the load forecasting results and some assumptions that we need to add in order to obtain more accurate forecasting because it is difficult to control the development and change regularity of future load.

(3) Temporality. Timeliness means that the power system load forecasting is completed within a certain time range, that is, the power system load forecasting is required to be real-time, and the prediction time corresponding to the predicted value and the time range of this prediction shall be specifically indicated when making the power system load forecasting.

(4) Multi scheme. In power system load forecasting,
different forecasting methods can be selected according to different forecasting conditions, and then an appropriate load forecasting model can be established. Due to the inaccuracy and conditionality of load forecasting, different forecasting methods should be adopted for loads with different characteristics under the condition of comprehensive consideration of various influencing factors. Different mathematical models are established to make load forecasting more accurate.

With the deepening of Gru network layers, the change of square difference loss is shown in Table 2. It can be seen from Table 2 that when the number of layers \( D \) of Gru network is 3, the square difference loss gets the minimum value. When the number of layers is greater than 3, the square difference loss does not decrease but increases due to over learning of the model. Therefore, the optimal number of layers \( D^* \) is 3; Adjust the L2 regularization weight coefficient \( \mu \) of CNN network training to obtain the square difference loss, as shown in Table 3. It can be seen from Table 3 that when the L2 regularization weight coefficient \( \mu^* \) is set to 0.001, the square difference loss after iteration is the smallest. Based on this, in this experiment, the L2 regularization weight coefficient \( \mu \) trained by cnn-gru network model is 0.001, and the number of Gru network layers \( D \) is set to 3.

### 3.4 analysis of prediction results

For the three types of training set data, the CNN Gru hybrid neural network model proposed in this paper, LSTM model, Gru model, cnn-lstm model, SVR regression model and decision tree model are used to fit the training data. Classify the 146 day test set according to the clustering results obtained in Section 3.2, and then select the corresponding model for prediction. In the above network model, the LSTM network model is composed of three-layer LSTM units; The Gru network model consists of three layers of Gru units; The CNN architecture setting in cnn-lstm model is the same as the model proposed in this paper; The kernel function of SVR model adopts Gaussian kernel function, and the penalty parameter is set to 5.0; The penalty parameter of pre cut branches and leaves node of decision tree model is set to 0.01.

For the above prediction model, calculate the maximum percentage of prediction error on three types of test samples \( \{ y \}_{E-MAX} \), prediction accuracy \( \{ y \}_{FA} \), Percentage error between average value and average absolute value of text \( \{ y \}_{MAPE} \), and the comparison results are shown in Table 4. The comparison of prediction efficiency of depth model is shown in Table 5. According to the information in tables 4 and 5, compared with the traditional Gru network, the maximum percentage of prediction error of the cnn-gru neural network model proposed in this paper is reduced by 4.79%, and the prediction accuracy is improved by 1.11%; Compared with LSTM neural network, SVR regression and decision tree model, the maximum error percentage decreased by 2.9%, 12.05% and 15.81% respectively, and the prediction accuracy increased by 1.03%, 3.12% and 4.44% respectively; Compared with cnn-lstm model, on the basis of almost the same prediction accuracy, the proposed model can shorten the model training time by 50.43%, and greatly improve the prediction efficiency of the model.

### 4. CONCLUSION

A regional medium-term load forecasting method based on CNN-gru hybrid neural network model is proposed. Firstly, K-means clustering method is used to divide the sample data set, then CNN network is used to extract the relationship between features and load data in high-dimensional space, construct the feature vector of time series and input it into Gru network. Finally, Gru network is trained to output load forecasting value, which has the following advantages:

1) A feature set including date factors, climate factors and similar daily load factors is constructed, which can give full play to the advantages of CNN network in the field of data mining and extract the potential relationship of discontinuous data in high-dimensional space.
2) The Gru cyclic network model can fully consider the time series characteristics of load characteristics, has good time series data fitting and regression ability, and has high prediction efficiency.

3) Cnn-gru hybrid neural network combines the advantages of CNN network and Gru network. The experimental results show that compared with the existing artificial intelligence prediction methods, the proposed method not only maintains a faster model training speed, but also has higher prediction accuracy. The load forecasting feature set established in this paper does not take into account the diversified load types, and the price factor is not included in the feature set. Therefore, in the next work, this paper will further study the impact of load characteristic classification on load forecasting, build a richer feature set including electricity price factors, explore the internal relationship of input characteristics, and further improve the accuracy of load forecasting.

References

[1] North American Blackout Time Series Statistics and Implications for Blackout Risk[J]. Benjamin A. Carreras, David E. Newman, Ian Dobson. IEEE Transactions on Power Systems. 2016 (6)

[2] On Hurst exponent estimation under heavy-tailed distributions[J]. Jozef Barunik, Ladislav Kristoufek. Physica A: Statistical Mechanics and its Applications. 2010 (18)

[3] Power-Law Distributions in Empirical Data[J]. Aaron Clauset, Cosma Rohilla Shalizi, M. E. J. Newman. SIAM Review. 2009 (4)

[4] Large blackouts in North America: Historical trends and policy implications[J]. Paul Hines, Jay Apt, Sarosh Talukdar. Energy Policy. 2009 (12)

[5] Risk-management and risk-analysis-based decision tools for attacks on electric power.[J]. Simonoff Jeffrey S, Restrepo Carlos E, Zimmerman Rae. Risk analysis: an official publication of the Society for Risk Analysis. 2007 (3)

[6] Evidence for self-organized criticality in a time series of electric power system blackouts[J]. B.A. Carreras, D.E. Newman, I. Dobson, A.B. Poole. IEEE Transactions on Circuits and Systems I: Regular Papers. 2004 (9)

[7] Rescaled variance and related tests for long memory in volatility and levels[J]. Liudas Giraitis, Piotr Kokoszka, Remigijus Leipus, Gilles Teyssière. Journal of Econometrics. 2003 (2)

[8] Long-Term Memory in Stock Market Prices[J]. Econometrica. 1991 (5)

[9] Robustness of the rescaled range R/S in the measurement of noncyclic long run statistical dependence[J]. Benoit B. Mandelbrot, James R. Wallis. Water Resources Research. 1969 (5)

[10] Fractional Brownian Motions, Fractional Noises and Applications[J]. Benoit B. Mandelbrot, John W. Van Ness. SIAM Review. 1968 (4)