Article

A Review of Transformer Fault Diagnosis Based on Information System Theory and Machine Learning

Xiaoli Huang [1],*, **Yumiao Yuan** [2], **and Jingyu Li** [3]

[1] Research Institution of Signal Detection and Information Processing Technology, Xihua University, Chengdu 610039, China; 6607784@qq.com

[2] Research Institution of Signal Detection and Information Processing Technology, Xihua University, Chengdu 610039, China; yumiaoyuan2022@163.com

[3] Research Institution of Signal Detection and Information Processing Technology, Xihua University, Chengdu 610039, China; 3564594137@qq.com

* Correspondence: 6607784@qq.com; (X.H.)

Abstract: Safe, high-quality and economical electric energy transportation is the basic requirements of modern power system operation. Transformer, as one of the core equipment of power system and national grid system, has the characteristics of diverse types, variable models and wide deployment. It is the basic equipment for power system to realize voltage change and electric energy distribution. Because the power system transformer needs to operate with load for a long time, the probability of failure is usually higher than that of other power equipment in general. This paper starts from the transformer fault and fault diagnosis. Firstly, the transformer fault types, traditional transformer fault diagnosis techniques and the advantages of machine learning in transformer fault diagnosis are reviewed. Secondly, the application of information system theory and information entropy in transformer fault diagnosis is introduced. Then it introduces machine learning technology, feature selection and extraction technology in transformer fault diagnosis, and machine learning algorithms such as support vector machine and extreme learning machine for transformer fault diagnosis. Finally, the potential development trend of information system theory, information entropy and machine learning in transformer fault diagnosis is forecasted. The combination of transformer fault prediction and machine learning algorithm is helpful for power system maintenance personnel to accurately predict the running state of power equipment, and also provides a new method and technology for the safe and reliable operation and regular maintenance of power system.

Keywords: transformer fault diagnosis; information system theory; information entropy; machine learning

1. introduction

Safe, high-quality and economical electric energy transportation is the basic requirement of modern power system operation. Transformer, as one of the most critical core equipment in the power system and the national grid system, has the characteristics of diverse types, variable types and wide deployment. It is the basic equipment for power system to realize voltage change and electric energy distribution. Because the power system transformer needs to operate with load for a long time, the probability of failure is usually higher than that of other power equipment in general. At the same time, due to the upgrading of China's current power grid system, there are more and more cross-regional networking and dispatching. If the transformer fault cannot be diagnosed and repaired in time, once the fault occurs, it is easy to cause a local chain reaction in a certain power grid. Therefore, carrying out daily fault detection and diagnosis for transformers is a necessary means to help power grid staff to carry out repair before transformer faults occur [1].

Commonly used transformers in power systems are mostly operated under direct exposure and natural environment with load for a long time. Therefore, it is impossible to

avoid some failures caused by the aging of parts, such as damp terminal rows due to rain and snow weather, aging or cracks in casing, and oil leakage in tap switches. If these faults are not found and removed in time, it is easy to cause power system faults in the long-term risky operation process, and serious ones may even cause large-scale power outage of the power grid system [1].

Represented by the United States, Canada, Japan and other countries, have developed different types of online monitoring equipment, which have been applied to some extent in practice. Although online monitoring technology in China started a little late, it has developed rapidly in the past 20 years, and different types of monitoring equipment have been developed by relevant departments of power and scientific research institutions [2]. The original fault diagnosis of transformers began in the early 20th century, mainly relying on the senses and experience of professionals with simple meters to judge the fault. Due to the high technical level requirements of operators, this method is vulnerable to the interference of external environmental conditions, and has a large number of uncontrollable factors. In the 1960s, diagnostic technology based on sensors and computer applications began to flourish. The theory and practice of signal detection, data processing and signal analysis gradually constitute the main research and development content of diagnosis technology at this stage. However, because the guidance of experts and professional technicians is still required, the final decision is greatly influenced by the knowledge of the diagnosis domain and the strategy for dealing with the problem. In the 1980s, the research results of artificial intelligence began to be applied to the field of fault diagnosis, and intelligent diagnosis technology appeared. During this period, various effective methods and means of related disciplines were widely adopted [2].

As the core electrical equipment in the power transmission and conversion system [3], under the condition of the widespread access of a high proportion of renewable energy, the problems of power transformer such as state perception and fault early warning are increasingly exposed [4].The power transformer failure will directly lead to the power system can not run safely and stably [5]. Therefore, it is of necessity to study the fault diagnosis of power transformer [6].

2. Transformer fault diagnosis

Power transformers often appear in the winding inter-turn fault, impact, damp, over-heating, winding open, grounding, mechanical failure and other faults; In the terminal row connection loose, short circuit, damp, lead break and other faults; In the casing aging, chapped, damp, low oil level, flange grounding and other faults; In the core insulation fault, parts loose, core laminated short circuit and other faults; Mechanical failure, overheating, lead failure, electrical failure, physical damage and other faults occur in the tap switch [1].

Transformer fault detection is the premise of diagnosis, including traditional off-line detection and modern on-line monitoring. Among them, off-line detection is the most obvious method used in the field, and it is also the commonly used method of power transformer preventive test. Online monitoring can pay attention to the fault situation of the transformer in real time, so as to achieve early detection and early treatment, which has outstanding advantages compared with conventional detection methods [1].

Transformer fault diagnosis plays an important role in tracking monitoring data and transformer health status [7]-[8]. In the operation process of the transformer, because the insulation material is affected by the outside world, it will decompress to produce various gases, and the specific gas content will increase when the transformer is broken down. This kind of gas is called the characteristic gas, and the characteristic gas is CH_4, C_2H_2, C_2H_4, C_2H_6, CO, CO_2, H_2, etc. [9]. Under normal operating conditions, the amount of gas decomposed by insulating materials is very small, but when the power transformer is not running normally in the power system, the increase of the transformer oil temperature and current speed up the decomposition and aging deterioration of insulating oil and solid insulating materials, so that the content of dissolved gas in the transformer oil will change [10]. By monitoring and analyzing the content of dissolved gas in transformer oil,

the operation state of transformer can be judged with effect [11]. The main method of transformer fault diagnosis,dissolved gas analysis in oil (DGA), has been widely applied in transformer fault analysis of Chinese power system [12].

The traditional method is the method of transformer fault diagnosis by analyzing dissolved gas after people summarize the data obtained in continuous experimental research. Traditional transformer fault diagnosis methods mainly include characteristic gas identification method and ratio method.

The characteristic gas identification method is a traditional transformer fault diagnosis method based on the type of dissolved gas. The principle of this method is that when the transformer fault of different fault types occurs, the influence degree of the transformer overheating and electricity and other influencing factors is different, resulting in the main dissolved gas composition and secondary gas composition of the insulating material in the transformer oil are different, and the difference in the composition of the dissolved gas can be determined by analyzing the fault of the transformer [11].

The ratio method is a conventional method for transformer fault diagnosis based on the proportion of dissolved gases in the oil. The uncoded ratio method does not adopt the coding corresponding to the dissolved gas ratio interval, but adopts the direct use of a specific ratio range corresponding to a specific transformer fault type. Because the ratio method abandons the relatively small probability in the process of statistical analysis of a large number of transformer faults, the accuracy of transformer fault analysis based on the ratio method is not high [11].

Although the above traditional transformer fault diagnosis method has achieved some results. However, most of them rely on fixed weights and thresholds, which can not precisely mirror the relationship between the performance and the characteristics of the fault and the objective law. With the rapid development of science and technology, more and more machine learning algorithms are applied to the fault diagnosis of transformers. Therefore, it is particularly important to use intelligent algorithms combined with traditional transformer diagnosis methods for fast and accurate intelligent diagnosis and to predict the possible future faults of transformers through the flexible use of DGA method. For example, artificial neural network [13]-[17], Support Vector Machine (SVM) [18], Bayesian network [19]-[21] and random forest [22] et al.

Li et al. [23] proposed a transformer fault diagnosis model based on improved Grasshopper optimization algorithmOptimized Support Vector Machine (SVM). The improved Grasshopper Optimization Algorithm (IGOA) was used to optimize the kernel function parameters and penalty coefficient of SVM. A transformer fault diagnosis model based on Dissolved Gas Analysis (DGA) was established in the SVM optimized by oil-based IGOA algorithm, and the effectiveness and superiority of IGOA-SVM in identifying transformer fault states were verified by comparing with PSO-SVM and GOA-SVM. DemirciMerve et al. [24] combined gas data classified by a machine learning approach with a sensor fusion approach to improve the estimation accuracy by more than 0.1 using a sequential Kalman filter. Hu [25] et al. used slime mold algorithm (SMA) to select the characteristic wavelength of transformer oil fluorescence spectrum, and built a transformer fault diagnosis model on this basis. They demonstrated the advantage of SMA in the characteristic wavelength screening of transformer oil fluorescence spectra, and used SMA to screen the characteristic wavelength of transformer oil LIF spectra and used it for transformer fault diagnosis. Liu Rongsheng et al. [26] combined the advantages of neural network and evidence theory through cloud computing method to compare chromatographic data and electrical test data of transformers. A transformer fault comprehensive diagnosis method based on multi-neural network and evidence theory based on cloud computing is proposed. The fault diagnosis performance of a transformer is measured by inspecting the faulty transformer. The results show that compared with the traditional single data alignment, the proposed method can improve the reliability and accuracy of diagnosis by comparing multiple kinds of data.

The most basic requirement of modern power system operation is more safe, higher quality and more economical, and transformer is one of the most nuclear equipment of power system. The traditional transformer fault detection and diagnosis technology is single, the engineering practicability is not strong, and the fault recognition rate is low, which can not meet the current power grid system upgrade in China. Carrying out daily fault detection and diagnosis for transformers is a necessary means to help power grid staff to carry out repair in time before the transformer fault occurs, which has important guiding significance for the power system. Therefore, the use of machine learning algorithms for transformer fault diagnosis needs further research and improvement [1].

3. Transformer fault diagnosis based on information system theory

Information Systems Theory (IST) is a conceptual framework for analyzing and designing information systems to support business processes. IST can also be applied to transformer fault diagnosis, providing a systematic approach to modeling and analyzing information flow in power systems. The primary goal of IST in transformer fault diagnosis is to identify the key information needed to make an accurate and timely diagnosis.

The basic idea of applying information system theory to fault diagnosis is to use an evaluation function, that is, to provide the information about the fault location, take the maximum information as the target, and find the largest as the object of doubt. The user then tests the part, stops if it is a failure, finds another fault with the most information, and repeats until the fault is found. This process uses numerical calculation instead of symbol matching, so the speed of reasoning is improved [27].

Information entropy can measure the uncertainty of random field. Entropy is a measurement unit of disorder degree, and its concept is derived from thermodynamic statistics. Thermodynamic entropy and information entropy express different contents and are fundamentally different in definition. In information theory, entropy aims at describing source uncertainty and quantifying the change degree of system state with entropy function. The mathematical expectation of self-information is defined as information entropy, that is, the average information of information source [28].

The physical meaning of entropy is to measure the average uncertainty of the whole system, which reflects the change degree of the statistical characteristics of the information source system. The different occurrence probability of an event makes the uncertainty of the event itself different. From the definition of information entropy, it can be represented that each event has different information entropy. The entropy value is not determined by the actual value of, but by its probability distribution. The entropy value also reflects the average amount of information [28].

In transformer fault diagnosis, the application of entropy has been widely studied and applied. Through the entropy analysis of transformer internal signal, the transformer fault type and health state can be judged, and the optimal maintenance strategy and the minimum maintenance times can be obtained [28].

Transformer fault is mainly divided into winding fault, lead fault, iron core insulation fault, sleeve fault, protection system fault, cooling system fault, lax seal, other faults main insulation fault, box fault, etc. According to the historical data of transformer faults, the information entropy values represented by various faults are calculated. By combining the information entropy theory with the transformer fault model, the transformer fault diagnosis model based on information entropy is obtained. By calculating the information entropy of various fault events, the importance of various faults can be quantitatively analyzed. By analyzing the event importance, the maintenance strategy can be optimized, the reliability of the system can be improved, and the important parts of fault detection can be determined [28].

Entropy can also be used to assess the operating condition and health of transformers. By long-term monitoring and analysis of the entropy value of the transformer, the operation trend and rule of the transformer can be found, and the potential fault risk can be

determined. Maintenance measures can be taken in time to ensure the normal operation and extend the service life of the transformer.

The normal operation of all the equipment in the power grid is a strong guarantee for the safety and stability of the power system. As the most important equipment in the system, the power transformer needs to carry out a lot of detection tests in its production and operation process [28]. In the process of equipment manufacturing, it is necessary to carry out transformer type test, equipment routine test and transformer factory test. A handover test should be carried out before entering the operation phase; For transformers in operation, preventive tests, diagnostic tests and on-line monitoring are required according to regulations. The purpose of the test is to grasp the transformer insulation situation in time, find out the possible problems of the equipment in advance, and conduct the corresponding maintenance, to ensure the normal operation of the equipment, reduce the operation risk and maintenance cost. Based on the fault that can be detected by each detection method in transformer test item, the information entropy that each detection item can provide is obtained. The maintenance model based on information entropy is obtained by calculating the maintenance method, and the optimal transformer maintenance strategy based on information entropy is obtained according to the order of the size and amount of information of maintenance information entropy.

The traditional transformer maintenance strategy is mainly based on preventive test and diagnostic test related experiment items, some experiment items repeatedly detect the same fault resulting in excess maintenance costs. This phenomenon is caused by the lack of overall consideration of maintenance strategy in the field design of transformer maintenance test scheme. To design transformer testability, it is not only necessary to consider the effectiveness of detection methods, but also more important to sort transformer maintenance means, and maximize the accuracy of diagnosis and prolong the maintenance cycle of equipment. In order to optimize transformer maintenance strategy, it is necessary to fully consider maintenance cost, detection difficulty and other objective factors, reasonable use of maintenance resources. It is hoped that the fault location can be found accurately with the least number of maintenance times to ensure the safe operation of the equipment [28]. In addition, in order to rank maintenance methods, it is necessary to evaluate the changes in information entropy brought by each test method and determine the maintenance methods used below according to the changes in different entropy values. On this basis, the set of minimum maintenance times under the same changes in information entropy is found.

All test methods are set according to the accuracy of fault diagnosis. After the test, it can be determined whether the corresponding fault is diagnosed by the detection method. The information drop value of each detection method is determined according to the information entropy value of fault diagnosis by the corresponding detection method. The number of detection required and the ranking of detection means will also vary accordingly. After comprehensive consideration of various maintenance schemes, the most appropriate maintenance strategy is found.

One of the main advantages of applying information system theory and information entropy to transformer fault diagnosis is that it can improve the accuracy of fault diagnosis while reducing the risk of false positives and false negatives. By modeling the information flow between the components of a power system, the key information needed for accurate fault diagnosis can be identified. These key information can be used for accurate and timely diagnosis, improve the reliability and efficiency of the power system [28].

4. Transformer fault diagnosis based on machine learning

Machine learning processes and transforms information and knowledge from the external environment into effective information content. Learning is the core of machine learning, including information collection, supervision and guidance, reasoning of learning and modification of knowledge base of learning system. Machine learning is purposeful, structured, effective and open. Machine learning can understand the content of learning, learning behavior has a certain purpose. In structure, the structure and organizational

form of knowledge can be modified and perfected systematically. In terms of effectiveness, the knowledge learned by machine learning system can adapt to the needs of practice and constantly improve the machine learning system. Openness is reflected in machine learning's ability to interact with the external environment and evolve itself [29].

Machine learning, as the core technology of artificial intelligence, uses computers to learn rules and patterns in massive data and dig out potential information from it [30]. At present, widely used machine learning methods mainly include traditional machine learning, integrated learning [31] and deep learning. Traditional machine learning can be divided into supervised learning, semi-supervised learning and unsupervised learning depending on whether the data is labeled or not. Supervised learning uses a set of labeled data sets to train the algorithm, and obtains an optimal hypothesis from it, and then uses the hypothesis to process the unknown data to realize the corresponding function [32]. In addition to labeled training data, semi-supervised learning also uses some unlabeled data for training. Unsupervised learning only receives the unmarked data and processes the input data based on the assumptions of the data to achieve certain functions [33].

The basic idea of ensemble learning [34] algorithm is to combine multiple isomorphic weak learners into a strong learner so that its performance exceeds that of any weak learner. Multiple weak learners are constructed through multiple iterations, and the weights of the next generation learner samples are affected according to the target test results [35]. Different weights are allocated to the samples for in-depth data mining. Finally, weights are voted to build a stronger learner to complete the specified task [32].

Hinton and his colleagues proposed deep belief networks [35] in 2006, which mimic brain mechanisms by building neural networks with multiple hidden layers to further interpret data, including images, sounds and text. Deep learning methods can learn more features and parameters without over-fitting. Therefore, it is widely used in image processing, speech recognition, text recognition and other fields [36].

At present, machine learning and related technologies have achieved great success in many fields, such as computer vision, artificial brain and speech recognition. Machine learning models are also widely used for important real-world tasks [37]. For example, face recognition [38]-[40], automatic driving [41], malware detection [42] and intelligent medical analysis [43]. It is foreseeable that machine learning will be the key technology and core creativity in the development of artificial intelligence, and will play an important role in promoting it.

With the increasing number of measuring devices in the power system, the amount of relevant power data available from the grid grows exponentially every year. In addition, more and more distributed energy is continuously connected to the grid, and the adjustable flexible load increases significantly, which makes the regulation and control of the grid more flexible. With the trend of power grid developing towards mass data and high intelligence, emerging technologies represented by machine learning have gradually become a research hotspot [44]-[45]. Machine learning analyzes, processes and feeds back information and data based on large amounts of data collected. It can make full use of multi-source heterogeneous mass data, and has a good effect in the learning and processing of nonlinear problems, so as to effectively coordinate the functions and demands of power generation, transmission, distribution and use in the power grid. In this context, domestic and foreign scholars have carried out relevant scientific research on how to use machine learning to promote the development of power system. More and more new techniques and methods of machine learning are applied in power system. Including state estimation, transmission and distribution equipment fault diagnosis, power system prediction, power distribution system operation optimization, etc., machine learning methods are popular research directions in power system [32].

With the increasing of distributed generation, the measurability and controllability of distribution network become more and more important. In order to control and optimize the distribution network and realize the interaction with users, it is necessary to use automatic means to collect a large number of network measurement data to complete various

calculation and analysis functions. However, problems such as loss of topology information and uncertainty of measurement set often occur in power distribution system. LeiWang et al. propose a physically-guided model that combines machine learning approaches with established physics-based approaches in a hybrid model to enhance the interpretability of data-driven models [46]. Compared with the traditional method, the learning ability of deep neural network is fully utilized to model the time correlation between system states. In addition, unlike pure data-driven supervised learning, nonlinear power equations are added and the deviation of estimated measurements from actual observations is taken as the loss of the trained neural network. The proposed method shows higher accuracy and is more robust to corrupted data.

Fault diagnosis of transmission and distribution equipment is the core task related to the safe operation of power grid. It is difficult to use traditional mathematical methods for modeling. However, machine learning technology can quickly judge the running state of equipment by mining the information in monitoring data. WanghaoFei et al. proposed a current tracking method by modeling a single distribution feeder as several independent parallel virtual lines to track the detailed contributions of different current sources to the line current [47]. Then the enhanced current information is used as the extended feature space of support vector machine, which significantly improves the ability of fault current monitoring on the power line. Simulation results show that the proposed method is sensitive to low level fault current. CepedaCristian et al. proposed a microgrid intelligent fault detection system based on local measurement and machine learning techniques [48]. An integrated machine learning model provides a level of intelligence for smart electronic devices installed on microgrids. Allowing each smart electronic device to autonomously determine if a fault has occurred on the microgrid eliminates the need for communication infrastructure between electronic devices for microgrid protection.

With the continuous development of machine learning technology, more and more scholars combine load prediction, power generation prediction and electricity price prediction with machine learning. These methods play a vital role in dealing with uncertainty and risk management in the distribution network. SanaMujeeb et al. used deep long and short term memory neural network to predict short-term load and electricity price [49]. The adaptive and automatic feature learning mechanism of deep neural network is used to learn real power data. The simulation results show that the proposed method is effective in forecasting. JaeikJeong et al. used location information and historical power generation data of multiple photovoltaic sites to predict short-term photovoltaic power generation [50]. A greedy adjacency algorithm is proposed to preprocess PV data into spatio-temporal matrix, which is learned by convolutional neural network. A large number of experiments on photovoltaic power generation at multiple sites in three states of the United States show that the proposed spatio-temporal convolutional neural network achieves quite accurate photovoltaic prediction, and the average absolute percentage error is reduced by 0.33 compared with existing methods.

Feature selection and extraction techniques for transformer fault diagnosis are used to identify the most relevant features or variables in the data, which can predict fault conditions. Commonly used feature selection and extraction techniques include principal component analysis (PCA), independent component analysis (ICA) and wavelet analysis. Cao Hao et al. [51] de-noised the acoustic signals collected during normal operation and abnormal transformer of a substation in Beijing. The MFCC features of the signal were extracted by Meir frequency cepstrum coefficient method, and the transformer acoustic signal fault detection was carried out by principal component analysis under unsupervised and semi-supervised learning modes respectively, and the transformer fault monitoring effect of principal component analysis under the two modes was compared. Yang Zeyu [52] proposed a fault detection method for complex systems which combines kernel independent component analysis algorithm and support vector data description algorithm. Firstly, the kernel independent component analysis method is used to extract independent components from process data. Then, the leading independent component is modeled by

introducing support vector data description, and the corresponding statistics and control limits are calculated to detect faults in nonlinear non-Gaussian systems. Simulation experiments are carried out and compared with kernel independent component analysis method and introduced support vector data method. The results show that the proposed method reduces the fault error rate and missing rate, and verifies its feasibility and effectiveness. Cui Dongjun et al. [53] proposed a new class of weighted wavelet functions based on subdivision rather than Fourier transform. Taking the wavelet function as the hidden layer function of the feedforward neural network and optimizing the learning rate of the network, the weighted wavelet neural network is constructed to process the dissolved gas content data in transformer oil.

Support vector machines (SVM) are a supervised learning approach based on statistical learning theory. Introduced in 1995 by professors Cortes and Vapnik from the former Soviet Union, SVM quickly became a mainstream machine learning technique due to its superior performance in classification tasks. Different from traditional learning methods, support vector machines improve the generalization ability of learning machines by seeking the minimum structural risk, minimize the range of empirical risk and confidence, and achieve good statistical rules under the condition of small statistical sample size, which is mainly used for classification and regression problems.

The ratio of H_2, CH_4, C_2H_6, C_2H_4, C_2H_2 and other gases in oil chromatography data (DGA) to total hydrocarbon is used as the evaluation index of fault judgment. Transformer faults are divided into six types: normal, low energy discharge, high energy discharge, partial discharge, medium and low temperature overheating and high temperature overheating [54]. Support vector machine (SVM) is a binary classifier, which can be used to distinguish transformer faults by constructing multi-classification SVM. In order to solve the problem of unequal cost of sample misclassification between classes, a cost sensitive mechanism was introduced to improve the misclassification sensitivity between faults during the construction of classification hyperplane, and a cost sensitive support vector machine with optimized nodes was designed. On the other hand, the data of fault types are not balanced, which will lead to the deviation of the partition superclassification plane when the classifier is trained. Therefore, an artificial intelligent-based optimization algorithm can be combined to optimize the design of misjudgment penalty factors, enhance the role of idling samples in constructing classification hyperplanes, and inhibit hyperplane migration.

Li Yunhao et al. [56] proposed a fault diagnosis method for power transformer based on improved Grey Wolf algorithm coupled with least square support vector machine. In order to improve the accuracy of fault diagnosis, the Grey Wolf algorithm is improved to find the optimal penalty coefficient C and kernel function parameter g in the least squares support vector machine. Feng Zhiliang et al. [57] proposed a Seagull algorithm to optimize support vector machines. Different gas fraction ratio features were added to expand the information features contained in transformer fault data. Principal component analysis (PCA) was used to extract input variable features to reduce the dimension of feature variables and the correlation between variables. Seagull optimization algorithm (SOA) was used to optimize the kernel parameters and penalty factors of support vector machine to improve the accuracy of support vector machine modeling. Compared with particle swarm optimization (PSO) and genetic algorithm (GA), Seagull optimization algorithm optimized support vector machine (SOA-SVM) can significantly improve the accuracy of transformer fault diagnosis, and improve the reliability and generalization performance. Zheng Yesheng et al. [58] proposed a multi-strategy improved Seagull optimization algorithm (ISOA) to optimize SVM transformer fault diagnosis method. Firstly, a multi-strategy improvement method is proposed to improve the optimization performance of SOA in all aspects. Then, the internal parameters of SVM were optimized using ISOA, and the transformer fault diagnosis model based on ISOASVM was constructed. Finally, the feature extraction results of DGA data were input into ISOA-SVM model for transformer fault diagnosis.

Extreme learning Machine is a new single hidden layer feedforward neural network learning algorithm [59].The Extreme Learning Machine algorithm has been applied to different degrees in fault diagnosis, image classification and speech recognition,which has accurate learning ability and good generalization performance. So far, many researches have applied extreme learning machine to power transformer fault diagnosis [60]. According to the relationship between transformer fault and characteristic gas, a suitable sample structure is selected as the input characteristic vector of the network, and its validity is verified. The results of various activation functions are studied in the power transformer fault diagnosis method based on extreme Learning Machine, and the influence of the number of hidden layer nodes on the fault diagnosis performance is studied [55].

In order to improve the accuracy of transformer fault diagnosis, Wang Rui et al. [61] adopted Ant-lion optimization algorithm to optimize the weights and thresholds of extreme learning machine and established a transformer fault diagnosis model based on Ant-lion algorithm to optimize extreme learning machine. The training sample set data is used for training, and the ELM network structure is determined according to the training error. The trained ALO-ELM model is used for fault diagnosis of the training sample set data, and the comparison is made with other fault diagnosis methods.Using Tent chaotic mapping and Chi-square probability density function, the original Tianying optimization algorithm (AO) is improved, and the improved algorithm effectively improves the convergence speed and optimization accuracy [62].Wang Chunming et al. [63] proposed a transformer fault diagnosis method for deep noise reduction extreme learning machine, aiming at the problems of long training time, easy overfitting and noise sensitivity in transformer fault diagnosis. Combining the extreme learning machine and the noise reduction autoencoder, the self-encoding extreme learning machine was constructed, and the deep noise reduction extreme learning machine model was stacked to carry out feature extraction. The extracted features were input into the conventional extreme learning machine for classification, and the deep noise reduction extreme learning machine classification algorithm was constituted as a whole.

To sum up, machine learning technology has achieved considerable success in transformer fault diagnosis [55], which can be applied to transformer fault classification, prediction and diagnosis. Fault diagnosis methods based on machine learning can use feature selection and extraction techniques to improve the performance and interpretability of the model. With the development of data acquisition and processing technology, the role of machine learning technology in transformer fault diagnosis will be enhanced.

5. Summary

As the "lifeblood" of power system, transformer undertakes irreplaceable work. Its stable operation is an important premise for the smooth operation of the whole power system. Once the failure occurs, it will not only bring a huge impact on social residents, but also affect the development of the country and the improvement of social economy. It can be seen that how to quickly discover and solve the transformer fault and how to accurately predict the transformer fault is particularly important in the power system development industry [64].

At present, there are three future development trends of transformer fault diagnosis:

5.1. Improve alarm data information

No matter which method is used to carry out the fault diagnosis of power system, the next operation must be carried out according to the alarm information inside the system. However, the alarm information provided by the power system itself is usually missing or even wrong, which is not conducive to the development of normal diagnosis. Therefore, rational use of information system theory and information entropy may significantly improve the accuracy and reliability of transformer fault diagnosis. The future developing trends of power system fault diagnosis is making fully use of information system theory and information entropy and other concepts to constantly improve the alarm data information

provided by itself, and strive to achieve high quality fault diagnosis under the condition of imperfect information. The latter requires the staff to carry out the fault diagnosis work, first of all, under the premise of imperfect information to carry out hypothesis research, and then effectively combine a variety of technical methods to make up for the loopholes brought by imperfect information, and then complete the high-quality diagnosis work [65].

5.2. Adopt multiple machine learning intelligence methods

In the process of power system evolving to smart grid and comprehensive energy, it has become a giant dimensional system with complex structure, various equipment and complicated technology. Traditional methods relying solely on physical modeling analysis have been difficult to meet the operational analysis requirements of multi-physical field coupling systems [66]. Machine learning relies less on specific mathematical models and is good at self-learning from data, which provides an effective solution to overcome the above technical difficulties. With the large number of power grid operation and monitoring equipment put into operation, the continuous integration of monitoring system and information system, the trend of sea quantification of power data will be more and more obvious. Collecting massive power data and using machine learning technology to mine its hidden value is the inevitable choice of intelligent and modernized power system. Therefore, machine learning will be very useful in the field of energy and power system in the future, and it is an inevitable demand in the era of informatization and big data.

5.3. Promote the practical research of the theory

At present, there are a large number and a wide range of researches in the field of power system fault diagnosis, but most of them only stay in the theoretical level, lacking of practical research [65]. At the same time, since the fault diagnosis of power system itself has strong practical characteristics, it is also one of the future development trends of power system fault diagnosis to promote the theoretical research in this field and turn to practical research. All the technologies applied in power system fault diagnosis are closely related to modern information technologies such as information system theory, information entropy and machine learning integration. Therefore, promoting the practical research on information system theory, information entropy and machine learning is conducive to the high-quality prosperity of power system fault diagnosis. It will also play an important role in promoting the whole intelligent process of modern society.

References

1. Lin Jingcha. Research Status of Transformer Fault Detection and Diagnosis Technology. *Electrical Switches* **2021**, *59(03)*, 101-103.
2. Liu Chunpeng, Wang Xiaopan. Research Status of Transformer Fault Detection and Diagnosis Technology. *Electrical Switches* **2020**, *(17)*, 152-153.
3. Hao Lingling, Zhu Yongli, Wang Yongzheng. Transformer Fault Diagnosis Method Based on DCAE KSSELM . *China Electric Power* **2022**, *55(2)*, 125-130.
4. Li Feng, Chen Wanwan, Li Xiaohua, et al. An intelligent fault diagnosis method for transformer in novel power system based on SVMD-CMSEE and GSA-SVM. *Electrical Measurement and Instrumentation* **2022**, 1-10.
5. Qin Peng, Zhao Feng. LS-SVM probability output diagnosis of power transformer faults using Bayesian framework. *China Electric Power* **2015**, *48(5)*, 41-45.
6. Song Hui, Yuan Longxiang, Guo Shuangquan. Grey Wolf Optimization Algorithm Based on Data Enhancement and Feature Attention Mechanism-Optimization Residual Neural Network Transformer Fault Diagnosis Method. *Modern Electric Power* **2023**,1-9.
7. Dai X W,Gao Z W.From model,signal to knowledge:A data-driven perspective of fault detection and diagnosis. *IEEE Transactions on Industrial Informatics* **2013**, *9(4)*, 2226-2238.
8. Stetco A,Dinmohammadi F,Zhao X,et al.Machine learning methods for wind turbine condition monitoring:A review. *Renewable Energy* **2019**, *133(1)*, 620-635.
9. Wu Wenying, Yi Meiling, Wen Jun, et al. Analysis and Judgment of Dissolved Gas in Transformer Oil. *Energy Research and Management* **2019**, *(4)*, 57-60.
10. Song Zhijie, Wang Jian. Transformer Fault Diagnosis Based on BP Neural Network Improved by Fuzzy Clustering and LM Algorithm. *High Voltage Appliances* **2013**, *49(05)*,54-59.
11. Shao Huidong. Transformer Fault Diagnosis Based on Machine Learning. *Guangdong University of Technology* **2020**.

12. Wang Wei, Tang Qinghua, Liu Liqing, et al. Transformer Fault Diagnosis Based on Weighted Comprehensive Loss Optimization Deep Learning and DGA. *China Southern Power Grid Technology* **2020**, *14(3)*, 2934.

13. Seifeddine S,Khmais B,Abdelkader C.Power transformer fault diagnosis based on dissolved gas analysis by artificial neural network. *First International Conference on Renewable Energies and Vehicular Technology* **2012**, *(3)*,230-236.

14. Siddique M A A,Mehfuz S.Artificial networks based incipient fault diagnosis for power transformers. *India Conference* **2016**,1-6.

15. Adewole A C,Tzoneva R.Distribution network fault detection and diagnosis using wavelet energy spectrum entropy and neural networks. *International Review of Electrical Engineering* **2014**, *9(1)*,165-173.

16. Adewole A C,Tzoneva R,Behardien S.Distribution network fault section identification and fault location using wavelet entropy and neural networks. *Applied Soft Computing* **2016**, *46(C)*,296-306.

17. Aslan Y,Yagan Y E.Artificial neural-network-based fault location for power distribution lines using the frequency spectra of fault data. *Electrical Engineering* **2016**, *199*,1-11.

18. Yan Z,Zhang B,Yuan Y, et al.Transformer fault prediction based on support vector machine. *International Conference on Computer Engineering and Technology* **2010**.

19. Aiapurua J I,Catterson V M,Stewart B G,et al,Power transformer dissolved gas analysis through Bayesian networks and hypothesis testing. *IEEE Transactions on Dielectrics and Electrical Insulation* **2018**, *25(2)*,494-506.

20. Wei C,Tao F B,Lin Y D,et al.Substation equipment thermal fault diagnosis model based on ResNet and improved Bayesian optimization. *2019 9th International Conference on Power and Energy Systems* **2019**,1-5.

21. Su H S,A novel substation fault diagnosis and maintenance method based on ontology and Bayesian network. *International Conference on Complex Systems and Applications* **2006**.

22. Wang D W,Sun Z W.Parallelization analysis of dissolved gases in transformer oil based on random forest algorithm. *Applied Mechanics and Materials* **2014** *519/520*,98-101.

23. Li Wannian,Xiao Jinfeng. Research on Transformer Fault Diagnosis Model Based on Improved GOA-SVM. *Journal of Physics: Conference Series* **2023** *2465(1)*.

24. Demirci Merve,Gözde Haluk,Taplamacioglu M. Cengiz. Improvement of power transformer fault diagnosis by using sequential Kalman filter sensor fusion. *International Journal of Electrical Power and Energy Systems* **2023**,149.

25. Hu Feng,Hu Jian,Dai Rongying,Guan Yuqi,Shen Xianfeng,Gao Bo,Wang Kun,Liu Yu,Yao Xiaokang. Selection of characteristic wavelengths using SMA for laser induced fluorescence spectroscopy of power transformer oil. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy* **2023**, 288.

26. Liu Rongsheng,Cui Shuguang,Lin Anping,Liao Yong. Transformer Fault Synthetic Diagnosis Method Based on Fusion of Multi-Neural Networks and Evidence Theory in Cloud Computing. *Journal of Physics: Conference Series* **2023**, *2433(1)*.

27. Li Xu, Xu Xinhe. A Fault Diagnosis Expert System Based on Information Theory. *Northeastern University Press* **1997**,532-534.

28. Yu Hong. Application of Information Entropy in Transformer Fault Diagnosis and State Detection. *North China Electric Power University* **2014**.

29. Ma Hongbin, Yang Fei, Yang Rena, Zhang Ran. Application of Machine Learning in Transformer Fault Diagnosis. *Hebei Electric Power Technology* **2021**, *40(04)*,35-38.

30. Liu Yunpeng, Xu Ziqiang, Li Gang, et al. Application of Data Analysis Technology driven by Artificial Intelligence in Power transformer Condition maintenance. *High Voltage Technology* **2019**, *45*,337-348.

31. Hansen L K, Salamon P. Neural network ensembles. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **2002**, *12(10)*,993-1001.

32. Hai Chen. Research on Transformer Fault Diagnosis Method Based on Machine Learning. *North China University of Water Resources and Electric Power* **2022**.

33. Shen Xuehua, Zhou Zhihua, Wu Jianxin, et al. Boosting and Bagging. *Computer Engineering and Applications* **2000**, *36*, 31-32.

34. Breiman L. Bagging predictors. *Machine Learning* **1996**, *24(2)*, 123-140.

35. Hinton G E, Osindero S, Teh Y W. A fast learning algorithm for deep belief nets. *Neural Comput* **2006**, *18(7)*, 1527-1554.

36. Yin Baocai, Wang Wentong, Wang Lichun. A Review of Deep Learning Research. *Journal of Beijing University of Technology* **2015**, *41*, 48-59.

37. Ji Shouling, Li Jinfeng, Du Tianyu, et al. Research on Interpretability Methods, Applications and Security of Machine Learning Models. *Computer Research and Development* **2019**, *56*, 2071-2096.

38. Taigman Y, Ming Y, Ranzato M, etc. DeepFace: Closing the Gap to Human-Level Performance in Face Verification*IEEE Conference on Computer Vision and Pattern Recognition* **2014**.

39. Dastgiri A, Jafarinamin P, Kamarbaste S, etc. Face Recognition using Deep Neural Networks. *Journal of Mechanics of Continua and Mathematical Sciences* **2019**, *14(3)*,510-527.

40. Schroff F, Kalenichenko D, Philbin J. FaceNet: A unified embedding for face recognition and clustering. *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* **2015**, 815-823.

41. Geiger A, Lenz P, Urtasun R. Are we ready for autonomous driving? The KITTI vision benchmark suite . *2012 IEEE Conference on Computer Vision and Pattern Recognition* **2012**,3354-3361.

42. Tobiyama S, Yamaguchi Y, Shimada H, etc. Malware Detection with Deep Neural Network Using Process Behavior. *2016 IEEE 40th Annual Computer Software and Applications Conference* **2016**, 577-582.

43. Habib N, Hasan M M, Reza M M, etc. Ensemble of CheXNet and VGG-19 Feature Extractor with Random Forest Classifier for Pediatric Pneumonia Detection. *SN Computer Science* **2020**, *1(6)*, 359.

44. Deng Xinyu, Wang Shouxiang, Guo Luyang. Application of Artificial Intelligence Method in Power Distribution Field. *Power Supply* **2019**, *36(01)*, 3-9.

45. Ai Qian. Research and Application of Artificial Intelligence in Power distribution field. *Power Supply* **2019**, *36(01)*, 1-2.

46. Wang L, Zhou Q, Jin S. Physics-guided Deep Learning for Power System State Estimation. *Journal of Modern Power Systems and Clean Energy* **2020**, *8(4)*, 607-615.

47. Fei W, Moses P. Fault Current Tracing and Identification via Machine Learning Considering Distributed Energy Resources in Distribution Networks. *Energies* **2019**, *12(22)*, 4333.

48. Cepeda C, Orozco-Henao C, Percybrooks W, etc. Intelligent Fault Detection System for Microgrids. *Energies* **2020**, *13(5)*, 1223.

49. Mujeeb S, Javaid N, Ilahi M, etc. Deep Long Short-Term Memory: A New Price and Load Forecasting Scheme for Big Data in Smart Cities. *Sustainability* **2019**, *11(4)*, 987.

50. Jeong J, Kim H. Multi-Site Photovoltaic Forecasting Exploiting Space-Time Convolutional Neural Network. *Energies* **2019**, *12(23)*, 4490.

51. Cao Hao, Huang Tao, Zhou Zhou, Xie Jie. Acoustic Anomaly Detection of Transformer DC Magnetic Bias Based on Principal Component Analysis. *Hunan Electric Power* **2021**, *41(05)*, 1-6.

52. Yang Zeyu. Research on Fault Detection and Classification of Complex Systems Based on Kernel Independent Component Analysis. *Hangzhou Dianzi University*, **2017**.

53. Cui Dongjun, Liu Nian, Liu Xiulan. Fault Detection of Oil-immersed Power Transformer Based on Weighted Wavelet Neural Network. *Power System Protection and Control* **2010**, *38(18)*, 19-23.

54. Zhang Yong. On-line Monitoring and Fault Diagnosis of Transformer Based on Dissolved Gas Analysis in Oil. *North China Electric Power University* **2014**.

55. Liu Linfan. Research Progress of Power Transformer Fault Diagnosis Based on Machine Learning. *Electronic World* **2017**, *(15)*, 9-10.

56. Li Yun-hao, XIAN Min-Daily, ZHANG Hai-Qiang, ZHAO Fei-Long, LI Jia-Yang, WANG Wei, LI Zeng-Yue. Power Transformer Fault Diagnosis Method Based on Improved Grey Wolf Algorithm Coupled with Least Squares Support Vector Machine. *Power Grid Technology* **2023**, *47(04)*, 1470-1478.

57. Feng Zhiliang, Xiao Hanqi, Ren Wenfeng, Du Yanli. Fault Diagnosis of Transformer Based on Seagull Optimization Support Vector Machine Based on Principal Component Analysis. *China Test* **2023**, *49(02)*, 99-105.

58. Zheng Yeshuang, Li Shichun, Lu Ling. Research on Transformer Fault Diagnosis Based on Multi-Strategy ISOA Optimization SVM. *Intelligent Power* **2023**, *51(02)*, 38-44.

59. Li Donghui, Yan Zhenlin, Yao Lele, et al. Short-term Power Load Prediction Based on Improved Manifold Regularization Extreme Learning Machine. *High Voltage Technology* **2016**, *42(7)*, 2092-2099.

60. Yuan Haiman, Wu Guangning, Gao Bo. Fault Diagnosis of Power Transformer by Particle Swarm Extreme Learning Machine Based on DGA. *High Voltage Appliances* **2016**, *(11)*, 176-180.

61. Wang Rui, You Jingyao, ZHANG Xiaobin, Li Yong, Xu Zhifeng, Liu Chuang. Transformer Fault Diagnosis Based on Ant-Lion Algorithm Optimization of Extreme Learning Machine. *Heilongjiang Electric Power* **2023**, *45(01)*, 17-22.

62. Yi Lingzhi, LONG Jiao, WANG Yahui, HUANG Jianxiong, Sun Tao, Yu Huang, Yu Guo. Research on Oil Immersed Transformer Fault Diagnosis Based on Improved Tianying Algorithm Optimization of Nuclear Extreme Learning Machine. *Hunan Electric Power* **2022**, *42(06)*, 7-14.

63. Wang Chunming, Zhu Yongli. Transformer Fault Diagnosis Based on Deep Noise Reduction Extreme Learning Machine. *Electrical Measurement and Instrumentation* **2019**, *56(15)*, 143-147.

64. Feng Baiqing. Intelligent Diagnosis and Fault Prediction of Large Transformer. *Shenyang University of Technology* **2022**.

65. Ouyang Shenrui. Research Status and Development Trend of Power Syst3em Fault Diagnosis. *Electric Drive Automation* **2021**, *43(04)*, 40-42+64.

66. Yang Ting, Zhao Liyuan, Wang Chengshan. Application of Artificial Intelligence in Power system and integrated energy system. *Power System Automation* **2019**, *43(01)*, 2-4.