Transformers Fault Prediction: An Improved Ensembled Method

Lu Peiwen[1,2], Huang Yongjing[1], and Khushnood Abbas[3]
[1] Department of Electrical and Information Engineering, Chengdu Textile College, Chengdu, China
[2] School of Electrical and Electronic Information, Xihua University, Chengdu, China
[3] School of Computer Science and Technology, Zhoukou Normal University, Zhoukou, China
Email: {1094922147, 15935243}@qq.com, khushnood.abbas@gmail.com

***Abstract*—In this study we present a data driven prediction approach to early prediction of transformer's fault. To make such prediction we have collected dissolve gas data of transformer. We have solved this problem bagging based ensembled algorithm. Further we have found that our data has imbalanced class examples. To overcome this, we have removed class bias by using Synthetic Minority Over Sampling Technology (SMOTE). SMOTE is best known for generating synthetic data for minority classes. It is also proven to be better than random sampling. SMOTE oversamples the minority classes data by fitting the linear lines among them. In that way we can generate as many data as we want. Thus, it helped us in avoiding overfitting problem. Our empirical results show that proposed framework outperforms the state-of-the-art methods such as BP neural network, and support vector machine. Our method achieves 90.67 % precision accuracy which is better than the base lines.**

***Index Terms*—transformer fault prediction, Dissolved Gas Analysis (DGA), bagging, support vector machine, ensembled learning**

I. Introduction

Your goal is to simulate the usual appearance of papers in the. We are requesting that you follow these guidelines as closely as possible. Transformers are one of the core equipments of power plants and substations.s. Transformers have a crucial role in reducing the incidence of faults and improving the safe operation of electrical power systems, which is why transformers play an important role in the whole power system supply. Most of the research on transformer fault diagnosis is based on the study of dissolved gas contents in the transformer's oil [1]. Dissolve Gas Analysis (DGA) is being done for a long time [2], [3] for the health diagnosis of transformers. DGA is based on concentration of the gases discharged by the transformers such as: hydrogen (H_2), methane (CH_4), ethane (C_2H_2), carbon monoxide (CO), carbon dioxide (CO_2), ethylene (C_2H_4) and acetylene (C_2H_2) [1]. CO and CO_2 are considered as responsible for cellulosic insulation. Along with CO and CO_2, generally small amount of H_2 and CH_4 is also associated with cellulosic decomposition. In normal functioning, transformers discharge these gases. It is found that when transformers discharge these gases abnormally is because of local overheating, electrical overload, and so on. Many methods are used to measure the anomalous behavior of the gases, such as threshold-based IEEE C577104 [4]; and ratio based [5], [6]. Gas ratio based measured are used for quantitative measurement while threshold-based used for classifying normal vs. abnormal. Although the amount of gas discharge depends on the transformer's age. Therefore age is also an important factor. The other factor may be loading, thermal factor, presence of one or more faults and voltage surge, etc.

When the transformer fault occurs, the dissolved gas characteristics in the oil tank of the transformer is different from that in the normal working state. Therefore, the transformer fault prediction and degree of damage to the transformer are judged by exploiting Dissolved Gas Analysis (DGA) [7]. In DGA, the transformer's oil concentrations and other properties are used to make such predictions. We extract some quantitative measures so that numerical predictions can be made. However, the judgment of fault types depends too much on expert experience. The transformers go into different fault types. There is no extensive study found for making such predictions. Previous method cannot identify the defects of multiple fault categories.

Based on DGA, Yin Haojie organically combines four common single prediction methods artificial neural network, time series, grey prediction, and regression prediction introduces induced ordered weighted average operator and Markov theory and establishes a new predictive model for fault predictions in transformer [8]. Liu Jiajia proposes the index weighted method, as the method of on-line monitoring transformers faults based on DGA technology is established [9]. Ge Xuliang uses the method of Deep Belief Network (DBN) in a deep neural network to construct a diagnostic model of Classified Deep Belief Network (CDBN), transformer on-line monitoring, and fault diagnosis method based on DGA [10]. Zhu Yuanye used BP neural network and grey theory to build a combined model to predict the changing trend of five characteristic gases in transformer's oil and then used the multi-level fault diagnosis with the help of a probabilistic neural network. The gas prediction value obtained by the above predictive model is used to

Manuscript received September 2, 2020; revised November 23, 2020.

doi: 10.18178/ijeee.8.4.82-87

diagnose the nature of the transformer fault [11]. Jia Jinglong selects BP neural network, extreme learning, and deep learning to make such predictions. In addition, Li Chunmao, Zhang Dongbo and Xu Mu used ,other predictive methods such as rough set [12], neural network [13], [14], support vector machine [15], [16], expert system [17], fuzzy theory [18] to diagnose transformer faults and achieved good diagnostic results. However, in the process of using these methods, they are affected by random factors and have poor stability. Moreover, the random initialization weights are defective and easily fall into the limit value. Their reliability and adaptability need to be improved.

A. Bagging Algorithm

1) Introduction of bagging algorithm

The Bagging algorithm (bootstrap aggregating) is a method to improve the accuracy of a learning algorithm. First, a series of predictive functions are constructed and then they are combined into a predictive function in a certain way. The Bagging algorithm uses an unstable weak classification method. Instability lies in that small changes in the data set can cause significant changes in the classification results. The accuracy of a single weak learning algorithm is not high. However, the learning algorithm is used many times to obtain a set of prediction function sequences. Finally, prediction is made according to the characteristics of the prediction samples. Voting choice can improve the accuracy of prediction. By introducing voting choice can improve the precision of expectation. Sensibly, the Bagging model trains a few weak classifiers in parallel to frame a solid classifier.

2) Bagging algorithm process

The Bagging algorithm completes the training of a sample set by calling a weak learning algorithm repeatedly, and then uses a strong learning machine as a model to integrate multiple weak classifiers into strong classifiers. Through repeated learning of the model, the prediction accuracy of the ensembled model is improved, and the deep regularity characteristics of the sequence to extracting fault identification for the record is diagnosed [16]. The algorithm consists of two steps: 1) training of base classifier: each base classifier is sampled and trained by bootstrap sampling method; 2) result integration: the results of base classifier are combined by vote or weighting.

The specific algorithm is as follows:

Input: Training set S, test data x, classification method C;

Output: Class discrimination R for x;

```
Begin
(1)     for i=1, 2,…, T do
(2)     S=bootstrap (x);
(3)     C=C(S);
(4)     R=C(x);
(5)     endfor;
(6)     R= ( )
```

Among them, represents the support category that seeks the most, which is obtained from the decision results of T classifiers.

B. Bagging Improved Algorithm Based on SMOTE

The traditional Bagging algorithm obtains the difference between member networks by the randomness of sample extraction. It has high classification accuracy on unbalanced data sets. However, a large number of candidate member networks need to be generated and trained beforehand, which can easily lead to network redundancy and a huge integration scale. It is not an ideal result. A Synthetic Minority Over Sampling Technology (SMOTE) is a Virtual Minority Classes Sampling method proposed by Chawla *et al.* [19]. It is a heuristics-based method that helps in avoiding the situations of over fitting, which is generally caused by non-heuristics based methods. The center thought of SMOTE is to embed new examples that produced randomly between minority class samples and their neighbors, which can expand the quantity of minority class tests and improve the circumstance of class imbalance [20], [21].

Assuming that X is the input of a few samples, it has a similar nearest neighbor, $a_1 \sim a_n$, randomly selected, and interpolated randomly between X and, a new minority sample can be constructed. The new sample is $a_1 \sim a_2$.

$$X_{new} = X + u(0,1) * (X - a_i) \tag{1}$$

where $u \in (0,1)$ is a random number sampled from uniform distribution $(0,1)$. $a_i \in a_1 \sim a_2$. The SMOTE algorithm does not merely copy samples according to a random oversampling method but adds new and non-existent samples so that the classification plane of the classifier extends to most classes space, which can avoid over-fitting of the classifier to a certain extent. In order to improve the accuracy of the whole classification, an improved Bagging algorithm based on SMOTE is proposed. Firstly, the SMOTE algorithm is used to increase the number of minority samples and weaken the imbalance between minority and majority classes. Secondly, the minority samples are weighted to improve their contribution to the Bagging-based classifier. Finally, the Bagging algorithm is used to integrate the improved algorithm with the classification accuracy of the base classifier as the weight. Definition X : sample set; n: sample array; X_{min}: minority class samples in X; X_{max}: majority class array samples in X; k: count array; ω: sample weight; R: arbitrary sample; IB, OB: sample set; $Mean$: Finding mean value; σ^2 variance; $abs()$: absolute value. The algorithm steps are as follows:

Identification of minority samples in SMOTE:

Step 1:

1) Record the number of samples k, and sample array n in each category of sample set X.

2) Calculate the mean value of k as $mean(k)$, variance as σ^2.

3) If

$k[i] < Mean(k)$ && $abs(k[i] - Mean(k)) > Square(k)$

$Xmin[j] = n[i]$ $Xmax[j] = n[i]$

4) EndIf

5) Return X_{min}, X_{max}

Step 2: Increase the number of minority samples in SMOTE:

1) five similar nearest neighbors $a_1 \sim a_5$ are calculated for any sample R in Xmin

2) one of them is selected to calculate the attribute difference vector $R - a_i$

3) a number u is sampled form $\sim$ uniform(0, 1)

4) $R_{new} = R + u(0,1) \bullet (R - a_i)$

5) Return R_{new}

6) Determining the weight of training samples For any sample R in the new sample set X_{new}, if so, the weight of the sample is

$$\omega = \min\left(\frac{k[j]}{\sum k[i]}, 1 - \frac{k[j]}{\sum k[i]}\right) \tag{2}$$

$k[j] \in X_{max}$ and $\sum k[i]$ is number of all samples. The weight of sample ω is

$$\omega = \frac{\sum k[i]}{k[i]} \tag{3}$$

where, $k[j] \in X_{min}$ and $\sum k[j]$ is number of all samples.

7) Training the base classifier: To use the Bagging algorithm to extract samples from the new sample set X_{new}, OB is an unselected sample, IB is used to train the base classifier. And OB is used to set the base classifier. The test accuracy of the base classifier is taken as the weight w. After training, the final classification result is determined by voting on the test set according to the weight w of each base classifier.

In order to verify the performance of Bagging's improved algorithm, the area under the receiver operating characteristic curve (ROC curve) is chosen as the evaluation index [22]. The AUC of the Bagging algorithm is 0.798, while that of the Improved Bagging algorithm is 0.948. The test results are shown in Fig. 1.

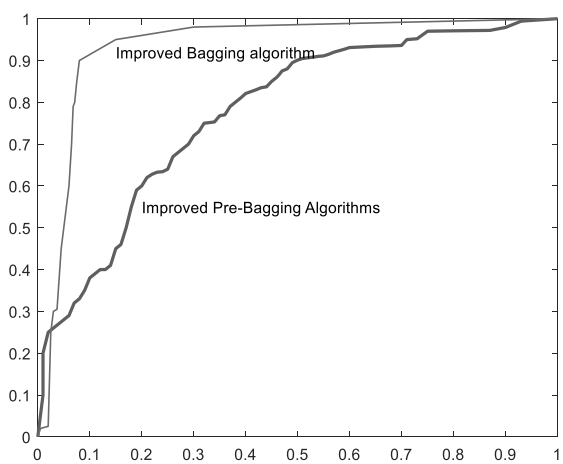

Figure 1. ROC of data set and curve comparison before and after improvement.

Fig. 1 shows that the improved Bagging algorithm shows good performance on the whole, and the corresponding Y value near 0.3 is the best. It can be seen that the improved Bagging algorithm based on SMOTE can balance the gap between the minority samples and the majority samples, increase the contribution of the minority samples, change the weight of the samples by using the Bagging algorithm, and further improve the impact of the minority classes. Finally, the weighted form is used to integrate, which can ensure the accuracy of the overall classification. The accuracy of classification of a few samples is improved.

C. *Transformer Fault Diagnosis Based on Improved Bagging Algorithm*

1) *The judgment basis of transformer fault category*

Transformer failure will be accompanied by a discharge or exothermic process. Transformer oil will dissolve and release five characteristic gases: hydrogen, methane, ethane, ethylene and acetylene. According to the characteristics of the normal state of a transformer and the content of five characteristic gases when the fault occurs, the content of each gas component will be changed. Transformer fault can be judged by analysis. The types of transformer faults can be divided into: 1) partial discharge; 2) low-energy discharge; 3) low-energy discharge and overheating; 4) arc discharge; 5) arc discharge and overheating; 6) fault-free; 7) low-temperature overheating; 8) medium-temperature overheating; 9) high-temperature overheating.

2) *The flow chart of bagging improved algorithm*

For transformer fault diagnosis, the rationality of data sample selection is closely related to the accuracy of fault prediction. After data sample selection, data should be processed first in order to establish a reliable training model. Then the data samples are trained, classified and finally determined according to the steps of Bagging improved algorithm based on SMOTE. The flow diagram of the improved Bagging algorithm is shown in Fig. 2.

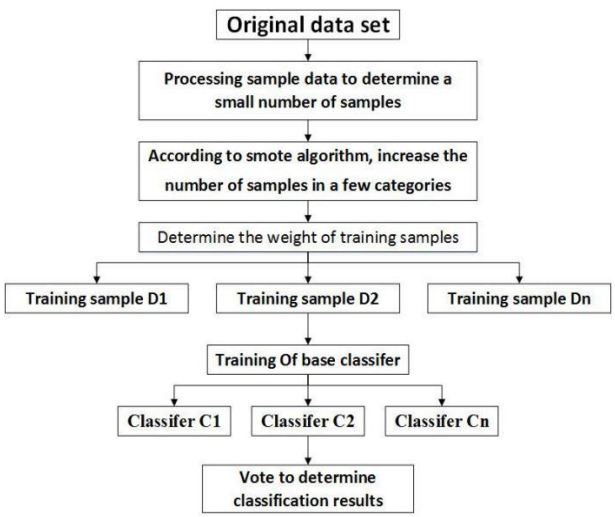

Figure 2. Flow diagram of Bagging improved algorithm.

This paper used MATLAB for performing experiments, divided the data into two parts: training set and prediction

set. Finally, the data are normalized by *X = normr (X)* to standardize the rows or columns of the data.

3) Simulation experiment

In this paper, the difference of gas content in transformer oil is taken as the basis of fault classification prediction, and the original data in reference [12] is taken as sample set. There are 180 fault samples, 60% data for training and 40% for testing.

The improved Bagging algorithm is used to train and learn 105 training samples. The weak classifier includes neural network and k-nearest neighbor, support vector machine, Bayesian classification and so on.

The simulation model is implemented by fitcensemble in the integrated learning toolbox of MATLAB (matrix laboratory), in which "good results = {'Total Boost','RUSBoost',...,'LPBoost','AdaBoost M2,'Bag'; Mdl = fit ensemble (X, Y,'Method,'Bag,..'Num Learning Cycles', 500,'Learners', t)".

When establishing Bagging's prediction model, cross-validation learning is conducted 500 times on the sample data, so that the results can achieve the best training effect and the prediction accurately approaches the ideal value. The iteration curve is shown in Fig. 3.

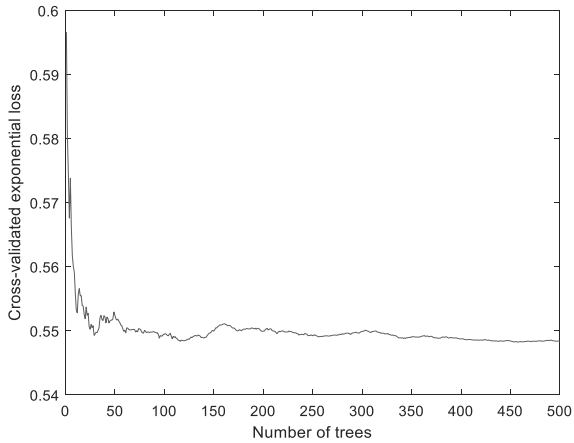

Figure 3. Iterative curve.

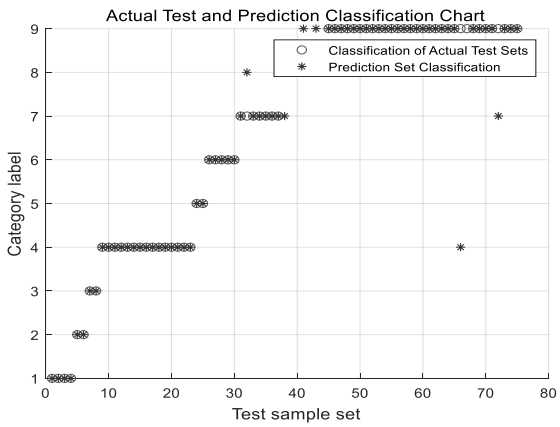

Figure 4. Classification chart of transformer fault prediction based on bagging improved algorithms.

In order to facilitate the visualization effect of experiment, the prediction results are visualized, as shown in Fig. 4. As can be seen from the figure, only 7 of 75 test samples do not match the actual results and they are concentrated in three fault states: low temperature overheating, medium temperature overheating and high temperature overheating. The remaining fault states can achieve 100% accurate prediction.

(There are nine different fault types which are as follows:1 partial discharge; 2 low-energy discharge; 3 low-energy discharge and overheating; 4 arc discharge; 5 arc discharge and overheating; 6 fault-free; 7 low-temperature overheating; 8 medium-temperature overheating; 9 high-temperature overheating.)

In order to prove the performance of the improved Bagging algorithm, this paper selects the same set of samples, and compares the accuracy of fault diagnosis and prediction by using Boosting derivative algorithms such as Total Boost, RUS Boost, LP Boost, Ada Boost M2 and SVMonly, BP Neural Network and Bagging, which are composed of single classifier. The results are shown in Table I.

TABLE I. COMPARISON OF FAULT DIAGNOSIS ACCURACY OF DIFFERENT METHODS

Method	Cross-validation	Train	Test
Total Boost	75.30%	100.00%	84.00%
RUS Boost	76.20%	100.00%	85.30%
LP Boost	77.10%	100.00%	85.30%
Ada Boost M2	76.20%	100.00%	86.30%
SVM only	73.30%	93.00%	81.30%
BP neural network	74.10%	100.00%	82.47%
Bagging	76.50%	100.00%	86.67%
Bagging Improved	79.10%	100.00%	86.67%

From Table I, it is easy to see that BP neural network and single classifier SVM only, have the worst prediction accuracy, which also shows that it is difficult to establish an accurate mathematical model for transformer fault diagnosis, and it is difficult to achieve good prediction results by using a single classifier, so it is necessary to use ensembled algorithm to improve prediction accuracy.

It can also be seen from Table I that under the same prediction conditions, the performance of Bagging algorithm and Boosting algorithm is basically close, and the prediction accuracy is in the range of 84%-86.3%. It shows that the integrated algorithm has superiority in transformer fault diagnosis and high prediction accuracy.

In this paper, an improved Bagging algorithm based on SMOTE is proposed. By increasing the number of samples of a few classes, the imbalance between a few classes and a majority of classes is reduced, and the weak learning algorithm is used to train the sample set to obtain a strong model. The strong machine learning model is used for transformer's fault diagnosis model to integrate multiple weak classifiers into a strong classifier. Through repeated learning of the model, the prediction accuracy of the ensembled model has improved, and the deep regularity features of the sequence are extracted to complete the fault identification of the records to be diagnosed. The simulation results show that the prediction accuracy of the improved Bagging algorithm

reaches 90.67%, which is better than other methods. It proves that this method has better adaptability and higher diagnostic rate. At the same time, it also proves the accuracy and validity of the algorithm.

D. Conclusions

The improved Bagging algorithm based on SMOTE proposed in this paper extends the classifier's classification plane to majority class space, avoiding the over-fitting of classifiers to some extent. While guaranteeing the accuracy of the whole classification, it also improves the performance of the minority classification. Different characteristics of five types of characteristic gas in oil before and after transformer failure, the improved Bagging algorithm proposed in this paper is applied to distinguish gas faults in transformer oil, which has good robustness and generalization ability and high fault diagnosis accuracy. The transformer fault diagnosis experiments with different methods show that the prediction accuracy of the improved Bagging algorithm is better than that other benchmark methods such as BP neuralnetwork and SVM , several boosting derivative algorithms, and classical Bagging algorithm. Its performance is more stable and reliable.

Acknowledgment

This work is supported by Sichuan university student innovation and entrepreneurship training program in 2018 (201811553020); Natural Science Foundation Project of Chengdu Textile College (2014fzlkb08); school level education and teaching reform research project of Chengdu Textile College (2014cdfzjj20) and Zhoukou normal university super scientific project 700.

Conflict of Interest

The authors declare no conflict of interest.

Author Contributions

Lu Peiwen, Huang Yongjing and Khushnood Abbas participated in the design of this study, and they both performed the statistical analysis. Huang Yongjing carried out the study and collected important background information. Lu Peiwen drafted the manuscript. All authors read and approved the final manuscript.

References

[1] P. Mirowski and Y. LeCun, "Statistical machine learning and dissolved gas analysis: A review," *IEEE Transactions on Power Delivery*, vol. 27, no. 4, pp. 1791-1799, 2012.

[2] H. Wang, H. Fan, Z. Yao, and C. Li, "Research of imbalanced data classification," *Application Research of Computers*, vol. 25, no. 5, pp. 31-41, 2008.

[3] J. J. Dukarm, "Transformer oil diagnosis using fuzzy logic and neural networks," in *Proc. Canadian Conference on Electrical and Computer Engineering*, 1993, pp. 329-332.

[4] E. Engineers and I. Board, *IEEE Guide for the Interpretation of Gases Generated in Oil-Immersed Transformers*, Piscataway, NJ, USA: IEEE, 2009.

[5] M. Duval, "Dissolved gas analysis: It can save your transformer," *IEEE Electrical Insulation Magazine*, vol. 5, no. 6, pp. 22-27, 1989.

[6] R. Rogers, "IEEE and IEC codes to interpret incipient faults in transformers using gas in oil analysis," *IEEE Transactions on Electrical Insulation*, vol. EI-13, no. 5, pp. 349-354, 1978.

[7] W. Zhu, D. Li, Q. Yan, and M. Pang, "Support vector machine based prediction of faults," *Coal Mine Machinery*, vol. 32, no. 3, pp. 253-254, 2011.

[8] H. Yin, "Transformer fault prediction based on dissolved gas in oil," master's thesis, Southwest Jiaotong University, 2018.

[9] L. Jiajia, "On-line monitoring method of transformer faults based on DGA technology," master's thesis, Jilin University, 2017.

[10] X. Ge, "Research on transformer online monitoring and fault diagnosis method based on DGA," master's thesis, Anhui University of Science & Technology, 2018.

[11] Y. Zhu, "Transformer fault prediction based on dissolved gas in oil," master's thesis, North China Electric Power University, 2013.

[12] Z. L. Wu, J. Yang, Y. L. Zhu, and Z. Liu, "Power transformer fault diagnosis based on rough set theory and support vector machines," *Power System Protection and Control*, vol. 38, no. 18, pp. 80-83, 2010.

[13] J. Jinglong, "Applied research of several neural network methods in transformer fault diagnosis," master's thesis, South China University of Technology, 2017.

[14] W. Chen, C. Pan, Y. Yun, and Y. Liu, "Wavelet networks in power transformers diagnosis using dissolved gas analysis, " *IEEE Transactions on Power Delivery*, vol. 24, no. 1, pp. 187-194, 2008.

[15] X. Mu, "Research on transformer fault diagnosis based on SVM," master's thesis, Anhui University of Science & Technology, 2017.

[16] M. Dong, Y. Y. Meng, C. X. Xu, and Z. Yan, "Fault diagnosis model for power transformer based on support vector machine and dissolved gas analysis," *Proceedings of the Chinese Society for Electrical Engineering*, vol. 23, no. 7, pp. 88-92, 2003.

[17] M. B. Jain, M. B. Srinivas, and A. Jain, "A novel web based expert system architecture for on-line and off-line Fault Diagnosis and Control (FDC) of transformers," in *Proc. Joint International Conference on Power System Technology and IEEE Power India Conference*, 2009, pp. 1-5.

[18] H. Xiong, W. G. Li, G. H. Chang, and H. M. Guo, "Application of fuzzy rough set theory to power transformer faults diagnosis," *Proceedings of the Chinese Society for Electrical Engineering*, vol. 28, no. 7, pp. 141-147, 2008.

[19] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: Synthetic Minority Over-sampling Technique," *Journal of Artificial Intelligence Research*, vol. 16, pp. 321-357, 2002.

[20] P. Jeatrakul, K. W. Wong, and C. C. Fung, "Classification of imbalanced data by combining the complementary neural network and SMOTE algorithm," in *Proc. International Conference on Neural Information Processing*, 2010, pp. 152-159.

[21] J. Wang, M. Xu, H. Wang, and J. Zhang, "Classification of imbalanced data by using the SMOTE algorithm and locally linear embedding," in *Proc. 8th International Conference on Signal Processing*, 2006.

[22] X. Zhu and Y. Yang, "A lazy bagging approach to classification," *Pattern Recognition*, vol. 41, no. 10, pp. 2980-2992, 2008.

Lu Peiwen was born in March 1998, Chengdu, Sichuan Province, China. She has completed her undergraduate degree from the School of Electrical and Information Engineering at Chengdu Textile College 2019. Now she is a master student in the School of Electrical and Electronic Data of Xihua University.

She has won two national inspirational scholarships and won the second prize in the National College Students' Electronic Competition.

Ms. Lu Peiwen is a member of ORCID. Her research field are electrical engineering and automation, transformer fault prediction, transformer fault overview.

Huang Yongjing has earned his Ph.D from Sichuan University in 2015. Currently he is an associate professor in Chengdu Textile College. His research fields are mechanical and electrical integration and control technology.
Mr. Huang Yongjing is a member of ORCID.

Khushnood Abbas received his Bachelor and Master degree from Aligrah Muslim University, India. He has completed his doctoral degree from School of Computer Science and Technlogy, University of Electronic Science and Technology in 2018. Currently he is working as an Assistant professor in Zhoukou Normal Univeristy, Henan China. His research fields are neural network and computer technology.
Mr. Khushnood Abbas is a member of ORCID.