Research on Fault early Warning Technology of Wind Turbine Main Bearing by Stacked Auto Encoder

Yuhao Fan*, Chenying Qian
Institute of electrical engineering, North China Electric Power University, Baoding 071000, China

*Corresponding author: 201801001505@ncepu.edu.cn

Abstract. With the rise of new energy power generation technology, the installed capacity of wind turbines continues to increase. At the same time, the potential faults of wind turbines have also increased with the increase of wind turbines. Therefore, early prediction of potential faults of wind turbines and ensuring the safe and stable operation of wind turbines is of great significance for improving power generation efficiency and reducing maintenance costs. In order to realize the fault early warning of the main bearing of the wind turbine, an early warning method of the main bearing of the wind turbine based on Stacked Auto encoder (SAE) is proposed.

Keywords: wind turbine; gearbox; main bearing; deep learning; fault early warning; stacked auto encoder.

1. Introduction

SAE models are stacked by multiple AE models. The AE model is similar in type to the traditional neural network model, but it differs from the traditional neural network in two ways: one is that the dimensions of the input layer and the output layer are identical; second, the implicit layer in the middle is lower than the dimension of the input layer. AE uses unsupervised learning, and the optimization goal during training is to make the values of the output layer as close or the same as those of the input layer as possible. The basic idea of AE is that the data of the input layer dimension represents certain information, and it is represented by the hidden layer of the smaller dimension in the middle by encoding. At this time, the dimension is compressed, and finally, the information represented by the middle hidden layer is restored to a data of the same dimension which is close to or the same as the original information by decoding. Therefore, the stacked auto encoder extracts the deep features of the sample data layer by layer, and transforms the scattered data into the intrinsic features that can deeply describe the sample data. To realize the function of fault early warning, first, the time domain characteristics of the vibration data of the main bearing of the wind turbine are extracted as data samples which directly reflect the operating states of the main bearing. Secondly, a fault early warning model for the main bearing of the wind turbine is constructed, and detailed steps for fault early warning are given; Finally, this method is tested through an engineering example, and the effect of the traditional neural network model on the early warning of the wind turbine main bearing is analyzed. The results show that the method proposed in this paper has better performance, higher accuracy and early warning of failures. The basic structure of SAE is shown in Figure 1.
2. Organization of the Text

2.1. Pre-training for SAE

Assuming that there are \{(x(1), y(1), (x(2), y(2)), ..., (x(m), y(m))\}, a total of M Training sets, the refactoring error is defined as:

$$
E(w, b; x^{(i)}, y^{(i)}) = \frac{1}{2} \| h_{w,b}(x^{(i)}) - y^{(i)} \|^2
$$

in which \(w\) is the weight of the network; \(b\) is the threshold of the network; \(h_w, b(x(i))\) is the reconstructed value of the network; \(x(i)\) equals \(y(i)\) is presented.

Define the loss function as

$$
J(W, b) = \left[ \frac{1}{m} \sum_{i=1}^{m} \left( \frac{1}{2} \| h_{W,b}(x^{(i)}) - y^{(i)} \|^2 \right) \right] + \frac{\lambda}{2} \sum_{l=1}^{\ell-1} \sum_{i=1}^{n_l} \sum_{j=1}^{n_{l+1}} (W_{ij}^l)^2 + \beta \sum_{j=1}^{p_l} KL(p||\hat{p}_j)
$$

The first is the constraint of auto encoder; The second item is a penalty item to prevent overfitting. The third is the sparse constraint, which is a measure of KL dispersion.

$$
KL(p||\hat{p}_j) = p \log \frac{p}{\hat{p}_j} + (1 - p) \log \frac{1 - p}{1 - \hat{p}_j}
$$

In order to minimize the loss function \(J_{\text{sparse}}(w, b)\), which contains \(w\) and \(b\), the gradient descent method is used to solve \(w\) and \(b\).
As for the gradient, it can be solved by the backpropagation (BP) algorithm, where, $\alpha$ for learning rate; $b_{i}^{(l)}$ is the 1st neuron threshold in the 1st layer.

### 2.2. Fine Tuning and Mean Drift Clustering Algorithm for SAE to Find Running Baseline

After pre-training the self-coded network with labeled data, the initial weights for each layer will be better placed in the parameter space than in random initialization. From these locations, the weights are further adjusted using the BP algorithm using tagged data. Since unlabeled data already provides a lot of prior information about the patterns contained in the input data, it is likely that the gradient descent from these locations will converge to a better local pole.

The mean drift clustering algorithm is based on a sliding window algorithm, using distance as a similarity index, moving the centroid along the direction of density rise through iterative calculation to find the cluster center. Therefore, the mean drift clustering algorithm is applied to SAE training, and the cluster centers of the reconstruction errors of the main bearings of wind turbines in stable operation state $(\delta_{t})$ can be iteratively calculated, and used as the benchmark value to warn the faults. Comparing the mean value of reconstruction error ($E_{av}$) of the field wind turbine output with the reference value $(\delta_{t})$ of stable operation state, considering both the sensitivity of early warning and the possibility of false warning, it is considered that when $E_{av}$ deviates from $\delta_{t}$ more than 60% and has a tendency to increase gradually, it will become a failure and vice versa, it will become a stable state.

### 2.3. Selection and processing of sample data

As the vibration signal is the most basic, the lowest level and the most able data to reflect the operation status of the mechanical equipment, the horizontal vibration signal of the main bearing of the wind turbine is selected as the source of the sample data first. In order to explain the relationship between vibration signal and operation status more simply and accurately, the time domain characteristic mean square value, peak value, margin, skewness and steepness of vibration signal are used as the characteristic data of wind turbine bearing status. At the same time, based on the structure and operation principle of wind turbines, considering the physical quantity related to the main bearing status of wind turbines and the complexity of the model and training time, a variable filtering method based on Mean Impact Value (MIV) is used. Six kinds of SCADA data (wind speed, cabin temperature, temperature of the main bearing on the rotor side, temperature of the main bearing on the gear box side, speed of the generator) are chosen to reflect the running status of the main bearing. In order to ensure the quality and quantity of sample data and avoid the decline of early warning effect due to data imbalance when selecting the data of the normal operation of the unit, the SCADA data of the main bearing of the wind power unit are selected evenly by season and wind speed and the unreasonable data is eliminated. There are large differences in these data, in order to reduce the imbalance of input data and reduce the calculation error, normalization of each data to $[0,1]$ is processed to improve the diagnostic accuracy.
2.4. Network Model Parameter Settings
The SAE parameters w and B are initialized to small random values with a Gaussian distribution, initial learning rate $\alpha$ set to 0.1, update rate to 0.01, maximum number of iterations to 1000, batch training to 100, other network parameters to default. In order to reduce the computational load, combined with the results from previous tests that the effect on network performance is not obvious when the number of neurons is similar, 10 is used for neuron magnitude 10-100 test interval and 100 is used for neuron magnitude 100-1000 test interval.

2.5. Project Example Analysis
In this paper, the SCADA on-line monitoring data of two wind turbines are taken from a wind power generator set 44 and 45 in a wind farm in China. The feature data of the main bearing status of the wind turbines are obtained through the above data filtering and processing methods. From February 2017 to February 2018, stable operation states of the wind turbines in the early 1 year are selected. A total of $11 \times 2800000$ feature data were used for pre-training and fine-tuning of SAE, and $11 \times 600000$ feature data were selected for a period of approximately 700 days between March 2018 and downtime due to faults, which are used as test data.

2.6. Failure Alert Test
Due to the constraints of different models, differences in geographic locations, external environment and other factors, different wind turbines take different reconstruction error benchmarks. In order to ensure the accuracy of early warning, the mean clustering algorithm is used to iteratively calculate the reference values of reconstruction error of No. 44 and No. 45 wind turbine, which are both in stable operation state. The reference value goes to 0.287 and 0.308 respectively. Then, these two wind turbines are tested for fault warning using the characteristic data of wind turbines, which has been collected for more than two years. Through the analysis of the feature data by SAE fault warning model of the main bearing of the wind turbines, the average value of the reconstruction error of the main bearing of the wind turbines is obtained, and the fault situation is predicted by comparing the relationship between the average value of the reconstruction error and the reference value of the reconstruction error.

The main bearing fault alarm of No. 44 wind turbine is shown in Figure 2. The average value of main bearing reconstruction error of No. 44 wind turbine represents the operation condition of this main bearing for more than two years. As shown in Figure 2, it fluctuates slightly near the reference value of reconstruction error and does not deviate significantly from the reference value of reconstruction error. Therefore, it is predicted that the main bearing of No. 44 wind turbine will run well. Unit 44 has been running steadily in the actual site, and no failure has been found.

![Figure 2. Fault warning of No. 44 wind turbine main bearing](image-url)
The main bearing fault alarm of No. 5 wind turbine is shown in Figure 3. It can be seen from the diagram that in the past more than one year, the average value of the reconstruction error of the main bearing of the wind turbine fluctuates slightly near the reference value of the reconstruction error, and the main bearing of the wind turbine is in normal operation. However, in the following period of time, it can be seen that the average fluctuation range of the reconstruction error of the main bearing of the 45 unit wind turbine increases gradually, and there is a gradual upward trend on the 60th day when it exceeds 60% of the reference value of the reconstruction error, so it is predicted that the main bearing of the wind turbine may fail. On the 700th day in the actual site, there was no grease fault in the main bearing roller holder and raceway space of the wind turbine, slight tilt of the roller and 3-4mm shaking of the roller, and there was a ring friction mark fault on the side of the roller body, and the machine stopped at this time. However, before the 700th day, the actual main bearing of the field turbine did not issue any fault alarm by the general fault diagnosis measures, so the operation status of the main bearing of the wind turbine set is reflected through the average of the reconstruction error, and the fault alarm can be given in advance according to its changing trend.

![Figure 3. Fault warning of No. 45 wind turbine main bearing](image)

### 2.7. Performance Analysis of Failure Alert

For other fail-free units of the same type in this wind farm, the warning results are similar to those of Unit 44, and those of Unit 45 for failures. In addition, 30 sets of fault early warning tests with different capacities, different types of failures and normal operation data of wind turbines are carried out. In order to verify the universality of the fault early warning methods, the selected fault units are the failures of different components, including gear box failures, bearing failures and blade failures. The proportion of samples of normal and faulty wind turbines is 2:1. Meanwhile, the BP or SVM based fault warning method for wind turbines is tested with the same samples of wind turbines. The test results of different algorithms are shown in Table 1.

| Algorithm | Fault warning time before the accident | Accuracy/% |
|-----------|--------------------------------------|------------|
| SAE       | About 30                             | 96.7       |
| BP        | About 8                              | 76.7       |
| SVM       | About 10                             | 83.3       |

By comparing the fault warning tests of different algorithms in the table, it is known that the SAE-based fault warning method for wind turbines can warn the faults earlier than the BP or SVM-based
fault warning method for wind turbines, and has a higher warning accuracy. Compared with other early warning algorithms, the deep learning-based network structure used in this method has significant advantages in dealing with large data. At the same time, in the selection of data, this method takes into account the various time domain characteristics of the vibration data from the wind turbine bearing bottom layer, and improves the warning effect. And for some potential and undetectable faults, such as wear and tear, aging of parts, they can accurately warn of faults on the basis of better model training.

3. Conclusion
For the problem of fault warning of main bearing of wind turbines, this paper puts forward a fault early warning method based on SAE. This early warning method overcomes the shortage of the conventional fault diagnosis method, which can only diagnose the fault of the bearing after an accident. The SCADA data of the main bearing can be processed and analyzed through the SAE fault warning model, in which the running state of the wind turbine bearing is reflected on the average value of the reconstruction error. And the fault can be predicted before it occurs when the average value of the reconstruction error is compared with the mean drift clustering algorithm. Therefore, this method can play a significant role in preventing the fault from occurring as well as the expansion of the fault. It also has been analyzed and tested with an engineering example, the result shows that the SAE-based method is feasible to alarm the fault of main bearing of wind turbines in advance.

References
[1] R. K. Ibrahim, S. J. Watson, S. Djurović and C. J. Crabtree, "An Effective Approach for Rotor Electrical Asymmetry Detection in Wind Turbine DFIGs," in IEEE Transactions on Industrial Electronics, vol. 65, no. 11, pp. 8872-8881, Nov. 2018, doi: 10.1109/TIE.2018.2811373.
[2] P. F Odgaard, J. Stoustrup and M. Kinnaert, "Fault-Tolerant Control of Wind Turbines: A Benchmark Model," in IEEE Transactions on Control Systems Technology, vol. 21, no. 4, pp. 1168-1182, July 2013, doi: 10.1109/TCST.2013.2259235.
[3] B. Yang, R. Liu and X. Chen, "Fault Diagnosis for a Wind Turbine Generator Bearing via Sparse Representation and Shift-Invariant K-SVD," in IEEE Transactions on Industrial Informatics, vol. 13, no. 3, pp. 1321-1331, June 2017, doi: 10.1109/TII.2017.2662215.
[4] R. Jiao, X. Huang, X. Ma, L. Han and W. Tian, "A Model Combining Stacked Auto Encoder and Back Propagation Algorithm for Short-Term Wind Power Forecasting," in IEEE Access, vol. 6, pp. 17851-17858, 2018, doi: 10.1109/ACCESS.2018.2818108.
[5] R. Karki, Po Hu and R. Billinton, "A simplified wind power generation model for reliability evaluation," in IEEE Transactions on Energy Conversion, vol. 21, no. 2, pp. 533-540, June 2006, doi: 10.1109/TEC.2006.874233.
[6] K. Leahy, R. L. Hu, I. C. Konstantakopoulos, C. J. Spanos and A. M. Agogino, "Diagnosing wind turbine faults using machine learning techniques applied to operational data," 2016 IEEE International Conference on Prognostics and Health Management (ICPHM), 2016, pp. 1-8, doi:10.1109/ICPHM.2016.7542860.