Fault Prediction of Fan Gearbox Based on Deep Belief Network

Hanfeng Zheng[1], Yiru Dai[2]

[1]CIMS Research Center, Tongji University, Shanghai, 201800, China

[2]CIMS Research Center, Tongji University, Shanghai, 201800, China

*1399766691@qq.com

Abstract. Due to the large amount of data in the actual fan equipment failure, the external noise is complicated, and there is a high degree of nonlinearity and complexity, which makes it difficult to extract the fault features. If the model is constructed by the traditional method, the accuracy of the fault prediction is poor. Therefore, considering the advantages of deep learning in data feature extraction, this paper proposes a wind fault prediction method based on deep belief network (DBN). The original raw data is firstly deleted and normalized, and then imported into the DBN for training. The internal parameters of the network are adjusted by reverse learning to improve the feature extraction accuracy. Finally, the BP neural network is used to predict the fault. Comparing the prediction results with the SVRM method, we can find that the method has certain advantages in the fault prediction for the data.

1. Introduction

With the advancement of technology and the increasing demand of people, the scale of use of equipment has been continuously expanded, the level of technology integration has been continuously improved, and its value has been rising, resulting in high maintenance costs for faults, which has led to the loss of downtime caused by equipment failure. Huge [1]. Equipment failure prediction, as an important part of Prognostic and Health Management, has a major impact on maintenance acquisition strategies, spare parts supply, operating conditions, and system revenue. At present, the fault prediction methods for equipment can be roughly divided into two categories: one is based on the mechanism model of the device, and the other is based on data driving [2]. The mechanism model based on equipment considers the strain, stress, damage, etc. of the internal parts and mechanism structure of the equipment over time, using cumulative damage theory, Paris formula, fracture mechanics, crack propagation energy theory, etc. [3-5] Predict device failure. The data-driven approach can be divided into the following two depending on the method of use: based on statistics and machine-based learning. The statistical-based method fully considers the characteristics of historical accumulated data. For example, Baigen Cai et al. used the Weibull distribution model to analyze the statistical distribution of historical sample sequences, and summarized the sample regularity prediction faults [6]. Yanming Yang et al. used the reliability data analysis method based on exponential distribution to fit the different distributions of fault data to improve the accuracy and reliability of the analysis [7]. Machine learning is a computer that can self-learn from the available data without any explicit program. The learned model can respond to new data based on its past experience. Nadai, N et al. used a radial basis-based neural network model to predict the operating state of the equipment, providing effective support for later maintenance decisions [8]. Deep learning has the advantage of

being superior to the shallow machine learning algorithm [9], which can overcome the insufficiency of PCA and other statistical methods to identify faults [10]. Therefore, this paper decided to use the deep belief network to extract the features of the fault data of the wind turbine, and then predicted the faults of the extracted features through BP neural network.

2. Based on DBN predictive model
This article will be divided into three parts in this part to introduce the construction process and focus of the entire model.

2.1. Raw data processing:
According to the contents of the fault list provided by the analysis, it is found that the probability of failure of the fan caused by the gearbox failure is large, and in the gearbox fault data provided, the oil pressure of the oil port of the gearbox oil pump is basically different from the normal value. This failure prediction for the fan gearbox can be reflected by the oil pressure at the suction port of the gearbox oil pump.

Therefore, the fan data of the gearbox oil pump in the fan data used in this paper is the output target. The parameters that affect whether the fan is faulty include more than 50 currents and voltages of each phase. The above parameters are used as input variables of the model. Because this paper uses the daily operation data of the fan, the magnitude of the parameter data provided varies greatly, and there are some partial distortion phenomena, which are directly used as model input prediction, and the prediction effect is very poor. Therefore, this paper needs to pre-process the original data. After trying to compare several processing methods, the following methods are used to process the data: firstly, the missing data is filled, the local average is used for filling, and then the data is normalized by the operation. Scale between 0-1 to eliminate the difference between the magnitudes of the data. Secondly, the data set is further divided into a training set and a test set.

2.2. Restricted Boltzmann machine
In this paper, the deep belief network (DBN) is used to construct the prediction model, and the restricted Boltzmann machine (RBM) is the basic unit of the deep confidence network DBN. It is a generative stochastic neural network proposed by Hinton in 1986[11]. It is essentially an energy-based generation model and can be regarded as an undirected model graph model. There is no link between nodes, one layer is the visible layer, that is, the input data layer (v), and the other layer is the hidden layer (h), as shown in Figure 1.
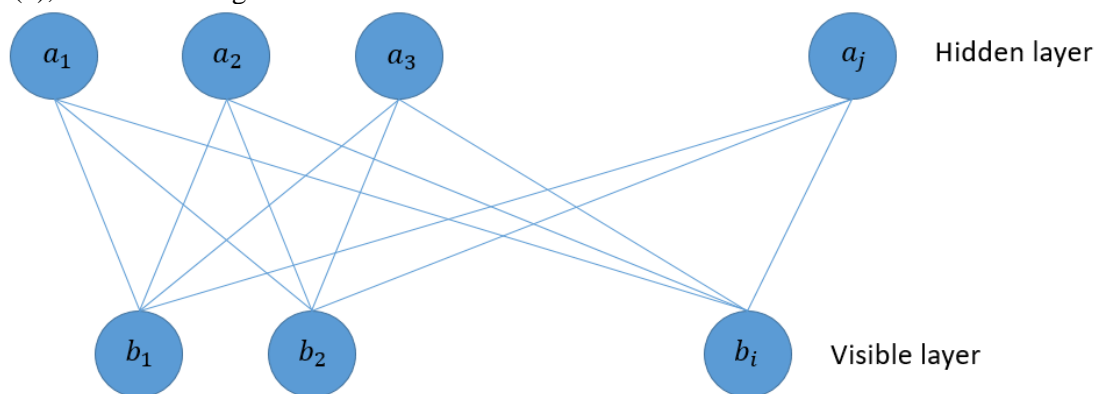

Figure 1. Network topology of the RBM

The network training for RBM is mainly for the bias b of its visible variable v, the offset a of the hidden variable h, and the optimal choice of the weight W between each layer. The energy formula for the joint configuration of the visible variable v and the hidden variable h is (1). Where θ is the parameter W, a, b of the RBM, W is the weight of the edge between the visible unit and the hidden

unit, and b and a are the offsets of the visible unit and the hidden unit, respectively. With the energy of the joint configuration of v and h, we can get the joint probability of v and h from formula (2), where $Z(\theta)$ in the equation is the normalization factor, also known as the partition function.

$$E(v,h \mid \theta) = - \left(\sum_{ij} W_{ij}v_i h_j + \sum_i b_i v_i + a_j h_j \right)$$

(1)

$$p_\theta(v,h) = \frac{1}{Z(\theta)} \exp(-E(v,h \mid \theta))$$

(2)

The training process of the entire RBM is to find a set of parameters, so that the joint probability distribution of RBM is the largest. The traditional training method is based on the Markov chain Monte Carlo method, but this method does not guarantee convergence. The current use of the CD-K method to train RBM ensures the fusion of files and increases speed and accuracy. The CD-k algorithm is essentially an improved algorithm that uses the training samples as the starting MCMC state. Usually only k (usually k = 1) steps Gibbs sampling is needed to obtain a sufficiently good parameter approximation. The steps of the corresponding CD-k algorithm are as follows:

①Given the input training sample X as the visible layer initial state v, the initial model weight W and the offset a, b, the given learning rate ε, the number of iterations training n.

②Calculate the activation probability of the hidden layer and the display layer according to formulas (3) and (4), respectively, where k=1, which requires cyclic sampling 3 times, and obtain corresponding outputs v^0, h^0, v^1, h^1 according to the input.

③According to the change of the weight W and the offsets a and b twice before and after, the three parameters are updated as in the formula (5) (6) (7).

④Cycle through steps ②-③ and continuously update the weights and offsets until the desired number of iterations n is reached.

$$h^k = p(h_j = 1 \mid v) = \sigma\left(a_j + \sum_i v_i W_{ij}\right)$$

(3)

$$v^{k+1} = p(v_i = 1 \mid h) = \sigma\left(b_i + \sum_j h_j W_{ij}\right)$$

(4)

$$W = W + \varepsilon(v^0 h^0 - v^1 h^1)$$

(5)

$$b = b + \varepsilon(v^0 - v^1)$$

(6)

$$a = a + \varepsilon(h^0 - h^1)$$

(7)

2.3. Building a DBN prediction model

The conventional DBN model can be regarded as a network composed of multiple RBMs interconnected. The specific model structure is shown in Figure 2. On the right side of the figure is a conventional DBN model structure. On the left is the corresponding RBM model layer. Each layer of DBN approximation can be considered as a single RBM, consisting of display layer v and hidden layer h, except that the model will be The hidden layer of one RBM layer serves as the display layer of the latter RBM (marked with a red box in the figure). During training, training data is entered from the display layer v_0 of RBM0. Then, the output of the hidden layer h_0 of RBM0 is used as the input of the display layer v_1 of RBM1. By analogy, RBM is unsupervised training from bottom to top, and the output of the hidden layer h_2 of the top RBM2 in the figure is an abstract representation of the input data, which can be approximated as a deep feature extraction of the original data.

After the training process is completed, the weight W of the entire network and the corresponding offsets b and a will be saved. At this time, for a single restricted Boltzmann machine, the offset a of the display layer v is no longer used, and the connection weight W is also changed from bidirectional

to unidirectional, as shown on the right side of figure 2, and its structural connection manner Same as forward neural network.
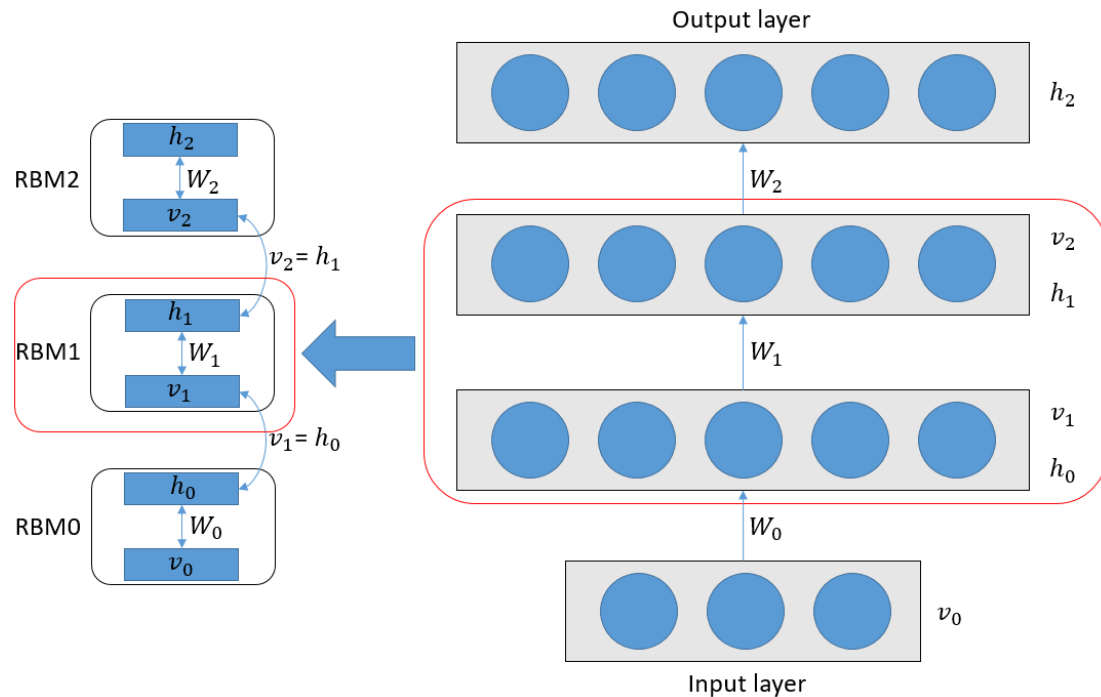

Figure 2. Network topology of the DBN

For the fault prediction problem of the fan gearbox involved in this paper, the diversity and complexity of the parameter dimension are considered. After a lot of training and testing, a suitable model is constructed, as shown in Fig. 3. The DBN model consists of three RBM layers, each with 250, 500, and 200 cells, and a BP neural network is added as the final regression output layer.
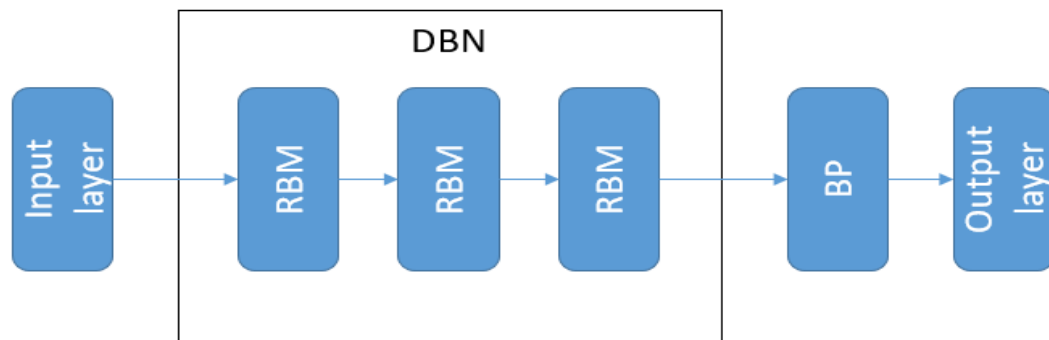

Figure 3. DBN+BP network structure

3. Formatting the text

The prediction model constructed in this paper is compared with the conventional SVRM prediction model. The RMSE and MAE values of the two results are used to reflect the pros and cons of the model. From the comparison of the indicators in the training set and the test set, the results are obvious. It can be seen that the prediction effect using DBN+BP is better.

Table 1. Training set and test set results

Method	Train		Test	
	RMSE	MAE	RMSE	MAE
SVRM	0.078	0.03	0.82	0.81
DBN+BP	0.018	0.013	0.17	0.11

In addition, it can be obtained from the error results of the two oil pressure prediction models for the oil pumping port of the gearbox shown in Figure 4. The orange line indicates the error of using DBN+BP to establish the prediction model, and the blue line indicates the error of using the commonly used SVRM to establish the prediction model. It can be seen from the two prediction curves that the error of the DBN curve is smaller.
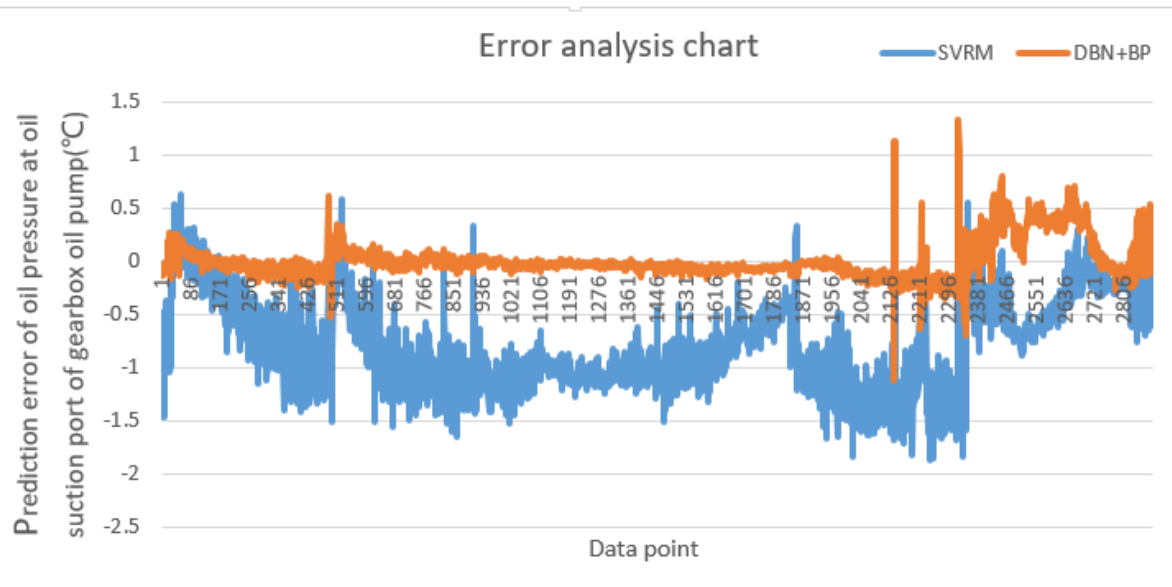

Figure 4. Error analysis

4. Conclusion
This paper describes some of the data problems that are often faced in actual wind turbine fault prediction and suggests solutions to these problems. Through the missing and normalized processing of the original data, the deep belief network is built to extract the fault features in the original data, and then the fault prediction of the fan gearbox is carried out. In the simulation, comparing the DBN+BP prediction model results with the commonly used SVRM prediction model results, it can be seen that the prediction effect after using DBN extraction features is good, and is significantly better than SVRM. However, it can be found from the error analysis graph that the error of the DBN+BP model is large in the later stage. In the later research, the DBN+BP model structure can be further improved to reduce the error in this aspect.

Acknowledgments

References
[1]   Lin S. (2013) Chemical Engineering and Equipment. J. Fault Diagnosis and Fault Control of Chemical Machinery. 09:84-86.
[2]   Heng A, Zhang S, Tan A.C.C, Mathew J. (2009) Rotating machinery prognostics: state of the art, challenges and opportunities. J. Mechanical System and Signal Processing. 23:724-739.
[3]   Yu J. (2010) Hidden Markov models combining local and global information for nonlinear and multimodal process monitoring. J. Journal of Process Control. 20(3):344-359.

[4]   Peng Y, Dong M. (2011) A hybrid approach of HMM and grey model for age-dependent health prediction of engineering assets. J. Expert Systems with Applications, 38(10):12946-12953.

[5]   Kobayashi K, Kaito K, Lethanh N. (2012) A statistical deterioration forecasting method using hidden Markov model for infrastructure management. J. Transportation Research Part B, 46(4):544-561.

[6]   Cai B.G, Shi X.Y, Wei S.G, Li Z.J. (2018) Statistical Distribution Analysis and Prediction of Time between Failures in On-board Train Control Equipment In: Proceedings 2018 Chinese Automation Congress. Xian. pp. 767-772.

[7]   Nadai N, Melani A.H.A, Souza G.F.M., Nabeta S.I. (2017) Equipment failure prediction based on neural network analysis incorporating maintainers inspection findings In: Proceedings - Annual Reliability and Maintainability Symposium. Orlando. pp. 7

[8]   Yang Y.M, Zheng H.Y, Zhang R.L (2017) 2017 2nd IEEE International Conference on Computational Intelligence and Applications. Beijing. pp. 567-571

[9]   Sun Z.J, Xue L, Xu Y.M. et al. (2012) Review of deep learning research. J. Journal of Computer Applications, 29(8): 2806-2810

[10] Zhou F.N, Gao Y.L, Wang J.Y, Wen C.L. (2017) Early diagnosis and life prediction of slowly changing faults based on deep learning. J. Journal of Shandong University (Engineering Science). 47(05):30-37.

[11] Hinton G.E, Salakhutdinov R.R. (2006) Reducing the dimensionality of data with neural networks. J. Science 313(5786):504.