Applying Deep Learning Method to Data Analysis of Low Voltage Line Carrier Module

Guorui Cao, Li Yang, Xuejun Yu*, Ruijun Zhong and Jinyu Lu

Electric Power Science & Research Institute of Tianjin Electric Power Company, Tianjin, China

*Corresponding author e-mail: sword_2005@126.com

Abstract. With the increasing popularity of the global energy Internet, the smart grid in the city has developed rapidly, and the number of smart energy meters is increasing year by year. The demand for detecting the carrier module of the power meter has also increased to a new level. But the existing carrier module data analysis scheme is not perfect. It is urgent to need a complete set of carrier module data analysis scheme and new detection technology to realize the multiplexing of carrier module data, and to excavate the value of the carrier module to detect the data. Unlike previous studies, which using SVM and Bayesian networks, this paper proposes a scheme for classification of carrier module detection data using a deep learning network. GRU deep neural network is used to model the test data and identify the environment automatically. According to the recognition results, the carrier module can be adjusted accordingly. Experiments show that the scheme proposed in this paper has good results. Our scheme has very good effect in low voltage power line carrier communication data analysis task. It can excavate valuable information of low voltage power line carrier communication and has high economic value. In addition, the technical solutions proposed in this paper will not affect the existing business processes, and can be applied in large areas.

1. Introduction

In recent years, the information technology in the power industry has developed rapidly. With the promotion of the global energy Internet, the development of the urban smart grid is rapid, the use of intelligent energy meter is increasing year by year, and the detection demand for the power meter carrier module is also increasingly promoted to a new height. In China, from the initial management information construction of financial computerization represented by electric power production automation, and to the construction of large enterprise informatization in recent years, especially the construction of the next generation smart grid, it has made the power data resources increase rapidly, and has formed a certain norm. It is imperative to make use of the carrier module data. How to use this valuable data resource to excavate its data value is the focus of attention nowadays [1]. As we all know, every part of the useful value of data mining is about 100 million revenue increase for the power system.

After years of development, machine learning has become a multidisciplinary field [2]. It has been widely applied in various industries [3] and has played an important role in practice [4]. Traditional
machine learning methods include native Bayes [5], decision tree [6], boosted tree [7], and support vector machine [8] and so on. It has achieved good results in image processing [9], text classification [10] and other tasks [11]. In recent years, with the introduction and application of deep neural network, the effect of image processing, text classification and other tasks has been further improved. Deep neural networks include two categories: recurrent neural network (RNN) [12] and convolutional neural network (CNN) [13]. Because of the time character of carrier module detection, the recurrent neural network is more suitable for the task of carrier module detection data classification.

From the recurrent neural network to the present, many improved models have been produced, such as LSTM [14], GRU (Gated Recurrent Unit) [15], tree-LSTM [16], Nested-LSTM [17], etc. In this paper, we use GRU network for modeling. Experiments show that this scheme can solve the classification problem of carrier module testing well.

The contributions of this article are as follows:

1. We propose to use GRU neural network to analyze the detection data of carrier module. Our scheme can well detect the value of carrier test data.
2. In this paper, special processing of power grid data is carried out, so that power network data can be processed by deep neural network.
3. The experiment proves the validity of the GRU neural network, which has a certain guiding significance for the future power grid data processing.

2. Flow chart

![Diagram](image)

Figure 1. The flow chart of our paper.

First introduce the flow chart of this scheme. This article is shown as shown in Fig.1. It is mainly composed of two parts: one is data processing, the other is GRU neural network. In the following chapters, the two parts will be described in detail.

3. Data

3.1. Data sources
Because there is no open data set, manual data collection is required. Through the preliminary investigation and analysis, this paper determines the data collected from three typical stations, the three districts are: urban, rural and urban fringe.

Urban area: There are many kinds of electrical equipment in this area, and there are usually filter protection circuits. Most of them are cable wiring, with a small radius of the station, and large frequency conversion devices (such as frequency conversion pumps) around them.

Rural area: Generally speaking, the platform is single-phase power supply, with large radius of the station area and relatively old lines, and the noise interference is relatively small.

Urban-rural fringe: The station is more complex, the noise interference type is uncertain, and the noise infection is more serious.

3.2. Power carrier module
The power carrier module is mainly used to concentrate the meter reading. It can transmit the meter data to the upper level meter reading system. The basic principle is to coupling the signal to the transmission line to realize the transmission of the signal on the local area network.
3.3. Dynamic power test
The method of testing carrier module is divided into static power test, dynamic power test, carrier read positive active test, LED transceiver lamp test. This paper analyzes the data obtained from dynamic power test. The dynamic power test software sends the voltage sampling command and the current sampling command to the test panel.

The test panel receives the command of the voltage sampling and current sampling, and returns the sample value to the testing software to judge whether it meets the requirement of dynamic power. The principle of dynamic power testing is shown in Fig.2

![Figure 2. The principle of dynamic power testing.](image)

3.4. Data pre-processing
Direct test data need to be pre-processed before it can be analyzed and excavated. In this paper, data pre-processing includes two parts: data cleaning and data combination.

3.4.1. Data cleaning. In data mining, there are a large number of incomplete, inconsistent and abnormal data in the massive raw data, which seriously affect the efficiency of the implementation of data mining modeling, and may even lead to the deviation of the mining results, so it is particularly important to clean the data. On the one hand, data pre-processing is to improve the quality of data, and on the other hand, it is better to adapt data to specific mining techniques or tools.

Data cleaning is mainly to delete unrelated data from the original data set, duplicate data, smooth noise data, and filter out data unrelated to the topic of mining, and deal with missing values and abnormal values. Statistically, missing data may produce biased estimation, which makes the sample data not well represented, and most of the data in reality contain missing values, so it is important to deal with missing values. The common methods of dealing with missing values are deletion, substitution and interpolation. In this paper, the percentage of missing values is relatively small, so the deletion method is used.

3.4.2. Data combination. Since we use GRU to model, input needs time series. So, we have to combine data into time series. In this paper, we use each time period (every two hours for one time period) to select a data way, and combine the input sequence with data growth rate of 12 \{t1, t2, t3...t12\}. Such data has time characteristics and can be handled better by GRU.

4. GRUs
RNN can handle time series well, but there will be a problem of gradient disappearance. GRU is a very mainstream RNN derivative. The network structure of GRU can alleviate the gradient vanishing problem very well. The mathematical formalization of GRU is as follows:

\[
h_t = GRU(h_{t-1}, x_t)
\]
The hidden state $h_t$ decode the representation of the time $t$. The overall architecture of GRU is shown in Fig.3. More details about GRU can be got in (Chung J, Gulcehre C, et al. 2015) [15]. The training data $\{x_1, x_2, \ldots, x_N\}$ through GRU, getting output $\{h_1, h_2, \ldots, h_N\}$, and then the vector $h_N$ is obtained through the average pool operation. Finally, we get the class distribution vector $y$ through a softmax regression.

![Figure 3. The overall architecture of GRU.](image)

5. Experiment

5.1. Data
We annotate the measured data according to the type of area we selected before. The details of the dataset are shown in Table 1.

| Transformer area | Train | Test |
|------------------|-------|------|
| Urban area       | 3600  | 600  |
| Rural area       | 2600  | 800  |
| Urban-rural fringe | 4600  | 500  |

5.2. Experiment Setting
Usually, there is an objective function in every algorithm of machine learning. The process of solving the algorithm is through the process of optimizing the objective function. In classification problem, loss function is usually used as its objective function. The loss function is used to evaluate the difference between the predicted value and the real value of the model. The better the loss function is, the better the performance of the model is. The loss function used in this paper is as follows:

$$L(\theta) = -\sum_i \tilde{y}_i^j \log y_i^j + \beta \sum \|\theta\|^2$$

(2)

Where $\tilde{y}_i$ is the true distribution for input sequence $i$, $y_i$ is the predicted distribution, and $j$ is the classes index. $\beta$ is the coefficient for $L_2$ regularization, $\theta$ is the parameter set.

In our framework, all parameters are initialized from the uniform distribution sampling. The SGD algorithm is used to train the model, and learning rate is set to 0.1. We set $\beta$ at 0.005. In the training process, we use dropout operation with a probability of 0.4 before the softmax layer. The batch size of our model is 15. When the prediction accuracy does not increase 10 iterations, the training process is stopped.

5.3. Comparison with Baseline Methods
To evaluate the performance of our solution in a comprehensive way. Compare our solutions with baseline methods:
Naive Bayesian classifier: The naive Bayesian classifier is a weak classifier based on Bias theorem. All simple Bayesian classifiers assume that each sample feature is not associated with other features. The basic idea of naive Bayes is this: for the given item to be classified, the probability of each category under the condition of this item is solved, which is the largest, which is considered to which category to be classified.

RNN (Recurrent neural Network) depicts the relationship between the output of a sequence and its previous information. From the network results, RNN will remember the previous information and use the previous information to affect the output behind. That is to say, the nodes between the hidden layers of the RNN are connected, and the input of the hidden layer not only includes the output of the input layer, but also the output of the hidden layer at the last time.

We use accuracy to evaluate the performance of the model. The following is defined as follows:

$$\text{Acc} = \frac{C}{N}$$  \hspace{1cm} (3)

Among them, C is the number of samples correctly predicted, and N is the total sample size. In general, the better the model is, the higher the accuracy is.

We evaluate these models on our data set and the results are shown in Table 2.

**Table 2. Accuracy of different models on data sets.**

| Model                  | Acc  |
|------------------------|------|
| Naive Bayesian classifier | 0.56 |
| RNN                    | 0.63 |
| GRU                    | 0.74 |

From Table 3, we can see that the accuracy rate of the naive bayesian classification method is the lowest. Other methods are all based on the depth learning model, and the accuracy is higher than that of the naive bayesian classification. This indicates that the deep learning model can well extract useful features and effectively analyze the detection data of carrier modules.

The accuracy of RNN is lower than that of GRU, because RNN has the problem of gradient disappearance and can not solve long distance dependence well. GRU can ease the problem of gradient disappearance. GRU uses gate function to selectively allow partial information to pass. The results show that the detection data of carrier module can be analyzed well by using GRU.

6. Conclusion and Future Work

In this paper, a solution of carrier module detection data classification based on GRU deep neural network is proposed. The main idea of this paper is to combine test data into time series and extract features by using GRU network to classify. The results show that it is feasible and effective to introduce deep learning into the data analysis of carrier module. The GRU neural network model can effectively recognize the environment of the carrier module, and adjust the carrier module dynamically according to the recognition results. It has high economic value. In future work, we plan to use other deep learning methods to analyze the detection data of carrier module. In addition, we will use deep learning to solve more problems related to power system data analysis.

References
[1] Chen L L, Liu W. Carrier communication abnormal data detection system based on the Cloud computing design [J]. Ship Science & Technology, 2016.
[2] Sundermeyer M, Schlüter R, Ney H. LSTM Neural Networks for Language Modeling [C]/// Interspeech. 2012: 601-608.
[3] Qing, Xiangyun, Niu, et al. Hourly day-ahead solar irradiance prediction using weather
forecasts by LSTM [J]. Energy, 2018, 148.

[4] Tang Y, Huang Y, Wu Z, et al. Question detection from acoustic features using recurrent neural network with gated recurrent unit [C]// IEEE International Conference on Acoustics, Speech and Signal Processing. IEEE, 2016: 6125-6129.

[5] Mccallum A. A comparison of event models for naive bayes text classification [C]// Proc. AAAI-98 Workshop on Learning for Text Categorization. 1998: 41-48.

[6] Quinlan J R. Induction on decision tree [J]. Machine Learning, 1986, 1 (1): 81-106.

[7] Tu Z. Probabilistic Boosting-Tree: Learning Discriminative Models for Classification, Recognition, and Clustering [C]// Tenth IEEE International Conference on Computer Vision. IEEE, 2005: 1589-1596 Vol. 2.

[8] Cortes C, Vapnik V. Support-vector networks [J]. Machine Learning, 1995, 20 (3): 273-297.

[9] Reyna R A, Esteve D, Houzet D, et al. Implementation of the SVM neural network generalization function for image processing [C]// IEEE International Workshop on Computer Architectures for Machine Perception, 2000. Proceedings. IEEE, 2000: 147-151.

[10] Sun A, Lim E P, Liu Y. On strategies for imbalanced text classification using SVM: A comparative study [J]. Decision Support Systems, 2010, 48 (1): 191-201.

[11] Kim M Y, Xu Y, Goebel R. Legal Question Answering Using Ranking SVM and Syntactic/Semantic Similarity [M]// New Frontiers in Artificial Intelligence. Springer Berlin Heidelberg, 2014: 244-258.

[12] Williams R J, Zipser D. A learning algorithm for continually running fully recurrent neural networks [M]. MIT Press, 1989.

[13] Krizhevsky A, Sutskever I, Hinton G E. ImageNet classification with deep convolutional neural networks [C]// International Conference on Neural Information Processing Systems. Curran Associates Inc. 2012: 1097-1105.

[14] Hochreiter S, Schmidhuber J. Long Short-Term Memory [J]. Neural Computation, 1997, 9 (8): 1735-1780.

[15] Chung J, Gulcehre C, Cho K, et al. Gated Feedback Recurrent Neural Networks [J]. Computer Science, 2015: 2067-2075.

[16] Tai K S, Socher R, Manning C D.: Improved semantic representations from tree-structured long short-term memory networks.arXiv preprint arXiv: 1503.00075 (2015).

[17] Moniz J R A, Krueger D. Nested LSTMs [J]. 2018.