Research and Application of Server Health State Prediction Model Based on BP Neural Network

Qianjun Wu1,*, Dan Jin2, Yuhang Chen1 and Xiaolong Wang1

1Information System Integration Company, NARI Group Co., Ltd., Nanjing City, Jiangsu Province, China
2Gansu Tongxing Intelligent Technology Development Co., Ltd., Lanzhou City, Gansu Province, China

*Corresponding author

Abstract. As the core equipment of the data center, the importance of the operation and maintenance of the server is self-evident. In order to improve the operation and maintenance of data center server, this paper selected the feature quantity using the analysis method of grey correlation and established server operation parameters and health state prediction model using BP neural network algorithm in machine learning. The model can predict server health status and made fault early warning which provide decision support for operation and maintenance. This can reserve the responding time for problems and realize “active” operation and maintenance of data center servers.

1. Introduction
The rapid development of the company's business, business data has gradually accumulated, and the widespread use of load balancing has caused the number of servers to increase dramatically. The normal operation of the server is the fundamental guarantee for the normal operation of the business system, so the operation and maintenance of the server is particularly important [1]. The main method of server operation and maintenance is based on the monitoring of operating parameters and real-time status data of the server, so as to play a predictive role.

The server cannot work normally, and there is a possibility that the business system will crash. Operation and maintenance personnel can perform maintenance operations [2]. This is called "passive" operation and maintenance. Compared to "passive" operation and maintenance, "active" operation and maintenance can predict before the server fails. It allows time for operation and maintenance personnel to solve problems. On the one hand, it can ensure the continuous and stable operation of the service system carried by the server, and on the other hand, it can extend the service life of the server [3]. This paper uses machine learning algorithms to establish the relationship between server operating parameters and health status, establish a server health status prediction model, implement server failure early warning, and provide decision support for operation and maintenance personnel to implement "active" operation and maintenance of data center servers.

2. Related Technical Principles and Algorithm Flow

2.1. Grey Correlative Degree Analysis
The server operation will generate a large amount of data. This paper analyses the correlation between the data of each field of the collected data and the server health status data based on the grey correlative
degree analysis to determine the model input features\cite{4}. Grey correlative degree analysis is a new method developed on the basis of grey system theory. Correlative degree analysis is a method to analyse the degree of correlation of various factors in the system, or a method of measuring association. Through the comparison of the correlation degree, the primary and secondary factors can be found in the system, and the main factors affecting a certain variable are obtained. The algorithm flow is as follows:

Step 1: The dimensionlessness of the parameters.

Step 2: The correlation coefficient calculation.

Let $X_0 = \{x_0(1), x_0(2), \ldots, x_0(n)\}$ as be a data sequence, $X_i = \{x_i(1), x_i(2), \ldots, x_i(n)\}$ as a comparison sequence, the correlation coefficient between the two sequences is:

$$\xi_i(k) = \frac{\min_{k} \min_{i} |x_0(k) - x_i(k)| + \rho \max_{k} |x_0(k) - x_i(k)|}{\max_{i} \max_{k} |x_0(k) - x_i(k)|}$$

(1)

Where: $|x_0(k) - x_i(k)|$ is the absolute difference between the $x_0(k)$ and $x_i(k)$ of the point $k$, $\min_{k} \min_{i} |x_0(k) - x_i(k)|$ is minimum absolute difference, $\rho$ is the resolution coefficient, $0 < \rho < 1$ generally $\rho = 0.5$. The degree of association $\gamma_i$ between $X_i$ and $X_0$ is:

$$\gamma(x_0, x_i) = \frac{1}{n} \sum_{k=1}^{n} \xi_i(k) i = 1, 2, \ldots, n$$

(2)

Generally, the degree of association between the data column and the comparison sequence is small when the degree of association is less than 0.6.

2.2. BP Neural Network

BP neural network is a multilayer feedforward neural network trained according to the error back propagation algorithm\cite{5,6}. BP neural network has ability classification for arbitrarily complex pattern and mapping ability of excellent multi-dimensional function mapping ability. The BP neural network indirectly adjusts the weight of the hidden layer by obtaining the error of the output result and the expected output through the output layer\cite{7}. The learning process consists of two processes: forward propagation of the signal and backpropagation of the error: in the case of forward propagation, the input samples are passed in from the input layer, processed layer by layer through each hidden layer, and then transmitted to the output layer. If the actual output results of the output layer do not match the expected output, it goes into the backpropagation phase of the error\cite{8}. In the case of backpropagation, the output is back-transported to the input layer through the hidden layer, and the error is distributed to all the units of each layer, thereby obtaining the error signal of each layer unit, and the error signal is used as the basis for modifying the weight of each unit.

3. Construction and Training of Prediction Model for Server Health

3.1. Features Selection Based on Correlative Degree Analysis

The raw data fields collected by the company server are: [device type, operating system, device name, runtime, IP, CPU usage, memory usage, affiliated unit, business system]. The grey correlation degree is calculated one by one for the nine feature quantities (fields) and the server health status. The result is shown in Figure 1. It can be seen that the grey correlative degree of four feature quantities is less than 0.6 including of "Device Name, affiliated unit, business system, IP" which should be eliminated. There are five features after filtering and there are [device type, operating system, runtime, CPU usage, memory usage], and the five server operating parameters are used as model input features.
Figure 1. Gray correlative degree between parameters and status.

Figure 2. The principle of model construction.

3.2. Construction and Training of Server Operating Parameters and State Models based on BP Neural Network

Research and practice have proved that the 3-layer network can theoretically approximate any nonlinear function for BP neural networks, so the model selects a 3-layer network containing one hidden layer. The network hidden layer neuron transfer function uses the S-type tangent function, and the output layer neuron transfer function uses the S-type logarithm function. The number of hidden layer nodes affects the prediction accuracy. After screening and analysis one by one, the training sample error is the smallest when the number of hidden layers is 8; the number of hidden layer nodes continues to increase when the prediction error increases instead due to the poor network extrapolation ability caused by the over-fitting phenomenon. The model input is [device type, operating system, runtime, CPU usage, memory usage] and the model output is [health status]. The modelling principle is shown in Figure 2.

The processed 80% data was taken as the model training data set and 20% data was used as the model test data set. In order to ensure the stability of the model, the model training set and the test set are randomly assigned for training. The BP neural network prediction results are shown in Figure 3 and prediction error of BP neural network is shown in Figure 4.
The results show that most of the predicted output from the BP neural network is consistent with the expected output (actual value of the test data), and there are just a small number of deviation points. There is an analysis of the prediction error. At first, the error is defined as following:

$$\text{error}(i) = \text{pred}(i) - \text{real}(i), \quad i = 1, 2, \ldots, I$$  \hspace{1cm} (3)

Where $\text{pred}(i)$ is the predicted output and $\text{real}(i)$ is the expected output of the test data of $i$. The training accuracy rate is defined as: the number of accurate predictions divided by the total number of predictions (the number of test data). The training accuracy rate is 75.69%.

3.3. Improvement and Optimization

The training results show that there is improve performance of spatial for the prediction accuracy using BP neural network. The reason is that the convergence speed is slow, and it is easy to fall into the local minimum without obtaining the global optimal solution. The genetic algorithm is used in this paper to optimize the BP neural network. The characters of genetic algorithm is group search strategy and information exchange between individuals in the group. The search does not depend on the gradient information, and the algorithm is robust and not easy to fall into local optimum. After training test, the
training accuracy rate is equal to 85.42%, which is nearly 10% higher than that of BP neural network. As data training increases, the accuracy rate will be higher.

4. Conclusion
In this paper, a model of server operating parameters and health status is established by using BP neural network algorithm in machine learning. Rules relationship mining from a large data of historical operational data set are used to construct and optimize the server health state prediction model which realize the server fault warning. It can reserve the problem response time to provide decision support for the operation and maintenance personnel. The experimental results and practical applications show the high efficiency and practicability of the algorithm.

References
[1] Wen Xinxin, Qiu Yifan, Sun Jun. Analysis and Modeling of Component Prediction Model Based on Big Data and Poisson Distribution[J]. Computer and Digital Engineering, 2014, 42(8): 1412-1414.
[2] Peng Daogang. Modeling of Boiler Oxygen Content Based on Big Data and Neural Network[J]. Thermal Energy Engineering, 2018, 33(9): 86-92.
[3] Zeng Mingjie. Big Data Building Fire Risk Prediction Based on Machine Learning and Intelligent Modeling[J]. Modern Property: Mid-term, 2019, 446(1): 45-45.
[4] Liu Rui. Design of equipment status warning based on big data system[J]. Instrumentation Users, 2018, 25(1): 27-29.
[5] Dong Na, Liu Weina, Hou Botao. Modeling method of network anomaly behavior based on big data[J]. Electric Power Information and Communication Technology, 2018(1): 6-10.
[6] Li Chengdong. Building abnormal energy detection method and system based on big data and fuzzy model: CN106384300A[P]. 2017.
[7] Chen Jiawei. Research and application of support vector machine in human health state prediction [D]. University of Science and Technology of China, 2014.
[8] Su Xiaobo, Sun Mengmeng, Pan Erdong, et al. Battery Health State Assessment Based on Neural Network Modeling and K-Means Algorithm[J]. Value Engineering, 2016(2): 102-108.
[9] Jiafei Niu, Peiqing Niu. An intelligent automatic valuation system for real estate based on machine learning, Proceedings of the International Conference on Artificial Intelligence, Information Processing and Cloud Computing - AIIPCC, 2019.
[10] Liu, Rui Fang. Wind Power Generation Prediction by Particle Swarm Optimization Algorithm and RBF Neural Network, Advanced Materials Research, 2012.
[11] Tao Li, Liling Long. Imaging Examination and Quantitative Detection and Analysis of Gastrointestinal Diseases Based on Data Mining Technology, Journal of Medical Systems, 2019.
[12] Y Li, C Tang, L Xu, S Ye. Using System Dynamic Model and Neural Network Model to Analyse Water Scarcity in Sudan, IOP Conference Series: Earth and Environmental Science, 2017.
[13] He Yongxiang, Guo Hongwu, Zhou Bin. Nonlinear Modeling of Small Unmanned Helicopters Based on Double Neural Networks, Chinese Automation Congress (CAC), 2018.