Research on Equipment Fault Diagnosis Classification Model Based on Integrated Incremental Dynamic Weight Combination

Jing Liu, Zhengshan Jiao, Zhijie Wang, Haipeng Ji and Bing Yu

July 30, 2019

Research on Equipment Fault Diagnosis Classification Model Based on Integrated Incremental Dynamic Weight Combination

Jing Liu[1,2]
[1]School of Artificial Intelligence
HeBei University of Technology,
[2]Tianjin JingNuo Data Technology Co., Ltd.
TianJin, China
liujing@scse.hebut.edu.cn

Zhengshan Jiao[1]
[1]School of Artificial Intelligence
HeBei University of Technology
TianJin, China
pisces.zhsh@gmail.com

Zhijie Wang[1]
[1]School of Artificial Intelligence
HeBei University of Technology
TianJin, China
962402242@qq.com

Haipeng Ji[3,2,*]
[3]School of Mechanical Engineering
HeBei University of Technology,
[2]Tianjin JingNuo Data Technology Co., Ltd.
TianJin, China
sioxas@live.cn

Bing Yu[1]
[1]School of Artificial Intelligence
HeBei University of Technology
TianJin, China
1013141237@qq.com

***Abstract*—This study proposes a classification model of equipment fault diagnosis based on integrated incremental learning mechanism based on the characteristics of industrial equipment status data. The model first proposes a Dynamic Weight Combination Classification Model based on Long Short-Term Memory (LSTM) and Support Vector Machine (SVM). Referred to as DWCLS model, it is used to solve the problem of fault feature extraction and classification in high noise equipment state data. Secondly, based on this model, integrated incremental learning mechanism and unbalanced data processing technology are introduced to solve massive unbalanced New data feature extraction and classification and sample category imbalance problem prevalent under equipment status data. Finally, an equipment fault diagnosis classification model based on integrated incremental dynamic weight combination is formed Equipment Fault Diagnosis Classification Model Based on Integrated Incremental Dynamic Weight Combination referred to as DWCMI model, and proved by experiments that the model can effectively overcome the problems of excessive data volume, unbalanced, high noise, and inability to correlate data samples in the process of equipment fault diagnosis.**

Keywords—Neural network, Support vector machine, Integrated increment, Unbalanced data processing, Fault diagnosis

I. INTRODUCTION

In recent years, the rapid development of information technology and industrial Internet of Things has promoted revolutionary innovation and breakthroughs in manufacturing. Represented by Germany's "Industry 4.0" [1], a number of countries have introduced various measures to attract manufacturing returns and improve the level of manufacturing intelligence. Industry 4.0 is a technological transformation of an industry and a change of industry. This paper, based on a large number of domestic and foreign literatures, combined with the characteristics of mechanical equipment fault diagnosis data, by studying the traditional Support Vector Machine (SVM), It is found that for SVM, its generalization ability is strong, only a small batch of sample data is needed for training to ensure good classification effect, but at the same time, support vector machine also has the problem of relying on parameter selection and weak fault tolerance; long-term and short-term memory Although the Long-Short Term Memory (LSTM) does not need to specify the prior probability in advance, it has good fault tolerance and nonlinear processing ability. However, training the neural network requires a large amount of sample data, and it is easy to appear. Based on the above research, an Equipment Fault Diagnosis Classification Model Based on Integrated Incremental Dynamic Weight Combination (DWCMI model) is established, and experiments show that the model can not only effectively overcome The problem of large amount of fault data, unbalanced, high noise, and inability to correlate data samples during equipment fault diagnosis, and effectively saves time cost.

The rest of the paper is structured as follows: Section 2 summarizes the research methods of fault diagnosis methods, unbalanced data processing and incremental learning in the field of equipment fault diagnosis; the third section analyzes the equipment fault diagnosis classification model based on integrated incremental dynamic weight combination .The theoretical basis, method flow, and implementation steps of the DWCMI model are elaborated. In the fourth section, the DWCMI model is applied to the fault diagnosis process of rolling bearing equipment to achieve reliable feature extraction and fault classification. The validity of the DWCMI model is proved by experimental comparison. Finally, in the fifth section, the proposed method is briefly summarized, and the research direction that needs to be improved in the future is proposed.

II. LITERATURE REVIEW

The research status of equipment fault diagnosis, unbalanced data processing and incremental learning are summarized as follows:

A. *Equipment troubleshooting*

In the current development of the industrial field, smart devices play an extremely important role. At the same time, the consequences caused by equipment failures become more and more serious. Therefore, how to carry out effective and accurate equipment fault diagnosis has been widely used by everyone. Pay attention to it. In the development of the fault diagnosis discipline, the most important and most critical issue is the extraction of fault feature information, which is directly

* Haipeng Ji is the corresponding author.

related to the accuracy of fault diagnosis and the reliability of early fault prediction. In order to solve the problem of feature extraction of fault diagnosis, people mainly use signal processing, especially the theory and technical means of modern signal processing, to diagnose equipment by vibration signal, which is one of the most effective and commonly used methods in fault diagnosis [2] . In order to meet the needs of analyzing complex system faults, the use of deep learning to establish complex nonlinear fault diagnosis models based on the unique advantages of feature extraction and pattern recognition has begun to gradually show its characteristics [3].

B. Research Status of Unbalanced Data Processing in Equipment Fault Diagnosis

There are also many algorithms that have been improved for the lack of improved unbalanced data processing methods for classifiers. Qian et al. combine sampling methods with integrated learning theory, in which the sampling used is undersampling and SMOTE algorithm [4]. Han Hui et al. proposed to improve the AdaBoost algorithm and then combine it with the oversampling algorithm [5]. Chawla et al. combine Boosting and oversampling SMOTE algorithms. The basic idea is to incorporate the comprehensive sampling technique into each iteration process, making the following classifiers more concerned with a few classes [6].

C. Research Status of Incremental Learning in the Field of Equipment Fault Diagnosis

Incremental learning technology is a widely used intelligent data mining and knowledge discovery technology. The idea is that the learning accuracy should be improved when the samples are gradually accumulated. The incremental learning algorithm for pattern classification can be applied to applications with large data volumes and data streams. For incremental problems, Gauwenberghs G et al. proposed an exact solution for incremental training[7], adding one training sample or reducing the effect of a training sample on Lagrange coefficients and support vectors. In reference [8], an incremental solution to the exact solution of the global optimization problem is proposed.

III. Equipment Fault Diagnosis Classification Algorithm Based on Integrated Incremental Dynamic Weight Combination

This section introduces the DWCMI method in detail. This method firstly combines the long-short-term memory neural network and the support vector machine dynamically, and uses the support vector machine to dynamically adjust the weights of the respective classifiers in the combined model to improve the classification of rolling bearing fault diagnosis. Accuracy; secondly, the non-equilibrium processing model is used to solve the problem of fault category imbalance in massive equipment state data; finally, the fault learning feature extraction and classification problem of newly added equipment state data is processed in real time through the incremental learning mechanism.

A. Unbalanced Data Processing Model

Based on oversampling and undersampling, this paper proposes a new data sampling model, which is based on the oversampling and undersampling fusion data sampling model (New Kernel Synthetic Minority Over-Sampling Technique And Tomek links Based on K-Nearest Neighbors, Referred to as NKSMOTE-NKTomek model). Oversampling refers to the improved SMOTE algorithm, namely the NKSMOTE algorithm[7]. The algorithm mainly solves the effective. Reference [9] introduced the PSVM-Proximal Support Vector Machine, which is extremely fast.

In addition to the above methods for implementing incremental learning by improving the SVM algorithm, we can also implement incremental classification learning through multi-classifier integration. The multi-class integration method is divided into three categories [11].

1) When a new set of data is encountered, a set of fixed classifiers updates the joint rule or voting weight.

2) The existing classifier updates the parameters each time a new data set is encountered.

3) Each time a new data set is encountered, a new classifier is trained with the new data set and added to the current integration.

For the data of types 1) and 2), 2 samples are randomly selected among the k samples, and n new samples are synthesized according to a certain rule among the 3 samples, wherein the n value is the upsampling magnification.

If the selected two samples y_1 and y_2 are majority samples, use the following formula to generate n samples:

1) Generate n temporary samples $t_j(j=1,2,...,n)$ according to y_1 and y_2:

$$t_j = y_1 + rand(0,0.5) \times (y_2 - y_1) \qquad (3\text{–}1)$$

2) Generate a new minority class sample $X_j(j=1,2,...,n)$from t_j and x:

$$X_j = x + rand\left(0,1\right) \times \left(t_j - x\right) \qquad (3\text{–}2)$$

If there are a few samples in the selected two samples y_1 and y_2, use the following formula to generate n samples:

1) Generate N temporary samples $t_j(j=1,2,...,n)$ according to y_1 and y_2:

$$t_j = y_1 + rand\left(0,1\right) \times \left(y_2 - y_1\right) \qquad (3\text{–}3)$$

2) Generate a new minority class sample $X_j(j=1,2,...,n)$ from t_j and x:

$$X_j = x + rand\left(0,1\right) \times \left(t_j - x\right) \qquad (3\text{–}4)$$

For the 3) type of data, n is set to 1 in order to reduce the risk caused by the noise data. At the same time, a few classes y are randomly selected from a small number of samples, and a new sample is randomly generated using the following formula.

$$X = x + rand(0.5,1) \times (y - x) \qquad (3\text{–}5)$$

When obtaining the K-nearest neighbor, the nonlinear mapping function is first used to map the sample to the kernel space. The distance between the samples in the kernel space is called the kernel distance. The calculation formula is:

$$\begin{aligned} d(\varphi(x),\varphi(y)) &= \sqrt{\|\varphi(x)-\varphi(y)\|^2} \\ &= \sqrt{K(x,x)+K(y,y)-2K(x,y)} \end{aligned} \qquad (3\text{–}6)$$

Where $\varphi(x)$ is a nonlinear mapping function, $K(x,y)$ is a kernel function, and the kernel function used here is a Gaussian kernel function, and its calculation formula is:

$$K(x,z) = e^{-\frac{\|x-z\|^2}{2\sigma^2}} \qquad (3\text{–}7)$$

(2) Tomek links Based on K-Nearest Neighbors (NKTomeK model): For minority and majority samples, a small number of samples are divided into different categories of data based on k-nearest neighbors. Reduce the number of samples calculated by the Tomek links algorithm and increase the efficiency of undersampled samples.

1)According to the minority class sample class in step (1), the sample set synthesized in step (1) is re-divided into a minority class sample and a majority class sample, and k neighbors are obtained for each sample in the minority class sample, according to k The sample categories of the majority and the minority samples in the nearest neighbor are the safety samples, the boundary samples and the noise samples. The category judgment criteria are as shown in step (1).

2)Remove noise samples.

3)Assume that the boundary sample set in the minority sample is D, and the majority sample set is U, where the number of U, is N.

$$for\ i = 1,2,...,N$$

4)The distance between x_1 in the majority sample set U and x_2 in the boundary sample set D, is $d = d(x_1, x_2)$.

5)The distance between x_1 in the majority sample set U and each sample in the synthesized sample set is the distance data set F. If $d < F$, the line where x_1 is located is returned.

6)At the end of the loop, delete the sample set of the returned rows. Combine the majority class set U with a few class sets.

B. Integrated incremental learning mechanism

The algorithm steps of the integrated incremental model based on learn++ are as follows:

(1) Assume that the input is the training data set d^t to be processed at the tth time, d^t is composed of the instance $x^t(i)$, and the examples $i = 1,2,...,m^t$, a total of m^t.

(2) When $t = 1$, set instance weight w_i^1, penalty weight D_i^1 is equal weight:

$$D_i^1 = w_i^1 = \frac{1}{m^1}, \forall i \qquad (3\text{–}8)$$

(3) At the subsequent t time, w_i^1, and D_i^1 are determined according to the classification accuracy of the integrated classifier on the current data set. t time training on the new data set to generate a new classifier h_{i^t}. Then, all currently generated classifiers need to calculate the classification error rate on the new data set:

$$\varepsilon_k^t = \sum_{i=1}^{m^t} D_i^t [h_k^t(x^t(i) \neq y^t(i))], k = 1,2,...,t \qquad (3\text{–}9)$$

Penalty weight is used to calculate the error rate. D_{i^t} weighting:

$$D_{i^t} = \frac{w_i^t}{\sum_{i=1}^{m^t} w_i^t} \qquad (3\text{–}10)$$

(4) The base classifiers that produce different error rates in different periods are handled differently. For the currently generated base classifier h_k^t,if the error rate $\varepsilon_{k=t}^t > 0.5$, the classifier is invalid and needs to be re-learned to generate a new classifier. For the base classifier h_k^t generated at the previous k time, if the error rate $\varepsilon_{k=t}^t > 0.5$, the error rate $\varepsilon_{k<t}^t$ of the classifier is set to 0.5. h_k^t, indicates the use of the classifier generated at time k at the current t time. Finally, all base classifiers are weighted and integrated to form an integrated classifier:

$$H^t(x^t(i)) = \arg\max \sum\nolimits_k W_k^t [h_k^t(x^t(i) = c)] \qquad (3\text{–}11)$$

The voting weight of each base classifier is determined by the weighted average error rate of the classifier:

$$W_k^t = -\lg \beta_k^t \qquad (3\text{–}12)$$

C. LSTM-SVM based dynamic weight combination classification model

The LSTM-SVM dynamic weight model relies on SVM to dynamically adjust the weight between the LSTM and SVM classifiers. This section focuses on the construction process of the combined model.

1) The SVM model construction process is as follows

(1) Extract feature vectors from known training samples to build training sample sets $\{(x_i, y_i) \mid i = 1,2,...,n\}$;

(2)Select the corresponding kernel function and corresponding parameters;

(3) Under the condition that the condition $\sum_{i=1}^{n} y_i \alpha_{i=0}$ is satisfied, the optimal Lagrangian parameter α^* is found;

(4) The support vector is searched from the training sample set, and the weight coefficient w^* and the classification threshold b^* of the optimal classification hyperplane are solved to obtain the optimal classification hyperplane.

(5) End the training process and get the SVM classification model.

2) The LSTM model construction process is as follows

a) Input layer

Let the training set $x \in R^{m \times n}$, m represents the number of samples, and n represents the data dimension. At the same time, the time dimension is added to the training set, and the training data is transformed into a three-dimensional matrix, that is $R^{m \times t \times i}$, where t represents the time dimension on the sample, that is, the sequence length, and i represents the input neuron at each moment. Dimensions. Map $x \in R^{m \times t \times i}$ to a linear input layer with weight $W^{(i)}$ and offset $b^{(i)}$, and change the data dimension i at each moment of the sample. Its formula is:

$$y^{(i)} = x \times W^{(i)} + b^{(i)} \tag{3–13}$$

among them: $W^{(i)} \in R^{i \times i_1}$; $b^{(i)} \in R^{i_1}$; $y^{(i)} \in R^{m \times t \times i_1}$.

b) LSTM network layer

It can be seen from (1) that the input of the LSTM network layer is $y^{(i)} \in R^{m \times t \times i_1}$, and the number of neurons of the LSTM is d, assuming the hidden layer output of the last time of each sample as the LSTM. The output of the network $y^{(h)}$, then $y^{(h)} \in R^{m \times d}$.

c) Output layer

Use the output of (2) as the input of the output layer, and use the softmax output layer to match the output dimension of the LSTM network layer with the last number of classifications.

$$y' = soft\max(y^{(h)} \cdot W^{(o)}) \tag{3–14}$$

among them: $W^{(o)} \in R^{d \times q}$,q is the number of categories. y' is the output of the network architecture, $y' \in R^{m \times q}$.

d) Cost function

The output probability distribution of the training is compared with the real data distribution, and the cross-entropy cost function of the predicted output and the actual output is calculated.

$$H(y) = \sum_m y' \log_2 y \tag{3–15}$$

Establish the basic structure of the LSTM as above, initialize the network parameters, and set the training times T. In each iteration, the cross-entropy cost function of the current iteration is obtained by forward propagation, and then the network parameters are updated by error backpropagation. Finally, iteratively T times, the cost function tends to converge gradually.

IV. EXPERIMENT

This section mainly analyzes the application performance of the proposed DWCMI model in the fault diagnosis of rolling bearings, and realizes the feature extraction and fault mode classification of bearing equipment. Firstly, the experiment selects the appropriate model parameters and model structure through experiments, and then demonstrates the effectiveness of the DWCMI model by comparing the experiments with BP, LSTM, SVM and other algorithms without adding unbalanced data processing and without adding incremental learning.

A. Data Description

The experimental data comes from the bearing status data of the Electrical Engineering Laboratory of Case Western Reserve University (CWRU). In order to test the effect of the NKSMOTE-NKTomek model on the classification of rolling bearing fault diagnosis in the case of unbalanced data sets, 80, 40, 20, 10 fault samples were randomly selected from the training samples and combined with 100 normal samples as 4 different. Training samples, 20 fault samples and 50 normal samples as test samples, each sample contains 1024 sample points, and four training samples are placed in the DWCLS model added to the NKSMOTE-NKTomek algorithm for training, get training results and test results. The specific bearing status data sample description is shown in Table I.

TABLE I. DESCRIPTION OF BEARING STATUS DATA

Group	Tilt Rate	Status Type	NumberOf Samples	Sample Points
DataSet1	1.25	Normal	100	1024
		Malfunction	80	1024
DataSet2	2.5	Normal	100	1024
		Malfunction	40	1024
DataSet3	5	Normal	100	1024
		Malfunction	20	1024
DataSet4	10	Normal	100	1024
		Malfunction	10	1024
TestSet	2.5	Normal	50	1024
		Malfunction	20	1024

B. Model structure

1) The proportion of noise in the data set

The noise specific gravity added to the model is one of the important factors affecting the model's effect. Adding too much noise may cause the model to lose some useful information and increase the training time of the model. Therefore, we determine the non-equilibrium and wavelet transform of the model. After the structure, the LSTM model structure, and the SVM model structure, it is necessary to determine the magnitude of the noise specific gravity through experiments. Figure 1 shows the variation of the correctness rate with the number of training times when the noise specific gravity is 10%, 20%, 30%, 40%. It can be seen from the figure that when the noise specific gravity is 20%, the correct rate of the model varies with the number of trainings. The increase, the convergence speed is faster, and the learning efficiency is higher. Therefore, 20% noise specific gravity is added to the original sample to verify the effectiveness of wavelet denoising and reconstruction in the data set processing.

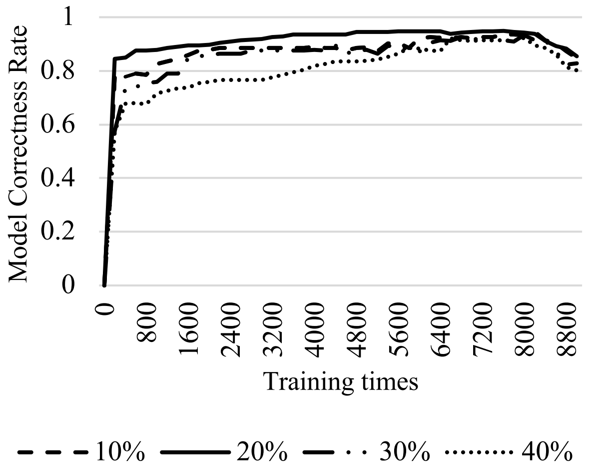

Figure 1. Noise specific gravity line chart

2) LSTM model

When establishing a bearing fault diagnosis and recognition model based on long-term and short-term memory neural network, some parameters need to be set by themselves. After continuous iteration, the optimal parameters are selected to obtain appropriate data values. The parameters that have a large influence on the LSTM model include the training times T, the learning rate η, the sequence length step, and the number of neurons d in the hidden layer unit. Each parameter has a great influence on the training effect, training time, and computational complexity of the LSTM model.

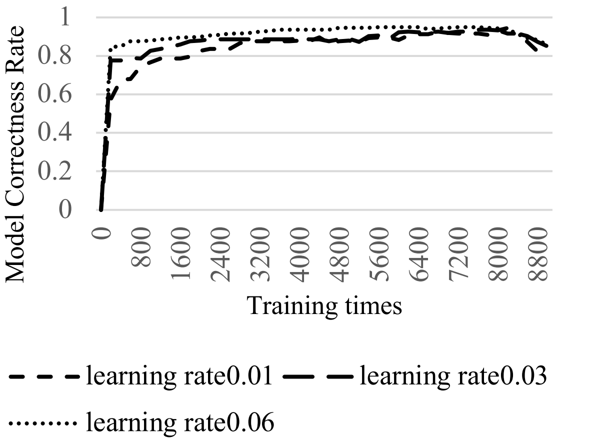

Figure 2. Learning rate line chart

Figure 2 shows the relationship between the correct rate and the number of training times under different learning rates. It can be seen from the figure that when the learning rate is 0.001, 0.003, 0.006, the correct rate is roughly the same when the number of iterations is less than 5000, and the number of iterations is 7000. And when the learning rate is 0.006, the accuracy rate curve shows a significant drop. The learning rate generally determines the speed at which the parameter is iterated to the optimal value. In general, the larger the learning rate, the larger the step size of the gradient in training, and the easier it is to skip the optimal solution.

Figure 3 shows the correct rate curve for the sequence length s of 256 and 1024 (learning rate 0.006, training times 8000). In the LSTM network, the longer the sequence of the sample, the more the number of gradient iterations of the backward propagation of the error, and the larger the calculation amount, which affects the convergence speed and learning efficiency. As shown in Figure 7, when the sequence length is 1024, the convergence speed is slow and the learning efficiency is low. When the sequence length is 256, the correct rate is significantly improved, and the convergence speed and learning efficiency can also be improved.

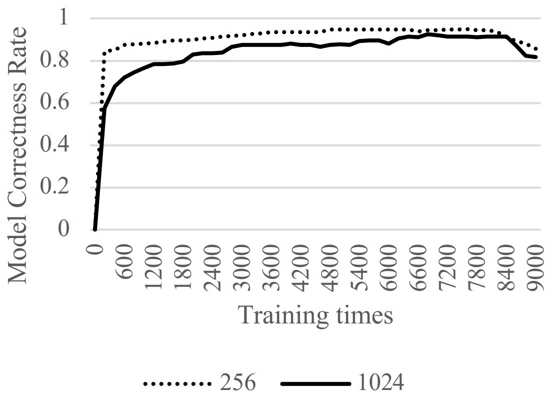

Figure 3. Sequence length line graph

From the above experimental results, it can be concluded that the training frequency T of the LSTM model in the rolling bearing fault diagnosis experiment should be set to 8000, the learning rate η is set to 0.006, the sequence length s is set to 256, and the number of neurons is set to 200.

C. Experimental results

After verifying the validity of the unbalanced data processing in the DWCMI model, a comparison experiment was performed on the performance of the integrated incremental learning algorithm in the DWCMI model, and four different training sets were divided into four groups, one of which was used as a training DWCMI model. The remaining three groups are used to add to the existing model for integrated incremental learning in three times. Use the DWCMI model proposed in this paper to compare the incremental learning with the DWCLS model, LSTM model, SVM model, and BP model, and use the test data set to test the model classification diagnosis effect, and record the accuracy of 20 experimental results for each set of incremental data. The G-mean value and the running time are averaged. The training average and test value comparison results of the four sets of incremental data are shown in Table II.

It can be seen from Table IV that the DWCMI model proposed in this paper is superior to the other four models in the training set of four different unbalanced tilt rates in terms of accuracy, G-mean and running time. From the final diagnosis of the model, the DWCMI model proposed in this paper maintains a high level of accuracy and G-mean values in different data sets of different tilt rates, which is significantly better than other BP models without incremental learning. Compared with the LSTM model, the DWCLS model and the SVM model have better training results, but the training time of the DWCLS model is significantly higher than that of other models. The SVM model has a large fluctuation in the test results of different tilt rate test sets. Therefore, the integrated incremental learning model in the DWCMI model can effectively incrementally train new data to obtain new models, and the DWCMI model can better adapt to changes in the environment through dynamic weights and penalty coefficients.

Table II. Comparison of Troubleshooting Results

Train Set	Method	Train Precision (%)	Train G-mean	Test Precision (%)	Test G-mean
1	DWCMI	91.38	0.9133	89.05	0.8774
	DWCLS	91.26	0.8991	88.23	0.8671
	LSTM	90.64	0.8923	87.05	0.8529
	SVM	89.83	0.8851	81.34	0.7745
	BP	90.25	0.8815	84.24	0.8306
2	DWCMI	91.03	0.8529	88.93	0.8441
	DWCLS	90.33	0.8485	87.36	0.7864
	LSTM	90.74	0.8513	86.85	0.7732
	SVM	87.24	0.7899	82.15	0.6892
	BP	89.71	0.8476	83.54	0.7014
3	DWCMI	91.56	0.8421	89.25	0.7436
	DWCLS	90.65	0.7585	87.93	0.6928
	LSTM	91.44	0.8402	87.87	0.6913
	SVM	89.19	0.7523	82.95	0.6708
	BP	90.63	0.7563	84.34	0.6864
4	DWCMI	91.91	0.7549	90.60	0.6892
	DWCLS	89.14	0.6819	88.24	0.6702
	LSTM	91.32	0.6981	88.05	0.6742
	SVM	92.24	0.7657	83.05	0.6257
	BP	91.34	0.6858	85.24	0.6454

V. CONCLUSION

In order to solve the problem of equipment fault diagnosis under the condition of massive unbalanced high-noise equipment status data, this paper proposes a classification model of equipment fault diagnosis based on integrated incremental dynamic weight combination. The model first uses the New Kernel Synthetic Minority Over-Sampling Technique And Tomek links Based on K-Nearest Neighbors (NKSMOTE-NKTomek model) to solve the category imbalance problem in the equipment operating state data. Secondly, the wavelet transform is used to remove the noise points in the vibration signal, and then the wavelet packet reconstruction is used to reconstruct the vibration signal after denoising into the original vibration signal; and the characteristic parameters of the vibration signal are realized by using the pole symmetric mode decomposition (ESMD). Extracting, normalizing feature parameters into feature vectors, using feature vectors to train long- and short-term memory neural networks and support vector machine models, and using support vector machines to dynamically adjust the weights of the respective classifiers in the combined model to achieve long-term and short-term The dynamic weight adjustment of the memory neural network and the support vector machine in the combined model. Finally, when there is new sample data, the feature vector of the newly added sample is put into the previously trained combination model as the test set. If the output result of the combined model meets the expected result, no processing is performed, if the output result does not meet the expected requirement. Integrate incremental learning for new sample feature vectors. After experimental verification, the DWCMI model makes the bearing fault diagnosis efficiency reach 89.45% on average, which is 3.93% higher than that of the non-incremental learning and unbalanced data processing process. Reliable classification of fault categories under massive unbalanced high-noise rolling bearing equipment status data.

Although the DWCMI model proposed in this paper has high classification performance for fault diagnosis of massive non-equilibrium and high-noise equipment operating state data, there are still some places where the model can be further improved. First, the model training time needs to be further shortened. In particular, the integrated incremental learning method can be further improved to obtain a faster integrated incremental learning model. At the same time, the unbalanced data processing in the DWCMI model is only for normal. Classes and fault classes are processed without further division of the different categories in the fault class. The next step is to refine the categories to obtain a more balanced data set and improve the classification of fault diagnosis.

ACKNOWLEDGMENT

This research is supported by Tianjin Science and Technology Project No.18YFCZZC00060 and No.18ZXZNGX00100.

REFERENCES

[1] Chang Shan. Industry 4.0: Intelligent Factory and Production [J]. Chemical Management. 2013(21)

[2] Yao Min. Digital Image Processing [M]. Beijing: Mechanical Industry Press, 2006: 140-142.

[3] He Xuewen. Research on Theory and Method of Fault Intelligent Diagnosis Based on Support Vector Machine, Central South University. 2004.

[4] BATISTAG E A P, PRATI R C, MONARD M C. A study of the behavior of several methods for balancing machine learning training data[J]. Association for Computing Machinery Special Interest Group on Knowledge Discovery and Data Mining Explorations Newsletter, 2004, 6(1):20-29.

[5] QIAN Yun, LIANG Yan-chun, LI Mu. A resampling ensemble algorithm for classification of imbalance problems[J]. Neurocomputing, 2014, 143(143): 57-67.

[6] Wang Li, Chen Hongmei. Non-equilibrium dataset classification method based on NKSMOTE algorithm. Computer Science, 2018, 45(9): 2-4.

[7] CHAWLA N V, LAZAREVIC A, HALL L O. SMOTEBoost: improving prediction of the minority class in boosting[J]. Lecture Notes in Computer Science, 2003, 2838: 107-119.

[8] Gauwenberghs G,Poggio T. Incremental and Decremental Support Vector Machines[J]. Machine Learning,2001,44(13):409-415.

[9] SOJODISHIJANI O,RAMLIAR.Just-in-Time Adaptive Similarity Component Analysis in Nonstationary Environments． Journal of Intel-ligent & Fuzzy Systems，2014，26(4) : 1745－1758.

[10] GOMES H M，BA R DDAL J P，ENEMB R ECK F． Pairwise Combi-nation of Classifiers for Ensemble Learning on Data Streams // Procof the ACM Symposium on Applied Computing． New York，USA:ACM，2015: 941－946.

[11] Haidong Shao,Hongkai Jiang,Xun Zhang,et al.Rolling bearing fault diagnosis using an optimization deep beliefnetwork[J].Measurement Science and Technology, 2015,26(11):115002.