Research on Equipment Fault Prediction Method Based on Industrial Big Data

Zihan Qi [1], Chang Tang [2], Luli He [3], Changjian Jiang [4], Jing Chen [5*], Yulin Huang [6, f]

[1]1569587263@qq.com

[2]1297605942@qq.com

[3]heluli17353605116@163.com

[4]2958831156@qq.com

[5*]Corresponding author: jingchen94@163.com

[6]hyl7423@163.com

[1] Department of Applied Statistics, Qilu University of Technology (Shandong Academy of Sciences), Jinan, Shandong, China

[2] Department of Applied Statistics, Qilu University of Technology (Shandong Academy of Sciences), Jinan, Shandong, China

[3] Department of Applied Statistics, Qilu University of Technology (Shandong Academy of Sciences), Jinan, Shandong, China

[4] Department of Applied Statistics, Qilu University of Technology (Shandong Academy of Sciences), Jinan, Shandong, China

[5] Shangdong Computer Science Center (National Supercomputer Center in Jinan), Qilu University of Technology (Shandong Academy of Sciences), Jinan, Shandong, China

[6] Department of Applied Statistics, Qilu University of Technology (Shandong Academy of Sciences), Jinan, Shandong, China

Abstract—Aiming at the problems of low accuracy and long prediction time of industrial equipment fault prediction, a prediction method based on main feature extraction and back propagation neural network (MFE-BPNN) was proposed. This method firstly preprocesses the missing, abnormal and high-noised industrial equipment data, then uses the method of recursive feature elimination combined with cross validation to extract the main feature variables, then designs the numbers of hidden layers and neurons, and weights of training and learning rates. This method improves the accuracy of industrial equipment fault prediction by preprocessing industrial data and establishing a prediction model based on a neural network. The prediction time is reduced by extracting the main characteristic variables. The experimental results of fan blade icing fault prediction in the field of power generation verify the effectiveness of this method.

Keywords-industrial data; data preprocessing; main feature extraction; equipment fault prediction; BPNN

ICIDC 2022, June 17-19, Qingdao, People's Republic of China

DOI 10.4108/eai.17-6-2022.2322795

1 INTRODUCTION

Industrial equipment failures often cause enterprises huge losses. With the advent of industry 4.0, the industrial devices of intelligent factories are equipped with various sensors to collect data such as vibration, temperature, current and voltage. By analyzing these data and adopting the corresponding algorithms, the time of potential equipment failure can be predicted, and the early warning and treatment can be carried out in advance, so as to reduce the loss caused by equipment shutdown. However, with the rapid growth of industrial equipment data, there are often problems such as missing, invalid, abnormal values, high noise, excessive characteristic parameters and unbalanced data, which often lead to the further problems of low prediction precision and long prediction time. Thus, it is difficult to accurately and quickly predict failures and carry out operation and maintenance in advance.

The research on the equipment fault prediction at home and abroad mainly includes prediction methods based on time series and those based on machine learning models. The commonly used time series prediction models include autoregressive moving average (ARMA) model and autoregressive integral moving average (ARIMA) model. Erdem et al. [1] proposed four prediction methods based on the ARMA model for wind direction and speed prediction. Zhang et al. [2] used the ARIMA model to predict the concentrations of PM2.5 in Fuzhou, China. Although the ARIMA model has been applied in many fields and can well describe the linear relationship in time series, there are many nonlinear problems in real life, and there are limitations to solve them by linear approximation.

Machine learning provides effective methods for nonlinear prediction problems, such as neural network [3], Gaussian process regression (GPR) [4], decision tree [5,6], logical regression [7], multilayer perceptron (MLP) [8]. Zong et al. [9] applied multi-layer perceptron (MLP) neural network to predict the energy demand of South Korea and compared it with several models. The experiments show that MLP has good prediction precision. Liu et al. [10] established a multi-variable prediction model based on GPR and multiple imputations. The model can well predict the time series with the missing data. Guo Yu and Yang Yu [11] proposed the equipment fault prediction based on the grey rough set and back-propagation neural network (BPNN). BPNN is also combined with empirical mode decomposition (EMD) to be used for the cavitation damage fault diagnosis of hydro-power units [12] and wind power prediction [13]. Zhou Guangfei [14] proposed to apply the multivariate statistical analysis method to fan blade icing fault prediction. In addition to using dominant features significantly related to icing, Zhou also constructed a series of invisible features, which greatly improved the prediction effect. However, this greatly increased the prediction time.

In this paper, an industrial equipment fault prediction method based on the extraction of main feature variables and the back propagation neural network (MFE-BPNN) is proposed. The data is preprocessed through data screening, data down-sampling, singular value deletion and feature selection. The main feature variable are extracted by combining the recursive feature elimination method for cross-validation (RFECV). Taking the icing fault of fan blade in power generation field as an example, the effectiveness of the model is verified and evaluated by comparing with the decision tree and logistic regression classification algorithm.

2 MFE-BPNN PREDICTION METHOD

2.1Data Description and Preprocessing

The data in this paper sources from the SCADA system of domestic wind power plants. About 390,000 data and 26 characteristic variables were collected in real-time operation of the fans for two months in chronological order. The data type is floating point. The names of variables are shown in Table 1:

TABLE 1 FAN ICING DETECTION VARIABLES

Number	Variable Name	Number	Variable Name
1	Wind speed	14	Temperature of variable pitch motor 1
2	Engine RPM	15	Temperature of variable pitch motor 2
3	Power	16	Temperature of variable pitch motor 3
4	Paired wind angle	17	Acceleration in X-axis direction
5	Average wind direction angle	18	Acceleration in Y-axis direction
6	Yaw position	19	Ambient temperature
7	Yaw speed	20	Engine room temperature
8	Blade angle 1	21	Temperature of No. 1 ng5
9	Blade angle 2	22	Temperature of No. 2 ng5
10	Blade angle 3	23	Temperature of No. 3 ng5
11	Blade angle 1	24	Direct current of No. 1 ng5 charger
12	Blade angle 2	25	Direct current of No. 2 ng5 charger
13	Blade angle 3	26	Direct current of No. 3 ng5 charger

There are many invalid data in the data set; the time is not continuous; the data has many singular values; and the equipment status data is unbalanced. In this paper, several data preprocessing methods such as data screening, data down-sampling, singular value deletion and feature selection are used to preprocess the data to optimize and improve the prediction effect and efficiency of the model. The specific process is shown in Figure 1:

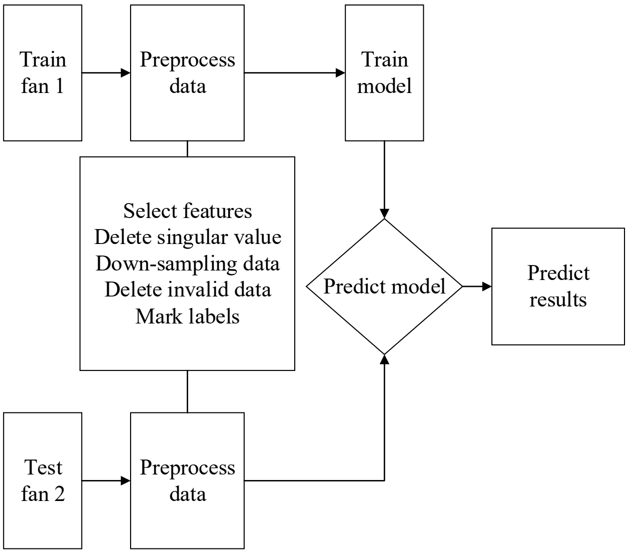


Figure 1 Ice Detection Process of Fan Blades

2.1.1Invalid Data Deletion and Down-sampling

There are three data set files: data.csv is the original data set; failure.csv is the data of fan during the fault icing; normal.csv is the data of fan during the normal operation. The data excluded in normal time period and icing time period of the fan are invalid data, which shall be filtered and deleted.

2.1.2Exclusion of Singular Values

Due to the noise and other factors in the data set that will cause the appearance of singular values, this paper uses the box plot method, i.e., quantiles to delete singular values. Basic principle and steps are as follows:

(1) Sort the sample data according to a variable from small to large;

(2) Calculate the upper quartile $Q2$, median and lower quartile $Q1$ of the selected variable respectively;

(3) Use the formula(1) to calculate the upper and lower limits:

$$U = Q2 + R \cdot \Delta, L = Q2 - R \cdot \Delta \qquad (1)$$
$$\Delta = Q1 - Q2$$

where R is the control limit. When $R = 1.5$, the singular value obtained when the variable value exceeds the upper or lower limit is called mild singular value. When $R = 3$, the singular value obtained is called extreme singular value.

(4) Delete the extreme singular value samples of the variable;

(5) Calculate the next selected variable and repeat the first step until all variables are verified. The box plot of feature variables is shown in Figure 2:

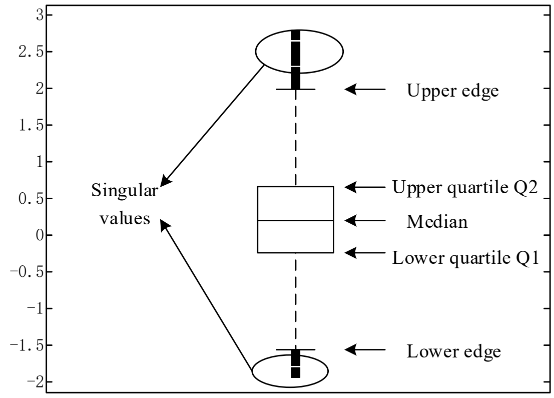


Figure 2 Box Plot

2.1.3Selection of Main Feature Variables

From the sample data, there are up to 26 variables of fan blades. If all variables are used for the model training, it will not only increase the calculation of data, but also will be difficult to accurately find the characteristic source of fan icing. If the features easy to classify can be extracted from the sample data, the fan with the icing fault can be detected effectively and accurately. This paper uses RFECV for feature selection. RFECV consists of the recursive feature elimination (RFE) and the cross-validation (CV). Firstly, RFE is used to sort the feature variables according to the given weight. The feature variables with low weight are excluded to select the important feature vectors, and then CV is used to optimize and select the best feature combination.

Phase I -- RFE phase

There are 26 feature variables in the initial data set. The current feature set is modeled and the importance of each feature variable is calculated. The least important feature variables will be deleted, and then the feature set is updated. The importance of the new feature collection is calculated until the importance rating of all feature variables is completed.

Phase II -- CV phase

According to the feature importance calculated in the RFE phase, different numbers of features are selected in turn. Then the selected feature set is cross-validated to finally determine the number of features with the highest average score, thus completing the feature selection. The output results are shown in Figure 3.

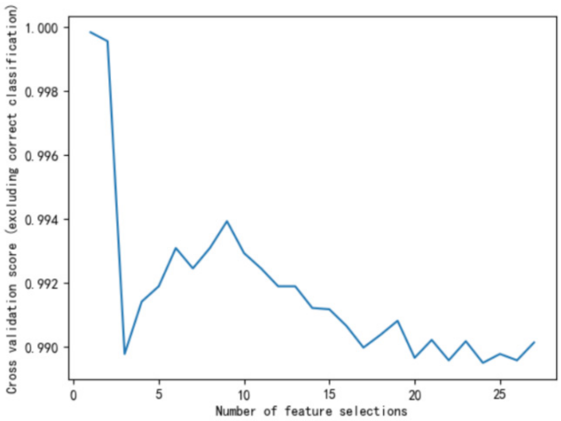


Figure 3 Feature Selection

As can be seen from Figure 3, the score drops dramatically in the wake of the third variable, which can ensure greater accuracy when selecting the top three variables. Therefore, the top three feature variables are selected to reduce the prediction cost.

2.2Fault Prediction Model of Industrial Equipment Based on BPNN

The BP neural network is a multilayer feedforward neural network trained according to the error back-propagation algorithm. Signals are input by the input layer, and then the signal is output by the output layer after being calculated by one or more hidden layers. The output values are compared with the expected value. If the two do not match (i. e., there is an error), the error is back propagated from the output layer to the input layer (error back-propagation). In the process of error back-propagation, the gradient descent algorithm is used to adjust the neuron weight. After a lot of repeated learning and training, the weight and threshold corresponding to the minimum error are finally determined, and the training stops.

The BP neural network has strong nonlinear mapping ability, which is conducive to solving the problem of complex internal mechanism of fan icing. BP neural network has high self-learning and self-adaptive ability and strong generalization ability, which can effectively determine whether the fan has fault and classify it correctly. BP neural network has certain fault tolerance ability. After its local neurons are damaged, it has little impact on the global training results, that is, it can stably judge the fan fault. If BP network has enough hidden nodes and hidden layers, it can approach all nonlinear mapping relations and fully reflect its generalization ability. See Figure 4 for details:

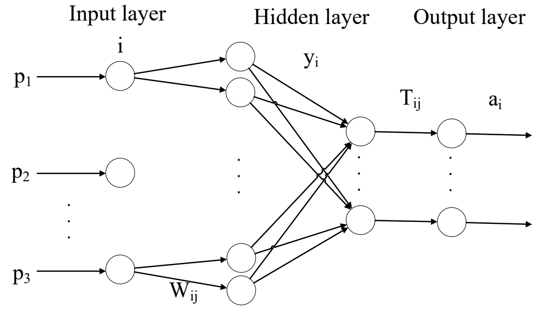


Figure 4 Structure of BP Artificial neural network

Let the input and output values of the BP artificial neural network be $(p_1, p_2, \dots p_m)$ and $(a_1, a_2, \dots, a_m)$ respectively. Then, the threshold is θ_i, and the connection weights between the nodes of the input domain and the output domain are W and T respectively.

Let the target value of the BP artificial neural network be t_i. Then the output value of the hidden layer of the BP artificial neural network is(shown in formula(2)):

$$y_j = f\big(\Sigma_i\big(w_{ij}p_i - \theta_j\big)\big) = f(net_i) \tag{2}$$

The output value of the output layer is(shown in formula(3)):

$$A_m = f\big(\Sigma_i\big(T_i y_j - \theta_j\big)\big) = f(net_i) \tag{3}$$

The error of the output layer is(shown in formula(4)):

$$\begin{aligned} E &= \frac{1}{2}\sum_I (t_i - a_m)^2 = \frac{1}{2}\sum_{\mathrm{i}} \left(t_i - f\left(\sum_j \left(T_i y_j - \theta_j \right)_j \right) \right)^2 \\ &= \frac{1}{2}\sum_{\mathrm{i}} \left(t_i - f(\sum_j T_i f(\sum_j \left(w_{ij} - \theta_j \right) - \theta_j)) \right)^2 \end{aligned} \tag{4}$$

3 Experiment and Result Analysis

3.1Data Preprocessing and Feature Extraction

The data used in this paper is about 370,000. According to the selection method of feature variables, the first three variables are finally used as the main variables to study the data -- “wind speed”, “generator rotation speed” and “active power”. The missing values, singular values and invalid values in the original data are processed, and the missing values are supplemented by the average value of the current dimension. The singular value is deleted by the box plot method, and a total of 1,103 data and invalid values are deleted. The data characteristics of the fan when it is not frozen or frozen are shown in Figures 5 and 6. It can be seen that there are obvious differences between the normal state and the frozen state. The experiment uses a notebook computer equipped with Intel i5-7300HQ CPU 2.5GHz with an 8G memory.

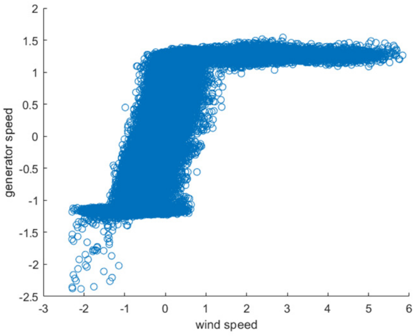


Figure 5 Scatter Diagram of Fan Blade Unfrozen Wind Speed -- Generator Speed

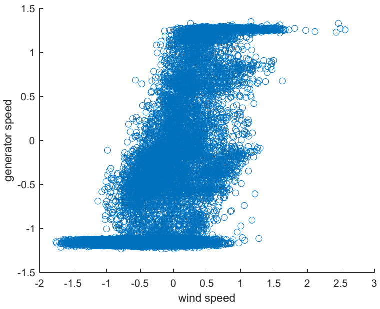


Figure 6 Scatter Diagram of Fan Blade Icing Wind Speed -- Generator Speed

3.2Model Training and Evaluation of BPNN

The BPNN has the abilities of nonlinear mapping, self-learning, self-adaptation, generalization and fault tolerance. Therefore, this algorithm is selected to classify the data here.

3.2.1Model training

a) Data preprocessing: the mapminmax function is used to standardize input data and output data based on Matlab to eliminate dimensional influence.

b) The BPNN neural network model is established. The transfer function is set to ‘tansig’; the training function of BPNN network is set to ‘trainlm’; the BPNN learning algorithm is set to ‘learngdm’; the network performance function is set to ‘mse’; the number of iterations is set to 1,000; the learning rate is set to 0.01; the minimum error of training goal is set to 10-7.

c) The training set is input into the model for training, and the BPNN prediction model of fan blade icing is obtained.

The testing set is input into the model for prediction. After prediction, it is compared with the real value to obtain the confusion matrix.

3.2.2Model Evaluation

After the training of the BPNN model is completed, it is necessary to evaluate the model with some index as the standard. The results can be divided into:

1) TP (correctly identify the blade as normal).

2) FP (misjudge the blade in non-icing state as icing state).

3) TN (misjudge the blade in icing state as non-icing state).

4) FN (correctly identify the blade as icing).

The total number of program recognition errors is recorded as N(formula(5)):

$$N = TN + FP \tag{5}$$

The total number of correct program identification is recorded as P(formula(6)):

$$P = TP + FN \tag{6}$$

The proportion of the quantity correctly identified as “icing” to the actual total number of“icing” is recorded as precision(formula(7)):

$$Precision\ = FN/(TN + FN) \tag{7}$$

The proportion of the quantity correctly identified in the test set to the total quantity is recorded as A (accuracy) (formula(8)):

$$A = \frac{P}{N+P} \tag{8}$$

3.3 Experimental results and analysis

3.3.1MFE-BPNN Experimental Results and Analysis of MFE-BPNN Method

After data preprocessing, the original sample set is obtained. There are 349,064 groups of unfrozen data, accounting for 93.38% of the total data, 24,729 groups of frozen data, accounting for 6.62% of the total data. The unfrozen sample size is 14.1156 times of the frozen sample size. The data are selected randomly to establish the original training set and the testing set. The ratio of the original training set to the testing set is 7:3; the number of hidden layers is set to 3; and the number of nodes in each layer is 8, 5 and 5 respectively. During the experiment, it is found that the proportion difference between the frozen and non-frozen data in the original training set is large, resulting in low prediction precision. Therefore, the proportion of non-icing: icing in the original training set is further adjusted to improve the model accuracy. Taking the icing data as the standard, the non-icing is sampled randomly: a training set is established according to the ratio of 1:1 of the icing data, so as to establish a balanced training set. In order to eliminate the error caused by different data, the training set and the testing set data used by each algorithm in the following paper are the same set of data. Table 2 shows the experimental results of BPNN prediction model.

TABLE 2 EXPERIMENTAL RESULTS OF MFE-BPNN PREDICTION METHOD

Sample set	Original training set		Balanced training set	
Variable category	All feature variables	Main feature variables	All feature variables	Main feature variables
N	472	5461	5520	12497
P	111997	107008	106949	99972
Precision	95.33%	42.20%	57.03%	75.51%
A	99.58%	95.15%	95.09%	88.89%
T(s)	90.4	54.6	12.72	4.7

It can be seen from Table 2 that even if only three main feature variables are extracted and tested after the model is trained by the BPNN network, there is little difference between the accuracy of correctly identifying the blade state and the prediction precision of the model trained by all 26 variables. For example, after the original training set and all feature variables are used to train the model, the prediction precision of the overall test data set is 99.58%, and the prediction precision of extracting three main feature variables is 95.15%. Using the balanced training set, it also reached 95.09% and 88.89% respectively. However, when the prediction precision is calculated only for the icing fault data, the result is low. For example, after the main feature variables are extracted from the original training set to train the model, the prediction precision of the calculated icing fault data is only 42.20% and 75.51% in the balanced training set, which is much lower than the accuracy of the total prediction testing set. In addition, it can be found that the prediction time T after the extraction of main feature variables is much lower than that using all feature vectors, from 90.5 seconds to 54.6 seconds, and from 12.72 seconds to 4.7 seconds.

3.3.2Experimental Results and Analysis of Decision Tree

Decision tree is a common machine learning classification algorithm. The initial decision tree is firstly established in the process, then is pruned, to draw the tree diagram and verify the model by test set. After the main feature variables are extracted from the original training set, an initial decision tree is established. The tree diagram drawn after pruning is shown in Figure 7. No branches are cut off, and all branches pass the test.

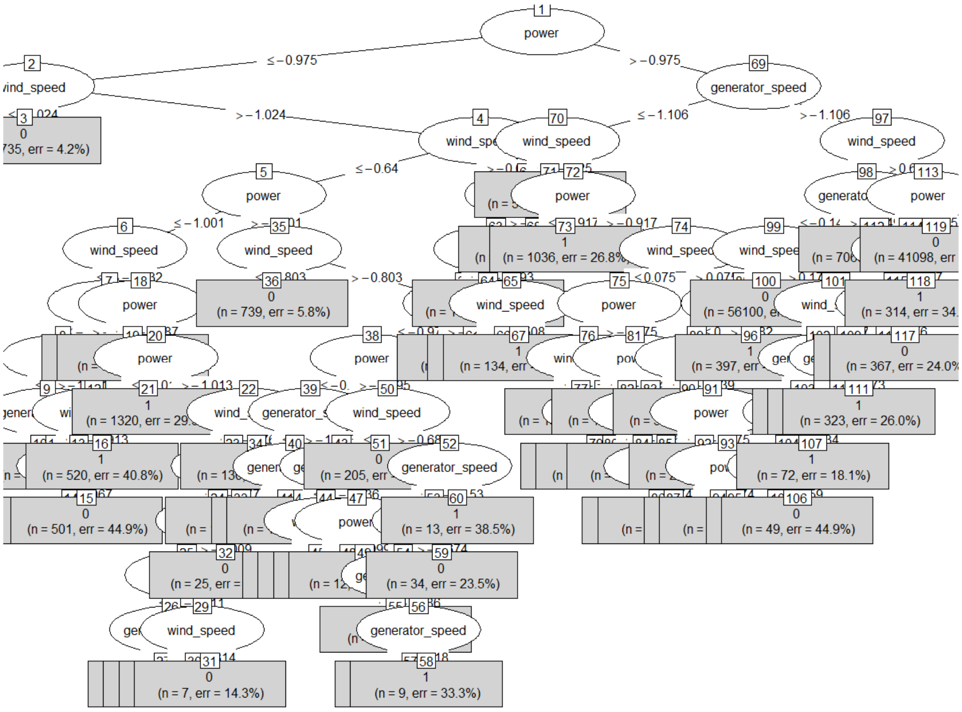


Figure 7 Tree Diagram After Pruning

According to this process, experiments are carried out for different training sets and different feature variables, and the results are shown in Table 3.

TABLE 3 EXPERIMENTAL RESULTS OF DECISION TREE PREDICTION MODEL

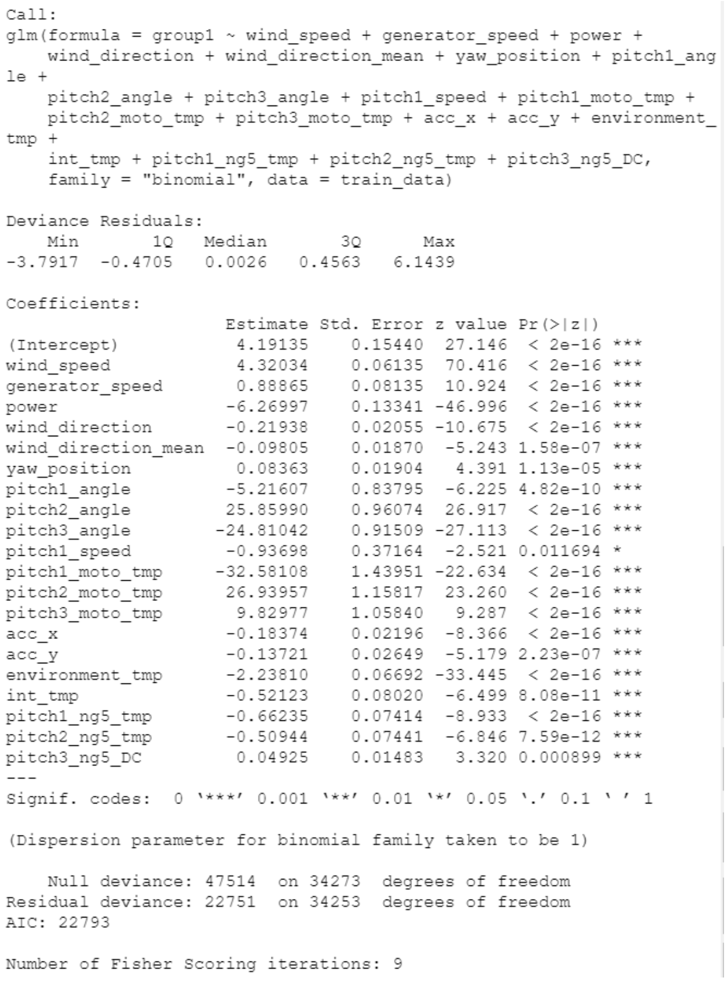
Sample set	Original training set		Balanced training set	
Variable category	All feature variables	Main feature variables	All feature variables	Main feature variables
N	59	5501	841	12124
P	112410	106968	11628	100345
Precision	99.54%	41.73%	99.57%	74.80%
A	99.95%	95.11%	99.25%	89.22%
T(s)	45	5.5	3	0.6

It can be seen that when the original training set of three main feature variables is extracted, the prediction precision of decision tree model for icing fault data is 41.73%, and the prediction precision of overall data is 95.11%. In the case of balanced training set, the prediction precision of decision tree model for the fault data is 74.80%, and the prediction precision of overall data is 89.22%, which is better than the model trained based on the original training set.

When all 26 feature variables are adopted, the prediction precision of the decision tree model for the fault data and overall data under the original training set and the balanced training set is almost the same. For example, the prediction precision of the model trained under the original training is 99.55% for the fault data and 99.95% for overall data, while the prediction precision of the fault data and overall data under the balanced training set is 99.57% and 99.25% respectively.

3.3.3Results and Analysis of Logistic Regression Experiment

Logistic regression is an algorithm that applies the idea of regression to classification problems. Firstly, the logistic regression model is established; then the logistic regression model is modified; and finally, the testing set is used for prediction. In the experiment, the significance test of setting parameters under the extraction of three main feature variables is p = 0.05, and the significance test of setting parameters under the selection of all feature variables is p = 0.001. The logistic regression model is established under the original training set. The output results of all 26 feature variables in R language are shown in Figure 8:



```
Call:
glm(formula = group1 ~ wind_speed + generator_speed + power +
    wind_direction + wind_direction_mean + yaw_position + pitch1_ang
le +
    pitch2_angle + pitch3_angle + pitch1_speed + pitch1_moto_tmp +
    pitch2_moto_tmp + pitch3_moto_tmp + acc_x + acc_y + environment_
tmp +
    int_tmp + pitch1_ng5_tmp + pitch2_ng5_tmp + pitch3_ng5_DC,
    family = "binomial", data = train_data)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-3.7917  -0.4705   0.0026   0.4563   6.1439

Coefficients:
                     Estimate Std. Error z value Pr(>|z|)
(Intercept)           4.19135    0.15440  27.146  < 2e-16 ***
wind_speed            4.32034    0.06135  70.416  < 2e-16 ***
generator_speed       0.88865    0.08135  10.924  < 2e-16 ***
power                -6.26997    0.13341 -46.996  < 2e-16 ***
wind_direction       -0.21938    0.02055 -10.675  < 2e-16 ***
wind_direction_mean  -0.09805    0.01870  -5.243 1.58e-07 ***
yaw_position          0.08363    0.01904   4.391 1.13e-05 ***
pitch1_angle         -5.21607    0.83795  -6.225 4.82e-10 ***
pitch2_angle         25.85990    0.96074  26.917  < 2e-16 ***
pitch3_angle        -24.81042    0.91509 -27.113  < 2e-16 ***
pitch1_speed         -0.93698    0.37164  -2.521 0.011694 *
pitch1_moto_tmp     -32.58108    1.43951 -22.634  < 2e-16 ***
pitch2_moto_tmp      26.93957    1.15817  23.260  < 2e-16 ***
pitch3_moto_tmp       9.82977    1.05840   9.287  < 2e-16 ***
acc_x                -0.18374    0.02196  -8.366  < 2e-16 ***
acc_y                -0.13721    0.02649  -5.179 2.23e-07 ***
environment_tmp      -2.23810    0.06692 -33.445  < 2e-16 ***
int_tmp              -0.52123    0.08020  -6.499 8.08e-11 ***
pitch1_ng5_tmp       -0.66235    0.07414  -8.933  < 2e-16 ***
pitch2_ng5_tmp       -0.50944    0.07441  -6.846 7.59e-12 ***
pitch3_ng5_DC         0.04925    0.01483   3.320 0.000899 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 47514  on 34273  degrees of freedom
Residual deviance: 22751  on 34253  degrees of freedom
AIC: 22793

Number of Fisher Scoring iterations: 9
```

Figure 8 Final Model Summary

The variables are screened according to the summary results, and the variables are eliminated according to the standard of p = 0.001. The modified logistic regression model eliminates "yaw_speed", "pitch1_speed", "pitch2_speed", "pitch3_speed", "pitch3_ng5_tmp", "pitch3_ng5_dc", indicating that the yaw speed, the blade speed, the temperature of ng5 3 and the DC of chargers have little influence on the fault. Therefore, such variables are eliminated.

Similarly, based on this process, experiments are carried out for different training sets and different feature variables, and the results are shown in Table 4.

TABLE 4 EXPERIMENTAL RESULTS OF LOGISTIC REGRESSION PREDICTION MODEL

Sample set	Original training set		Balanced training set	
Variable category	All feature variables	Main feature variables	All feature variables	Main feature variables
N	15607	6237	15625	19452
P	96862	106232	96844	93017
Precision	84.89%	29.27%	84.75%	77.77%
A	86.12%	94.45%	86.12%	82.70%
T(s)	5	2	1	0.3

As can be seen from Table 4, based on the original training set and the logistic regression prediction model of all 26 feature variables, the prediction precision of the fault data is 84.89199%, and the prediction precision of overall data is 86.12%. After selecting three main feature variables, the prediction precision of the trained prediction model for overall data is 94.45%, but the prediction precision of the fault data is only 29.27%. Under the balanced training set, whether based on all feature variables or the prediction model with main feature variables selected, the prediction precision of the fault data and overall data is almost the same; the prediction precision of the fault data is slightly less than that of overall data; the prediction time of extracting main feature variables is slightly less than that of all feature variables.

3.3.4Comparative Analysis of Prediction Results of Three Methods

Comparing the experimental results of the MFE-BPNN prediction model, the decision tree prediction model and the logistic regression in Table 2, 3 and 4, it can be found that the prediction time is greatly shortened after extracting feature variables, and that the prediction effect of the MFE-BPNN prediction method is better. For example, under the original training set, the prediction precisions for the fault data and overall data are 42.20% and 95.15% respectively in the MFE-BPNN prediction method, 41.73% and 95.11% respectively in the decision tree prediction model, 29.27% and 94.45% respectively in the logistic regression prediction model. Similarly, under the balanced training set, the prediction precisions for the fault data and overall data are 75.51% and 88.89% respectively in MFE-BPNN prediction method, 74.80% and 89.22% respectively in the decision tree prediction model, and 77.77% and 82.70% respectively in the logistic regression prediction model.

The fastest training speed is the logistic regression prediction model under the extraction of main feature variables from the balanced training set, which takes only 0.3 seconds. The prediction time after the extraction of main feature variables by the three algorithms is much lower than that of all feature variables. The MFE-BPNN, decision tree and logistic regression all retain the three variables of wind speed, generator power and power. It can be seen that these three variables are closely related to the occurrence of fault.

4 Conclusion

This paper presents a fault prediction and classification method of industrial equipment based on the selection of main feature variables and the back propagation neural network (BPNN). The method proposed in this paper improves the prediction precision and shortens the prediction time through data preprocessing, extraction of main feature variables and prediction model based on the neural network. The method has the characteristics of operability, intelligibility and comparability. The effectiveness of the model is proved by the comparison with the decision tree and the logistic regression classification algorithm. This prediction method, which can simulate multiple variables without inputting complex variables, provides a way to solve the problem of industrial equipment failures and reduce loss.

Acknowledgments. This work was supported by Project "Research on equipment fault prediction method based on industrial big data" supported by National College Student Innovation and Entrepreneurship Training Program (202010431010), and Project "Research on resource allocation method of cloud-edge collaborative computing for artificial intelligence image recognition” supported by Shandong Provincial Natural Science Foundation (ZR2020MF034).

References

[1] Erdem E, Shi J. (2010) ARMA based approaches for forecasting the tuple of wind speed and direction [J]. Applied Energy, 2011, 88: 1405- 1414.

[2] Zhang L Y, Lin J, Qiu R Z, et al. (2018) Trend analysis and forecast of PM2.5 in Fuzhou, China using the ARIMA model [J]. Ecological Indicators, 95: 702-710.

[3] Dazhong,L.Jiarui,L.Huaying,Z.(2019)Prediction method of fan blade icing based on deep fully connected neural network [J]. Electric Power Science and Engineering, 35(04): 39-44.

[4] Richardson R R, Osborne M A, Howey D A. (2017) Gaussian process regression fox forecasting battery state of health [J]. Journal of Power Sources, 357: 209-219.

[5] Wu Yiwen. (2017) Research on Classification Algorithm Based on Decision Tree [J]. http://kns-cnki-net-
s.vpn.qlu.edu.cn:8118/kcms/detail/detail.aspx?FileName=SZTJ201712233&DbName=CJFQ2017.

[6] Dazhong,L.Keyan, Z.Fang,L.(2020)Decision Tree Discrimination Method for Wind Turbine Working State Driven by Big Data[J]. Electric Power Science and Engineering, 36(02): 22-27.

[7] Chunshan,T.Xilin,L.Jia,W.(2016)Geohazard susceptibility assessment based on CF model and Logistic Regression models in Guangdong[J].
Hydrogeology and Engineering Geology, 43(06): 154-161+170.

[8] Hongxia,Y.Xing,L.(2016)Analysis and Forecast Model of Major Loss Flight Accidents Based on MLP Neural Networks Method[J]. Journal of Shanghai University of Electric Power, 32(05): 504-506.

[9] Zong W G, Roper W E. (2009) Energy demand estimation of South Korea using artificial neural network [J]. Energy Policy, 37(10): 4049-4054.

[10] Liu T H, Wei H K, Zhang K J. (2018) Wind power prediction with missing data using Gaussian process regression and multiple imputation [J]. Applied Soft Computing, 71(10): 905-916.

[11] Yu,G.Yu,Y.(2017) Equipment fault prediction based on grey rough set and BP neural network [J]．Application Research of,34(9):2642-2645.

[12] Yi,D.Jian-zhong,Z.Ya-hui,S.Shuai-xuan,L.Yan-he,X.Wei,J. (2018)Diagnosis of Hydropower Units Cavitation Signals Based on EMD-BPNN[J]. Water Resources and Power, 36(03): 157-160.

[13] Harendra Kumar Yadav, Yash Pal, Madan Mohan Tripathi. (2020) Short-term PV power forecasting using empirical mode decomposition in integration with back-propagation neural network [J]. Journal of Information and Optimization Sciences, 41(1): 25-37.

[14] Zhou Guangfei(North China Electric Power University (Beijing)). (2019) Ice Detection for Wind Turbine Blades via Multivariate Statistical Analysis Method[D]. http://223.220.252.171:81/KCMS/detail/detail.aspx?filename=1019237414.nh&dbcode=CMFD&dbname=CMFD2020.