Research on Power Data Mining Algorithm Based on Transfer Learning

Hongyang He

July 12, 2022

Research on Power Data Mining Algorithm Based on Transfer Learning

HE Hongyang, University of Birmingham , Birmingham, 999020, UK, hongyang_he@163.com

Abstract: *With the development of new energy power grid construction, the scale of the power grid is expanding day by day, and the accidents are also increasing by the failure of the power grid and the collapse of the whole network. Therefore, it is necessary to evaluate the power data diagnostic voltage measurement capability of the power grid and customize quantitative indicators based on the evaluation content. At the same time, massive data mining is carried out during the operation of the power grid to form visual data content, which is convenient for operation and maintenance personnel to maintain and monitor the power grid. This paper takes the supervised power data mining method as the technical core based on the above. It formulates the evaluation system of power-data-diagnosis and voltage measurement capability based on big data to conduct technical evaluation and related visualization technology research.*

Keywords: *Transfer learning; Power data mining; SVM*

1. Introduction

1.1 Autoencoder

Power grid data mining mainly consists of the following steps. The first is to determine the data source of the power grid. Doctors, power grid workers, and researchers need to obtain this data through detailed conversations. The second is the preparation work to ensure that the power grid data can be preprocessed in advance, and the different types of data can be processed in a unified format to ensure that the data mining process can be carried out smoothly. In the process of data mining, the mathematical model is determined. Then the specific data content required by data mining is analyzed in detail to ensure that the process is reasonable and effective. The results of data mining have a certain value.

The formula for the autoencoder when the input samples are $(x_1, x_2, x_3, \ldots x_i)$ is as follows. Where S is the activation function, W_1 is the connection weight, W_2 is the communication weight, b_i is the bias term, and $h_{w,b}(x)$ is the output layer activation value [1].

$$d = S\left(W_1 x + b_1\right) \quad\quad (1)$$

$$h_{w,b}(x) = S\left(W_2 d + b_2\right) \qu\quad (2)$$

$$h_{w,b}(x) \approx x \qquad\qquad （3）$$

1.2 Recurrent Neural Network

According to the above model, the mathematical model of the recurrent neural network is derived as follows.

$$z_h^t = \sum_{i=1}^{I} w_{ih} x_i^t + \sum_{h'=1}^{H} w_{h'h} s_{h'}^{l-1} \qquad\qquad （4）$$

$$s_h^t = f_h\left(z_h^t\right) \qquad\qquad （5）$$

$$z_o^t = \sum_{h=1}^{H} w_{ho} s_h^t \qquad\qquad （6）$$

$$\hat{y} = f_o\left(z_o^t\right) \qquad\qquad （7）$$

Among them, x is the target data of the long-short-term memory network, y is the specified data of the long-short-term memory network, w is the weight of connecting each long-short-term memory network, z is the weighted sum of the long-short-term memory network, f is the activation function, s is the activation value of the long-short-term memory network, t represents the current moment of the long-short-term memory network, I is the dimension of the target data phasor, and i is the subscript of the target data phasor, h is the phasor subscript of the hidden connection part unit at time t, h' is the subscript of the hidden connection part unit phasor at time $t-1$, and o is the subscript of the specified data phasor [2].

The overall system can include three content layers: the power data receiving layer, the power data processing layer, and the power data transmission layer. The main functions of these three layers are as follows: the main function of the power data receiving layer is to accept basic power data information and pull the picture power data; the power data transmission layer accepts the above information and distributes it; The power data processing layer analyzes and decomposes the power data by the algorithm model.

1.3 The Network Training Method of Back-propagation Neural Network

The flowchart of the neural network training method of back-propagation is shown in Figure 1 below. First, the input data required for iteration is iterated and arranged according to the gradient descent rule, and then the chain rule is used to derive the data, and the data is reduced and optimized. After that, the BP formula is used to update and feed-forward the data to ensure that it is in an iterative update state. Finally, the gradient descent method updates the parameters to complete the training process.

```
┌─────────────────────────────────┐
│    training sample set input    │
└─────────────────────────────────┘
                 │
                 ▼
┌─────────────────────────────────┐
│      Iteratively update the     │
│     data according to the       │
│    gradient descent method      │
└─────────────────────────────────┘
                 │
                 ▼
┌─────────────────────────────────┐
│   Chained Derivatives Using     │
│   Compound Mathematical         │
│           Models                │
└─────────────────────────────────┘
                 │
                 ▼
┌─────────────────────────────────┐
│    Data Recursion Using BP      │
│           Algorithm             │
└─────────────────────────────────┘
```
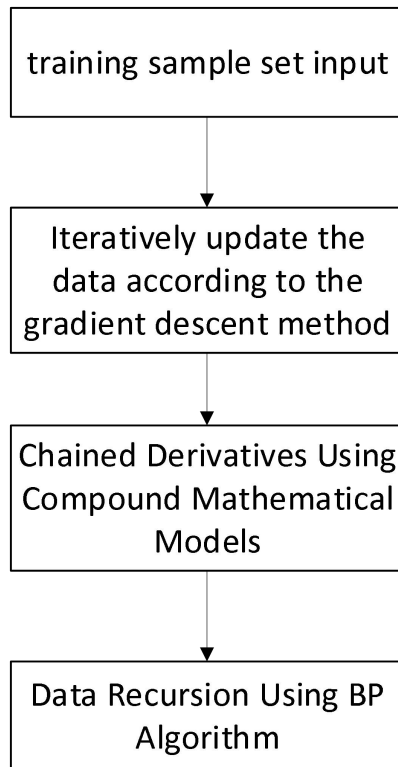
Figure 1 The training process of back-propagation neural network

The back-propagation neural network is used for feature extraction and transfers learning as a classic data algorithm. In this process, the back-propagation neural network completes related operations through a hierarchical algorithm. First, set the desired input and store it as input mode [3]. Then, the training sample information is sorted and injected into the back-propagation neural network in the form of tabular data, and this process is a forward process. When the expected value does not match the actual value, the network will apply the principle of the least square error to replace the data and enter the reverse process. The responses are weighted, and the path weights of the corresponding connection paths are modified according to the continuous feedback process of the error values in the forward and reverse processes. And the actual output of the network gradually approaches the preset expected value. The connection weights between each layer are output through the above learning process.

2. Grid Fault Detection Based on Self-encoder

2.1   Line Fault Detection Method of Sparse Auto-encoder

Based on the above research theories, the overall system of the sparse auto-encoder line fault detection method is summarized. The implementation process is as follows: The real-time communication server is used for data acquisition and analysis processing. The whole process relies on the power data center of the total relay protection equipment, and the obtained results are fed back to the power data center. After the data is processed, a carrier is needed for storage and backup, so the use of database management system software to manage power data in an orderly manner can greatly expand the storage capacity of power data. In the end, the technical difficulties of the client need to be solved. Therefore, in view of the advantages of extensive application, open-source, and cost-effectiveness of the power grid data server, the interaction of the power grid data server client is used to meet

the customers' needs for querying data. Based on the consideration of multiple factors of real-time, security, and economy, the three technical steps are matched with a server, respectively, and the free topology structure and on-site node decentralized autonomous method is used to process power data exchange macroscopically.

The line fault detection method of sparse auto-encoder is shown in Figure 2 below, and its specific mathematical model is as follows [4].
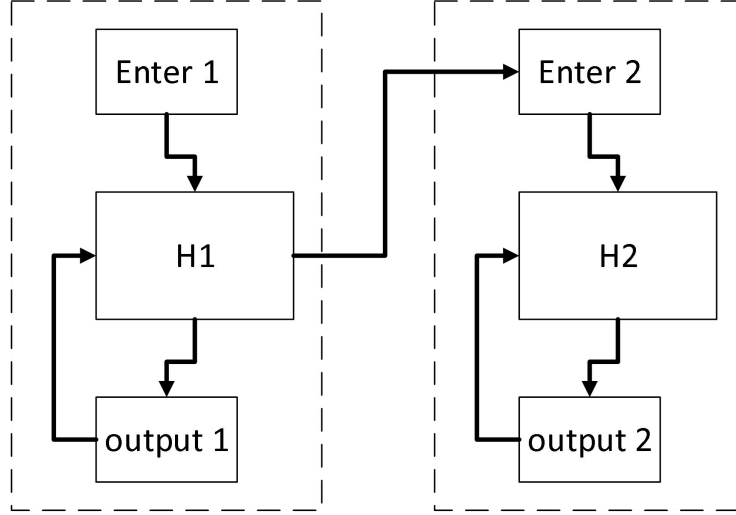

Figure 2. Flow chart of sparse auto-encoder line fault detection method

The feed-forward formula is as follows, where n_i is the number of permissions of the hidden neural network of the i-th layer, n_0 is the permissions number of the input layer neural network, Z is the weighted sum input of the hidden layer neural network permissions, and H is the hidden layer neural network permission value.

$$Z_1 = \left[z_1^{(1)}(X), z_2^{(1)}(X), \cdots, z_{n_1}^{(1)}(X) \right] = \left[\sum_{i=1}^{n_0} w_{1i}^{(1)} x_i, \sum_{i=1}^{n_0} w_{2i}^{(1)} x_i, \cdots, \sum_{i=1}^{n_0} w_{n_1 i}^{(1)} x_i \right] \quad (8)$$

$$H_1 = \left[h_1^{(1)}(X), h_2^{(1)}(X), \cdots, h_{n_1}^{(1)}(X) \right] = \left[f\left(z_1^{(1)}(X)\right), f\left(z_2^{(1)}(X)\right), \cdots, f\left(z_{n_1}^{(1)}(X)\right) \right] \quad (9)$$

$$= \left[f\left(\sum_{i=1}^{n_0} w_{1i}^{(1)} x_i \right), f\left(\sum_{i=1}^{n_0} w_{2i}^{(1)} x_i \right), \cdots, f\left(\sum_{i=1}^{n_0} w_{n_1 i}^{(1)} x_i \right) \right]$$

$$H_2 = \left[f\left(\sum_{j=1}^{n_1} w_{1j}^{(2)} f\left(\sum_{i=1}^{n_0} w_{1i}^{(1)} x_i \right) \right), f\left(\sum_{j=1}^{n_1} w_{2j}^{(2)} f\left(\sum_{i=1}^{n_0} w_{2i}^{(1)} x_i \right) \right) \cdots, f\left(\sum_{j=1}^{n_1} w_{n_2 j}^{(2)} f\left(\sum_{i=1}^{n_0} w_{n_1 i}^{(1)} x_i \right) \right) \right] \quad (10)$$

2.2 An Improved Fault Detection Model Based on Support Vector Machine

As a frontier branch of computer science, SVM technology has received extensive attention in the context of the current Internet big data. At present, the main research direction of SVM technology in China is to combine SVM with power data mining to effectively manage network information and prevent information loss. The main

technologies used in this process are machine learning, data mining, computer technology, computer management power system, and network server power system [5]. At present, the main research direction in this field is the combination of data clustering and power data mining. Therefore, this paper mainly focuses on applying SVM in power data mining to find a technology that combines effectively with SVM and data clustering. As the most commonly used protection power system for power data mining, the intrusion detection power system can identify and intercept internal and external interference to protect the power system in real-time. In this way, the abnormal behavior of the power system is discovered by collecting and statistical power data mining related data to detect the power data mining and the power system is disturbed. Intrusion detection power system is currently the main direction of domestic research, usually using the established normal server and related data between users to formulate abnormal activity data table to timely and effectively judge the situation of power system interference.

2.3 Algorithm Performance Test

Up to now, the method of detecting the entry of power data to the power data mining system is usually adopted, but this detection method is not efficient, prone to false alarms, and has low adaptability, so it is not suitable for the current network situation with a large amount of power data. Therefore, a new power data mining system is needed to effectively manage and prevent students' real-name power data [6].

In this process, clustering algorithms are usually used to perform multivariate analysis and statistics, which can be effectively classified by finding similarities between power data to form multiple adjacent digital power data sets. This type of learning usually uses observation to learn and train power data samples, from which the corresponding rules in the power data set can be deduced, and the clustering problem can be classified to evolve into an optimization problem.

Cluster analysis usually uses pattern analysis to classify power data, including supervised and unsupervised classification. Supervised analysis usually adopts the method of analyzing the known power data to analyze all the attributes of the power data. In the unsupervised analysis method, the clustering method is usually used to divide the power data set into a cluster to identify the degree of intimacy between them. This classification method does not require prior knowledge, so there is less demand for learning samples and a wider range of use.

3. Data-driven Fault Diagnosis of Power Grid Based on Transfer Learning Neural Network

3.1 Long Short-Term Memory Network Model

As a classic network security data processing algorithm in deep learning, the long-short-term power memory network has been widely used in the field of network security data processing and recognition. In this process, the long-short-term power memory network needs parameters compared with other neural networks. It can directly input raw network security data to carry out corresponding processing operations, so it has been widely and profoundly applied in this field. The long-short-term power memory network is used for feature extraction and transfers learning as a classical algorithmic neural network [7]. The neural network completes related operations through a hierarchical algorithm in this process. The entire learning process of a neural network is usually from forward to reverse. First, set the desired input and store it as input mode. Then, the training power data sample information is sorted and injected into the neural network in tabular form. This process is a forward process. At the

same time, during the learning process, the difference between the actual output and the expected output is calculated, stored, and recorded as the error value.

3.2 Fault Diagnosis Model Based on Transfer Learning

The fault diagnosis model based on transfer learning is shown in Figure 3 below. First, the electrical data is sampled, and then the fault data samples are preprocessed and input into the test data set for model training and testing. In this process, the unsupervised training process of electrical data is realized by dimensionality reduction and invisible features analysis. The trained data is subjected to support vector machine training, divided into normal and fault data information. Finally, measures are taken to the parsed data, and the results are output [8].
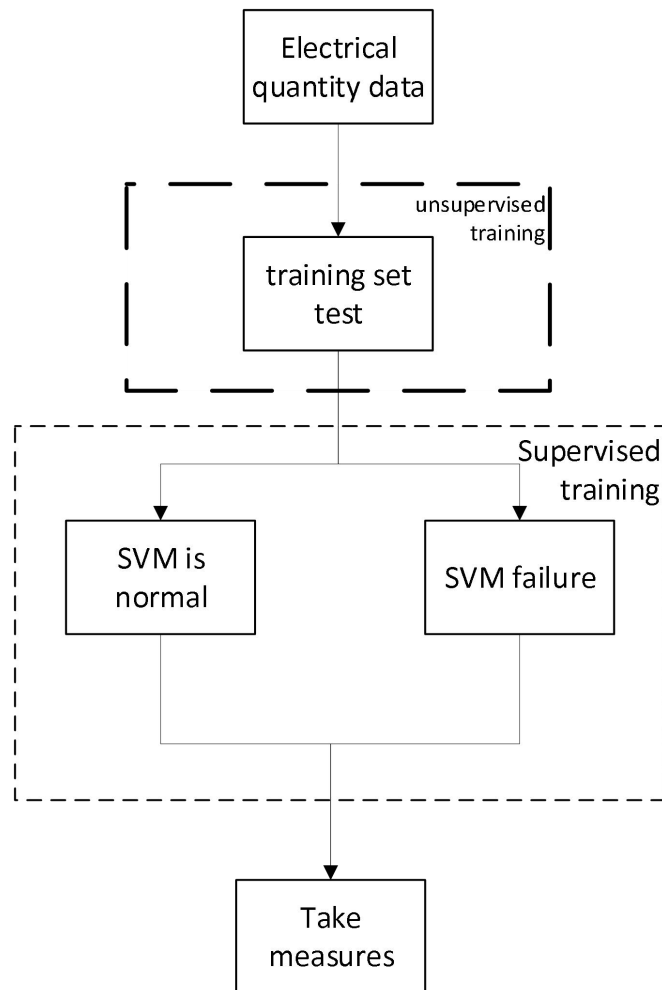

Fig. 3 Flow chart of fault diagnosis model based on transfer learning

3.3 Short-term Load Forecasting Method for Grid users based on GRU

The short-term load forecasting method model for grid users based on GRU is shown below.

$$y_p = \frac{y_{\max} - y_{\min}}{x_{\max} - x_{\min}}\left(x - x_{\min}\right) + y_{\min} \qquad （11）$$

Among them, y_p is the preprocessing result, $y_{\max}$ and $y_{\min}$ is the maximum and minimum value of the preprocessing result, $x_{\max}$ and $x_{\min}$ is the maximum value and minimum value of the unprocessed data. Where, $y_{\max}$ and $y_{\min}$ is set to 1 and 0.

The power grid data mining model is a multi-layer, feed-forward data analysis model. Its operation process is based on the algorithm error to conduct the reverse propagation algorithm research. This is a related data analysis model algorithm that is widely used. According to the error, reverse algorithm propagation can solve the above problems in the traditional power grid power data mining model by using the principal component guided analysis and training termination method [9]. In using the power grid power data mining model, there will be an over-fitting phenomenon, resulting in the feature not being fully extracted. However, applying the above-mentioned early termination method can prevent the feature from falling into the optimal local solution to solve the overfitting problem of the algorithm loading speed being too slow [10].

Power grid power data mining is usually a corresponding classification algorithm proposed for machine learning algorithms, and it is a binary linear supervised learning classifier. Make sure that the feature vector can completely cover its nodes. After designing its insensitive loss function, the support vector machine classification algorithm has gradually spread from the original classification problem to other nonlinear mining problems and can be used to solve the problem of event prediction sequences. At present, there are researches that can use grid power data mining to process grid power data grouping, which has a good development prospect.

4. Conclusion

This paper analyzes the causes of abnormal power data of power grid voltage, and at the same time uses comprehensive evaluation and big data visualization methods for the big massive data of the power grid, and conducts preprocessing and algorithm research on power grid data according to the core of data mining technology. The data is characterized by location mining and analysis processing. With the clustering method as the technical core, the relevant clustering effect indicators are discriminated using clustering interpolation and data preprocessing to formulate a power grid power operation analysis capability evaluation system based on big data, conduct technical evaluation, and carry out related research, and research on visualization technology. This paper establishes a complete set of power operation analysis capability evaluation systems through the basic structure of big data technology and smart grid data mining technology.

Reference

[1]. M. Jain, K. Gupta, A. Sathanur, V. Chandan and M. M. Halappanavar, "Transfer-Learnt Models for Predicting Electricity Consumption in Buildings with Limited and Sparse Field Data," 2021 American Control Conference (ACC), New Orleans, LA, USA, 2021, pp. 2887-2894.

[2]. S. Chan, I. Oktavianti and V. Puspita, "A Deep Learning CNN and AI-Tuned SVM for Electricity Consumption Forecasting: Multivariate Time Series Data," 2019 IEEE 10th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON), Vancouver, BC, Canada, 2019, pp. 0488-0494.

[3]. H. Xiaobo, D. Luping, C. Xiu, Y. Gen and X. Xiangkai, "A Predictive Model of Electricity Resident User Payment Based on Transfer Learning," 2021 6th Asia Conference on Power and Electrical Engineering (ACPEE), Chongqing, China, 2021, pp. 356-360.

[4]. L. Rodolph, M. S. Q. Zulkar Nine, L. Di Tacchio and T. Kosar, "Energy-saving Cross-layer Optimization of Big Data Transfer Based on Historical Log Analysis," ICC 2021 - IEEE International Conference on Communications, Montreal, QC, Canada, 2021, pp. 1-7.

[5]. L. Cai, J. Gu and Z. Jin, "Two-Layer Transfer-Learning-Based Architecture for Short-Term Load Forecasting," in IEEE Transactions on Industrial Informatics, vol. 16, no. 3, pp. 1722-1732, March 2020.

[6]. Y. Zhang et al., "FedNILM: Applying Federated Learning to NILM Applications at the Edge," in IEEE Transactions on Green Communications and Networking.

[7]. R. R. Bhat, R. D. Trevizan, R. Sengupta, X. Li and A. Bretas, "Identifying Nontechnical Power Loss via Spatial and Temporal Deep Learning," 2016 15th IEEE International Conference on Machine Learning and Applications (ICMLA), Anaheim, CA, USA, 2016, pp. 272-279.

[8]. M. Krishnamoorthy, K. Gunasekaran, B. R. T. Bapu, S. Balaji, V. Nagaraju and P. J. SathishKumar, "Bio Inspired FFA Algorithm for Efficient Data Transfer in WSN," 2022 4th International Conference on Smart Systems and Inventive Technology (ICSSIT), Tirunelveli, India, 2022, pp. 317-320.

[9]. D. L. Cavalca and R. A. S. Fernandes, "Deep Transfer Learning-Based Feature Extraction: An Approach to Improve Nonintrusive Load Monitoring," in IEEE Access, vol. 9, pp. 139328-139335, 2021.

[10].Jiang, N. Yan, B. Shen, C. Gu, H. Huang and H. Zhu, "Research on Lightweight Method of Image Deep Learning Model for Power Equipment," 2021 China International Conference on Electricity Distribution (CICED), Shanghai, China, 2021, pp. 334-337.