Application of Support Vector Machine (SVM) Algorithm in Fault Diagnosis of Smart Grid Transformer

Li Wan[1, a, *] and Guobin Wan[2]

[1]State Grid Binzhou Zhanhua Power Supply Company; 256800, China

[2]State Grid Guangrao Power Supply Company 256800, China

[a]762276848@qq.com

*Corresponding author

Keywords: DGA, support vector machine, transformer, fault diagnosis, parameter optimization, SVM model

Abstract: In view of the different types and contents of dissolved gases in transformer oil under different working conditions of oil-immersed power transformers, detecting the types and contents of different gases in transformer oil has become an important method to determine the working status of transformers. Based on the support vector machine (SVM) model, this paper uses the support vector machine's excellent solution to non-linear problems, and presents a method to determine the working mode of the transformer. After testing, the accuracy rate is as high as 96.8%. Accuracy requirements.

1. Introduction

The power transformer is an important part of the power system. The quality of its operation is related to the reliability of the power system. Once the power transformer fails, it will cause huge economic losses. There are various diagnostic methods adopted after a transformer failure. The Dissolved Gas Analysis (DGA) method is one of the most effective methods to detect a transformer failure. It can timely find the internal faults in the transformer and maintain the transformer at the same time. It can also eliminate hidden troubles in the transformer. When the transformer fails, different test methods are used for fault analysis and diagnosis through the generated gas. The characteristic gas method, the Rogers ratio method, and the improved three-ratio method are the main diagnostic methods for transformer faults based on DGA. There are codes for fault diagnosis. The problem of blind spots cannot diagnose multiple faults at the same time. Therefore, in order to solve the blind spot problem of the ratio method, a new fault diagnosis method is proposed to diagnose transformer faults [1].

Support vector machine algorithm is based on statistical learning theory. It can solve nonlinear and high-dimensional machines. Support vector machine is used to classify transformers for fault diagnosis. The selected parameters have a greater impact on the results of fault diagnosis. The DGA-based transformer fault diagnosis method cannot accurately diagnose the fault location of the transformer. Based on this, this paper proposes a transformer fault diagnosis method based on particle swarm optimization support vector machine and a differential evolution support vector machine method for transformer fault diagnosis. It diagnoses the fault of the transformer and determines the location of the fault as soon as possible. Ensure the normal operation of the power system [2].

2. Support Vector Algorithm Introduction

The support vector machine (SVM) method proposed by Vapnik et al. [3] is regarded as a good alternative to the traditional learning classification method, especially in small samples and non-linear cases, it has good generalization performance. The SVM method is based on the VC dimension theory of statistical learning theory and the principle of structural risk minimization, and seeks the best compromise between model complexity and learning ability based on limited sample information to obtain the best generalization ability. As a new learning classification method, SVM has been successfully applied in the field of fault diagnosis.

Support vector machines are proposed for two classification problems. For multi-classification problems, there are two main solutions. One is a one-time solution method, that is, directly establishing a multi-classification objective function for optimization. This method is used to solve the optimization problem. There are many variables used in the process, which is not practical due to the high computational complexity. The second is to transform the multi-classification problem into multiple two-classification problems for solving. This scheme mainly includes a class of pair-of-class method, paired classification method, error correction output coding method and hierarchical SVM multi-class classification method. The first three methods all need to construct more SVM two classifiers, and there are unrecognizable domains. Therefore, this paper designs a vector machine multi-class classifier for transformer fault diagnosis. In order to make the extracted features not affected by the amplitude of the transformer output signal, the signal is standardized before EMD decomposition. The specific feature extraction steps are as follows:

(1) Standardize the transformer signal $\boldsymbol{X}$:

$$\tilde{\boldsymbol{X}} = D_\sigma^{-1}[\boldsymbol{X} - E(\boldsymbol{X})] \tag{1}$$

Among them, $\boldsymbol{X}$ represents the output signal sequence of the transformer, $E(\boldsymbol{X})$ is $\boldsymbol{X}$ the corresponding mean value, and D_σ is the standard deviation of $\boldsymbol{X}$.

(2) Perform EMD decomposition on $\tilde{\boldsymbol{X}}$ to extract the first 5 IMF components, $\boldsymbol{C_1},\boldsymbol{C_2},\boldsymbol{C_3},\boldsymbol{C_4},\boldsymbol{C_5}$ is the G5 IMF component sequence, and $\boldsymbol{C_6}$ is the residual sequence.

(3) In order to enhance the fault characteristics of the IMF component, the IMF component and the residual term are reduced. Calculate the reduction threshold for each IMF component and residual term:

$$ThrC_i = \sqrt{\frac{1}{m}\sum_{j=1}^{m} C_{i,j}} \tag{2}$$

Where m represents the length of the IMF component and the residual term, and $C_{i,j}$ represents the value of the j point of the first IMF component.

Perform the following reduction processing on each IMF component and residual term, and find the reduction ratio of each IMF component and residual term [4].

$$\begin{gathered} \tilde{C}_{i,j} = \begin{cases} C_{i,j} & \left|C_{i,j} \geq ThrC_i\right| \\ 0 & \left|C_{i,j} < ThrC_i\right| \end{cases} \\ CutDC_i = \frac{Num(\tilde{C}_{i,j} \neq C_{i,j})}{m} \end{gathered} \tag{3}$$

Among them, $\tilde{C}_{i,j}$ is the reduced IMF component and the residual term, and $CutDC_i$ is the reduction ratio of the corresponding component, which is the ratio of the number of points that are cut and the total points.

3. Principle of transformer fault diagnosis system

Because the dissolved gas content in transformer oil is different under different working conditions of transformers, analysis of dissolved gas in transformer oil is needed, which can well reflect some potential failures of the transformer that are not usually detected. The improved three-ratio method proposed in has the highest judgment accuracy rate, namely the three contrast values of CH4 / H2, C2H2 / C2H4 and C2H4 / C2H6. In different working modes, the difference between the three teams' ratios is the largest. By analysing these three comparison values, the working mode of the transformer can be accurately identified, and the cause of the fault can be correctly judged. Support vector machine (SVM) is used to train and learn the known data, identify the differences between the three contrast values under different working modes, and finally output the analysis results. The SVM network design process is shown in Figure 1.

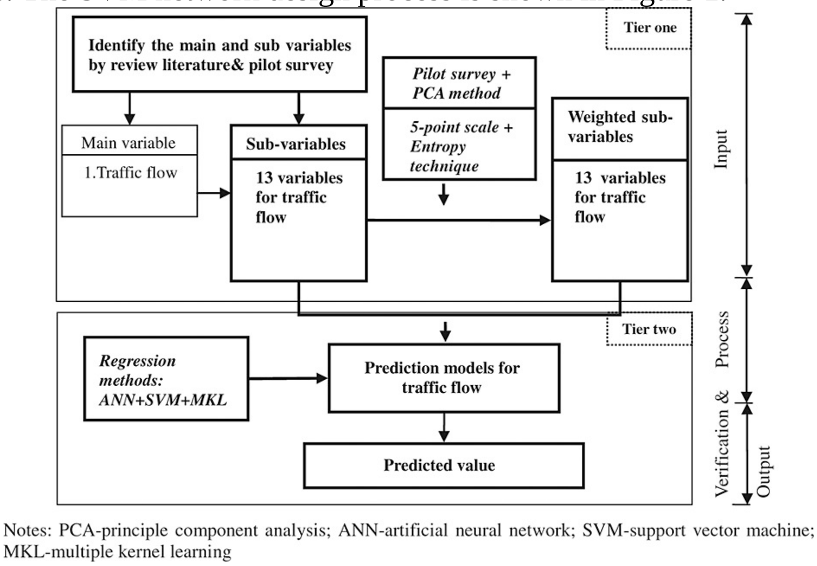

Figure 1. SVM network design flow

The working mode of the transformer is mainly divided into three failure working modes and normal working modes, as shown in Table 1.

Table 1. Classification of working modes

Operating mode	overheat	Low energy discharge	High energy discharge	normal work
Numbering	1	2	3	4

4. Troubleshooting methods

Based on DE's global search capabilities, a model is established based on the sample misjudgement rate as the standard, and the optimal kernel parameters and penalty factors are found for the support vector machine classifier, which overcomes the problem of low efficiency of the traditional support vector machine classifier and improves it To improve the practicality of support vector machines. The support vector machine based on DE focuses on the input and output, and the mapping relationship between the input and output is mainly completed by the support vector machine. In the process of modelling the support vector machine, the DE algorithm will generate penalty factors and kernel parameters, and the error between the output of the support vector machine and the actual value is modelled as the objective function. The modelling steps are:

1) Classify the sample parameters and divide them into training and test sets, and then unify them after classification; 2) Initialize them using differential evolution algorithm and set the size and dimension of the population; 3) Use the RBF function as the kernel function, input the training set into the support vector machine, compare the trained value with the actual value as the objective function of the DE algorithm, and determine whether the result meets the conditions to obtain the optimal parameter value; 4) Use the obtained optimal parameter values as the parameters of the support vector machine, enter the test set, and use the differential evolution algorithm to output the test results; 5) Compare the results of the test set with the predicted values, and judge the accuracy of the results. The parameter optimization process is shown in Figure 2 [5].

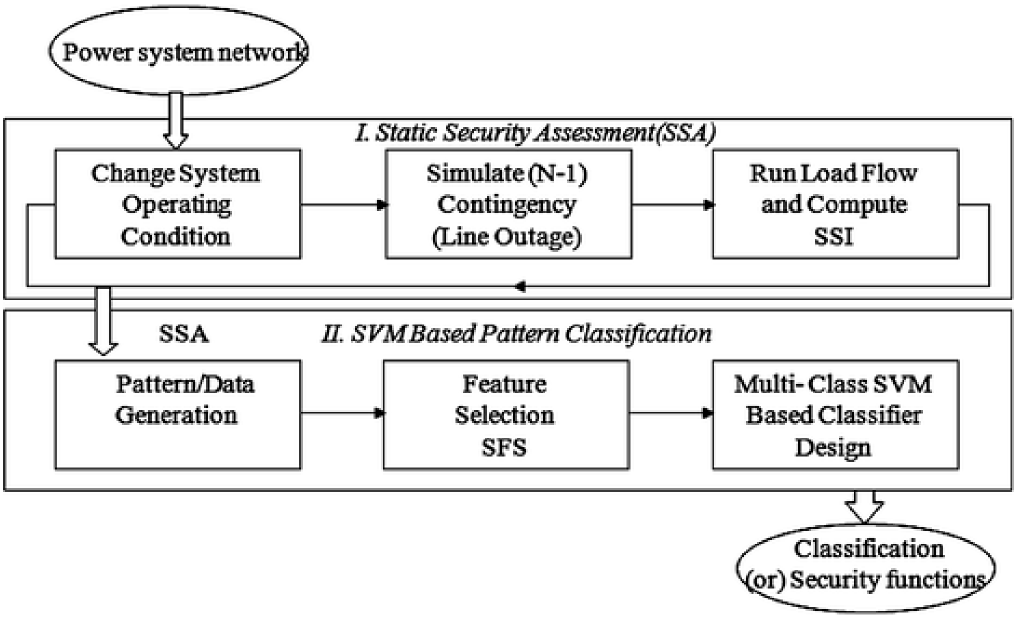

Figure 2. Parameter optimization process

5. Realization of transformer fault diagnosis

The characteristic gases dissolved in oil used for transformer fault diagnosis are: H2, CH4, C2H6, C2H4, and C2H2, which carry the high temperature overheating, high energy discharge, low energy discharge, partial discharge, and medium temperature of power transformers to a certain extent Low temperature overheating 5 kinds of fault information, so it can be used as a characteristic gas.

When using support vector machines to analyse and diagnose transformer faults, the data needs to be pre-processed. Some data may be large or small, and may be several orders of magnitude different from the data of other groups. Each parameter in different working modes collected by sensors and acquisition devices is a sample, and the corresponding working mode is the output result. In this paper, a total of 31 sets of data are extracted. The first 23 sets are selected for training and learning, and then all the data are collected. Disrupt and re-evaluate.

First load the data, select the training data and test data, divide the training data into training input vectors and training target vectors, replace the training data, and complete the prediction of the test data. The training results are shown in Figure 3. The abscissa is the sample number of each group, and the ordinate is the category. The asterisk line indicates the prediction operation mode, and the circle indicates the real operation mode. The resulting confusion matrix is shown in Figure 4.

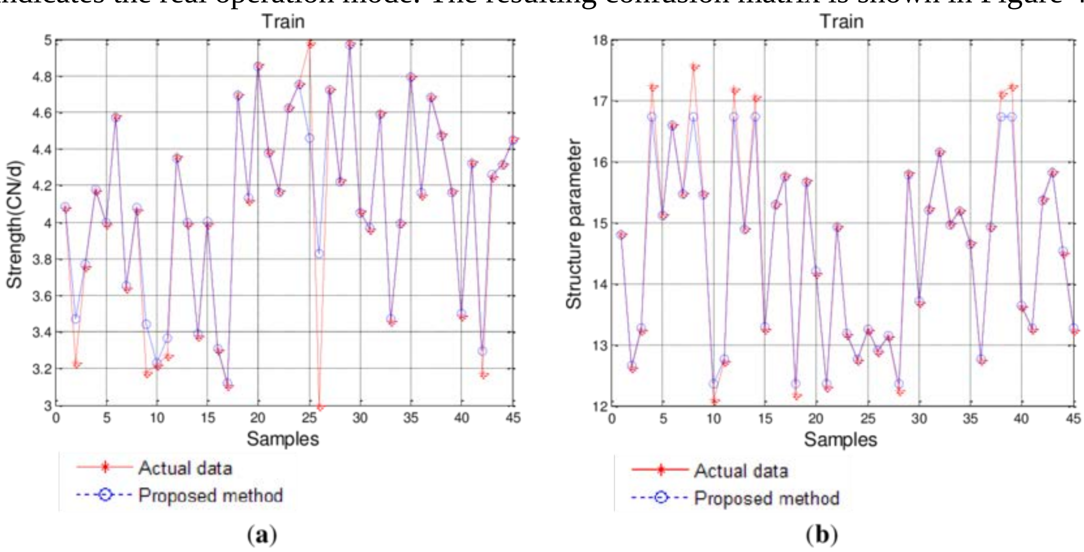

Figure 3. SVM training results

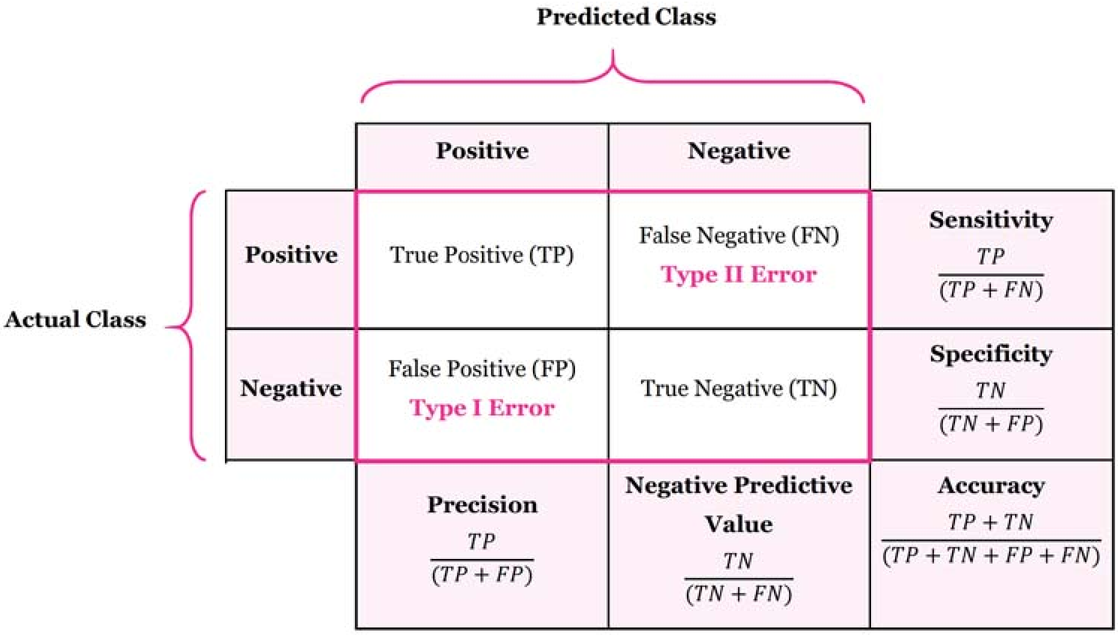

Figure 4. Confusion matrix

The first to fourth rows of the matrix represent the prediction and recognition of the four working modes. The light gravy squares represent the number and proportion of correct predictions, and the gravy squares represent the number and proportion of incorrect predictions for each category. The gravy squares represent the accuracy of each type of population, and the final white squares represent the prediction accuracy of the population. It can be seen from Fig. 3 that only the last sample has a wrong prediction, and the real working mode 1 is identified as the working mode 4. The confusion matrix in Figure 4 can clearly see the prediction accuracy rate of the various working modes and the overall accuracy rate of 96.8%. Meet the accuracy requirements for the actual transformer fault diagnosis.

The volume fractions of hydrogen (H2), methane (CH4), ethane (C2H6), ethylene (C2H4), and acetylene (C2H2) in large power transformer oils are used as feature values. The state output is the same as that in the particle swarm optimization algorithm. This is used as the state output of the SVM. The training set and the test set are each divided into 28 groups. According to the number of samples in the training set and the test set, the characteristic gas distribution shown in FIG. 4 is obtained. When the temperature is low, methane accounts for a large proportion; as the temperature gradually increases, the volume fraction of ethylene increases significantly; if the temperature is too high, methane and ethylene will be produced; when a local current occurs, hydrogen and methane will be produced; During the spark discharge, the gas dissolved in the oil is mainly ethane and hydrogen; when the arc discharge occurs, the gas produced is methane, hydrogen, and ethane. The experiment is consistent with the operation results, which indicates that the experiment of the sample data is time-effective.

6. Conclusion

Aiming at the shortcomings of traditional power transformer fault diagnosis, this paper proposes a support vector machine power transformer fault diagnosis model based on small samples and strong generalization ability. The fault diagnosis process includes the four steps of determining the fault state of the power transformer, normalizing the data, optimizing parameters, establishing a model, and testing. The diagnosis of actual fault samples shows that the model has high classification accuracy and strong practicability.

References

[1] H. Xue, K. Zhang, B. Li, & C. Peng. Fault diagnosis of transformer based on the cuckoo search and support vector machine. Power System Protection & Control, 43(8) (2015) 8-13.

[2] Su, G., Fang, D., Chen, J., & Rui, Z. Sensor multi-fault diagnosis with improved support vector machines., 14(2) (2017) 1053-1063.

[3] Ming, Y. C., & Hoang, T. T. A differential particle swarm optimization-based support vector machine classifier for fault diagnosis in power distribution systems., 17(3) (2017) 51-60.

[4] YUAN Haiman, LEI Fan, CHEN Yu, CHEN Mingxing, JIANG Danyu, & GAO Bo. Fault diagnosis of transformer using relevance vector machine with particle swarm optimization based on dga. High Voltage Apparatus, 53(2) (2017) 108-112 and 119.

[5] H. Fu, Z. Qi, & R. Ren. A new fault diagnosis method for transformer based on improved rvm. Journal of Liaoning Technical University, 36(9) (2017) 964-970.