Engineering Applications of Artificial Intelligence 127 (2024) 107357

Contents lists available at ScienceDirect

Engineering Applications of Artificial Intelligence

journal homepage: www.elsevier.com/locate/engappai

A literature review of fault diagnosis based on ensemble learning

Zhibao Mian [a,b,*], Xiaofei Deng [a], Xiaohui Dong [a], Yuzhu Tian [c], Tianya Cao [a], Kairan Chen [a], Tareq Al Jaber [b]

[a] School of Computer Science and Engineering, Northwest Normal University, Lanzhou, China
[b] School of Computer Science, FoSE, The University of Hull, Hull, UK
[c] School of Mathematics and Statistics, Northwest Normal University, Lanzhou, China

ARTICLE INFO

Keywords:
Ensemble learning
Fault diagnosis
Intelligent maintenance
System reliability

ABSTRACT

The accuracy of fault diagnosis is an important indicator to ensure the reliability of key equipment systems. Ensemble learning integrates different weak learning methods to obtain stronger learning and has achieved remarkable results in the field of fault diagnosis. This paper reviews the recent research on ensemble learning from both technical and field application perspectives. The paper summarizes 87 journals in recent web of science and other academic resources, with a total of 209 papers. It summarizes 78 different ensemble learning based fault diagnosis methods, involving 18 public datasets and more than 20 different equipment systems. In detail, the paper summarizes the accuracy rates, fault classification types, fault datasets, used data signals, learners (traditional machine learning or deep learning-based learners), ensemble learning methods (bagging, boosting, stacking and other ensemble models) of these fault diagnosis models. The paper uses accuracy of fault diagnosis as the main evaluation metrics supplemented by generalization and imbalanced data processing ability to evaluate the performance of those ensemble learning methods. The discussion and evaluation of these methods lead to valuable research references in identifying and developing appropriate intelligent fault diagnosis models for various equipment. This paper also discusses and explores the technical challenges, lessons learned from the review and future development directions in the field of ensemble learning based fault diagnosis and intelligent maintenance.

1. Introduction

Fault diagnosis mainly studies the detection, separation, and identification of system faults, including identifying whether and when a fault occurs, and fault location. It is a process of tracking faults based on fault symptoms, fault knowledge and analysis of test results. Fault diagnosis can be divided into analytical model-based and data-based diagnosis methods. Analytical model-based methods generally use accurate mathematical models of the system and observable input and output signal processing to construct residual signals. The residual signal can reflect the inconsistency between system expectations and actual conditions and can be used for fault diagnosis. Analytical model-based methods are highly dependent on an accurate mathematical model of the system being diagnosed. In practice, it is difficult to establish an accurate mathematical model of the system, especially for complex systems. In this case, the model-based method is no longer applicable. However, with the development of information technology, a large amount of system operation data can be stored and analyzed, which has resulted in data-based fault diagnosis methods. Data-based methods use artificial intelligence models to analyse system operation process data, so that fault diagnosis can be completed without knowing the precise analytical model of the system.

Currently, the utilization of artificial intelligence technology such as intelligent fault diagnosis (Lei et al., 2016) for critical equipment to assist next generation intelligent maintenance has become more and more promising (Shao et al., 2015). Intelligent maintenance is a system utilizing intelligent data analysis and decision-making algorithms or tools to predict and prevent the potential failure of machines. Model-driven engineering, especially the model-based dependability analysis methods (Sharvia et al., 2016) could benefit the design and development of intelligent maintenance in terms of automation and consistency. In our early research work, we combined model-based

* Corresponding author. School of Computer Science, FoSE, The University of Hull, Hull, UK.
E-mail addresses: z.mian2@hull.ac.uk (Z. Mian), 2021212127@nwnu.edu.cn (X. Deng), dxh@nwnu.edu.cn (X. Dong), pole1999@163.com (Y. Tian), caotianya1979@126.com (T. Cao), 2021222187@nwnu.edu.cn (K. Chen), T.Al-Jaber@hull.ac.uk (T.A. Jaber).

https://doi.org/10.1016/j.engappai.2023.107357
Received 29 March 2023; Received in revised form 11 October 2023; Accepted 23 October 2023
Available online 10 November 2023

dependability analysis methods with the traditional maintenance technologies and proposed a model-based system intelligent maintenance framework (Mian et al., 2020). This framework combines traditional maintenance theories such as reliability-centered maintenance (RCM) (Moubray, 2001) with model-based system dependability analysis techniques to achieve a model-based system maintenance. To enhance the analytical performance of the framework, this paper introduces a number of AI-based maintenance methods, especially intelligent fault diagnosis. The paper also explores a new intelligent maintenance model.

Intelligent fault diagnosis is an important component of an intelligent maintenance system. It helps to provide more accurate maintenance plans for equipment maintenance by monitoring online system operation states and analyzing various real time system data. There have been machine learning methods such as k-nearest neighbors (KNN) (Chen et al., 2020), support vector machine (SVM) (Zhong et al., 2020), convolutional neural network (CNN) (Xiong et al., 2015), extreme learning machine (ELM) (Shi and Zhang, 2020) and decision tree (Peng et al., 2020) used for equipment fault diagnosis. These methods suffer problems, however, including incomplete feature extraction (missing features and long model training and validation times when there is a large amount of training data or a dataset with many data features). The core issues of intelligent fault diagnosis are the speed of training, the accuracy and timeliness of fault diagnosis. Training speed may be affected by the size of the dataset used and the performance of selected learning methods based on that dataset. The larger the dataset size the more training time is required. Models that can use small datasets or perform well on small datasets are highly anticipated. Furthermore, Wen and Lv (2020) argued that single fault diagnosis technology has problems such as low diagnostic accuracy and weak generalization ability. Ensemble learning methods (Zhou, 2021a) tend to work better on smaller dataset (Xu and Yang, 2018) and have the potential to perform higher diagnostic accuracy and generalization capabilities. This motivates our research on ensemble learning.

Ensemble learning is a machine learning method that aims to capture the wisdom of a group decision. Ensemble learning integrates multiple learners, with the aim of compensating for the disadvantages of any single learner. Through this ensemble, it forms an ensemble model with a higher accuracy than any single learner (Zhou, 2021a).

Ensemble learning have been widely used in various industries. Fig. 1 shows the number of papers applied to ensemble learning in various fields classified by Web of Science. In Fig. 1, ensemble learning is more used in Computer Science and Natural Science disciplines (e.g., Biology, Chemistry and Physics). But it is interested to see that it has been used in various fields. Ansari et al. (2023) compared the depression detection of an ensemble learning model based on bagging with that of a hybrid model and demonstrates that the ensemble model outperformed the hybrid model for classification results. Mehta et al. (2022) applied bagging, pressurization and stacking methods to combine weak models to detect various types of cyber-attacks and achieve better security mechanisms. The results show that higher accuracy can be obtained by combining models using ensemble learning algorithms. Fang et al. (2023) proposed a knowledge-enriched ensemble method to combine information from both knowledge graphs and pre-trained word embeddings. They demonstrated that the proposed word embeddings outperform the state-of-the-art models in word analogy, word similarity and several downstream tasks. Xue et al. (2022) designed two new Gaussian process regression (GPR)-based ensemble learning models for soil moisture estimation, the bagging embedded with GPR (BAGGPR), and the gradient boosting (BOOST) embedded with GPR (GBGPR). The results demonstrate that the proposed ensemble learning methods outperform other state-of-the-art models. Kamble and Sengupta (2022) compared five ensemble learning-based machine learning (EML) algorithms (bagging, random forest, rotation random forest, extreme gradient boost, and adaptive boosting) with five conventional machine learning (CML) algorithms for recognizing multiple human emotions from electroencephalogram (EEG) signals. The results show that the mean accuracy of multiclass emotion recognition over five EML algorithms is higher than the mean accuracy of five CML algorithms, for both arousal (88.95% vs. 83.08%) and valence (88.90% vs. 82.81%) dimensions, respectively. Janet et al. (2022) evaluated educational datasets using various classifier models as well as ensemble learning models. The results show that the ensemble algorithms perform better than the base algorithms when making student grade predictions. To achieve higher accuracy and effectiveness in predicting the price of virtual currency, Dong (2022) proposed a segmented integrated learning (ensemble-SVR) method based on SVR algorithm. For some other applications, for example, Alotaibi (2020) proposed a stacking learning technique to detect the mode of transportation. The proposed method outperformed the other methods in terms of accuracy on the three datasets with two different data splits.

Ensemble learning has also been applied in the field of equipment fault diagnosis and has achieved remarkable diagnostic results. In this paper we review the literature on equipment fault diagnosis based on ensemble learning. Furthermore, we argue that the theory and practice of ensemble learning based fault diagnosis methods should be introduced into the field of model-based maintenance. This will lead to intelligent system maintenance and improve the dependability of the system.

Fig. 2 shows the framework of the fault diagnosis methods based on a single learning model (a) and an ensemble learning model (b). The

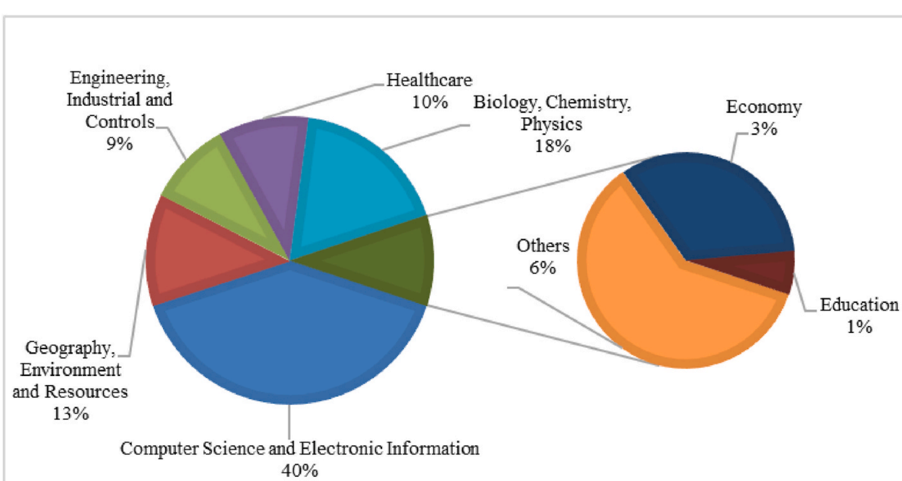


Fig. 1. Applications scenarios of ensemble learning in various fields.

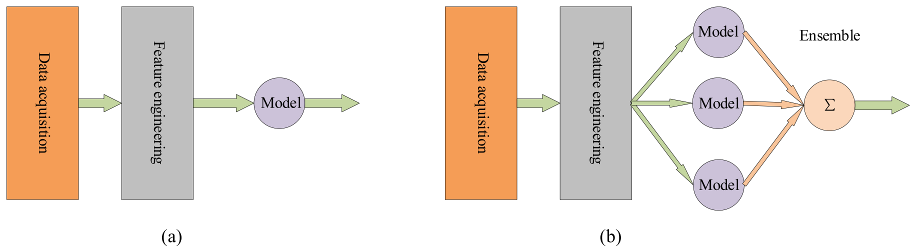


Fig. 2. The basic framework of fault diagnosis methods based on (a) single learning model; (b) ensemble learning model.

framework of ensemble learning based fault diagnosis methods are similar to those of the single learner-based methods. The main difference is that the ensemble learning model is a collective intelligence expression of multiple single learning models. The result of the fault diagnosis model based on ensemble learning is determined by the result of multiple trained single models, which is calculated by the ensemble strategies (e.g., bagging, boosting and stacking algorithms).

In detail, Fig. 3 shows the fault diagnosis framework based on ensemble learning. The first step is to collect relevant fault data and perform feature processing on the data, i.e., feature engineering. Feature engineering can be divided into traditional feature extraction methods, neural networks and deep learning methods. The traditional feature extraction method is very helpful to reduce the complexity of the model and improve the accuracy of the model, but it is relatively dependent on professional knowledge and prior information (Rong, 2021). Neural networks and deep learning methods are currently the mainstream research methods (Tang et al., 2022). They greatly remove the dependence on prior knowledge. Feature engineering is followed by the process of selecting proper leaners and ensemble strategies. An ensemble model is generated by integrating different ensemble strategies (e.g., bagging, boosting, stacking, and blending). Finally, the classification results of fault diagnosis are output through the assemble model.

This paper first surveys the literature review on fault diagnosis in recent years. Most of it summarizes the fault diagnosis research from the perspectives of technology or application. Wen and Lv (2020) reviewed fault diagnosis methods based on deep learning, including stacked self-coding, deep confidence network, convolutional neural network and recurrent neural network. Cai et al. (2017) reviewed the research literature on Bayesian networks and their optimization methods in engineering systems. Cen et al. (2022) reviewed the application of shallow machine learning (SML), deep learning (DL) and transfer learning (TL) in fault diagnosis. They pointed out that TL's diagnostic model can realize knowledge transfer across conditions, machines and even domains to solve the problems of data scarcity and sample imbalance that often occur in fault diagnosis. These reviews provide extensive and detailed discussions of the technical theories, research status, technical challenges and problems of various intelligent methods from a technical perspective. The reviews, however, lack a comparison between technologies, as well as the application situation.

Compared with a technology-based overview, a system equipment-based overview focuses on different intelligent diagnosis methods used for a specific critical system. This allows the researchers of the same

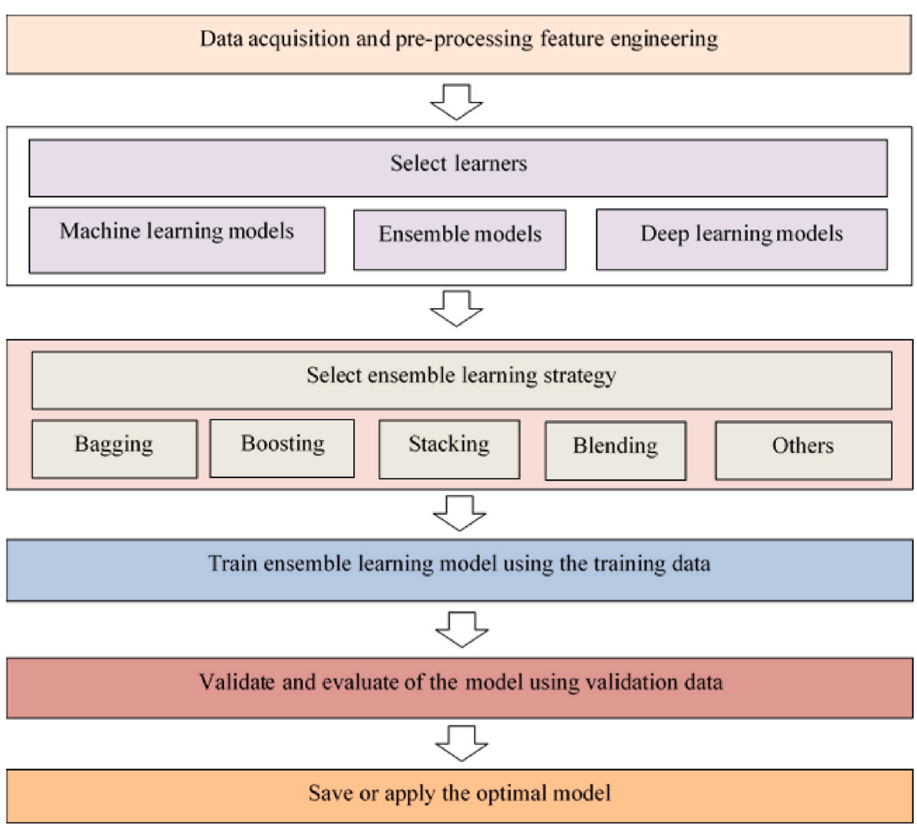


Fig. 3. The fault diagnosis framework based on ensemble learning.

equipment system to have a deep understanding of the fault diagnosis research in this field. Ding et al. (2022) reviewed the fault diagnosis techniques applied to key components of cold ironing equipment. They classified the literature according to the type of key components, and classified and analyzed the fault diagnosis methods applied to each type of components. Chen et al. (2019) summarized fault diagnosis methods for permanent magnet synchronous motors from three perspectives: model-based fault diagnosis, different signal processing methods and data-driven diagnostic algorithms. They also pointed out that ensemble learning is a promising method for improving fault diagnosis accuracy. Naveen and Sugumaran (2021) reviewed fault diagnosis methods for photovoltaic module. They targeted problems where traditional methods were labor-intensive and not cost-effective. They suggested the use of hybrid fault diagnosis techniques to obtain fast and efficient detection and classification of multiple faults. Zang et al. (2019) reviewed the fault diagnosis methods for high-speed railways. They conducted in-depth discussions on the fault diagnosis of high-speed railway turnouts from four technical perspectives: expert system-based model, fault tree-based model, Bayesian network-based model and neural network-based model. According to the structural complexity of high-speed railways, they proposed that each diagnostic method can be combined according to its merits to be applied to the target system.

To sum up, some of these reviews focus on fault diagnosis technology, and some focus on the application of fault diagnosis, but these reviews all have one thing in common, i.e., they all regard integrating or mixing various diagnostic methods as a future research direction. Exploring fault diagnosis methods based on ensemble learning is potentially a valuable work. In addition, there are relatively few reviews on ensemble learning for fault diagnosis. This paper reviews fault diagnosis based on ensemble learning.

This paper differs from other reviews in the same field is that: first, the diagnostic techniques discussed are different; this paper mainly focuses on ensemble learning methods. Secondly, this paper discusses ensemble learning based fault diagnosis methods from both application and technology perspectives. Finally, this paper has the purpose of identifying the general best ensemble learning based fault diagnosis models in a given application area so that they may be applied in other areas. Table 1 below summarizes the main differences between this paper and the other similar work.

To facilitate readers to understand the journals and conferences related to ensemble learning-based fault diagnosis research, this paper summarizes these journals and conferences. Fig. 4 shows the distribution of relevant journals and conferences; and the larger the rectangular area, the greater the number of references summarized from relevant journals or conferences.

The main contributions of this article are as follows.

1. This paper reviews ensemble learning based fault diagnosis methods from both application and technology perspectives. The paper summarizes the experiences and lessons learned from the reviewed literature.
2. From an application perspective, this paper reviews the ensemble learning based fault diagnosis methods and their application for some critical equipment including bearings, gears, transformers, new energy systems and some other key system equipment. The performance of applying those ensemble learning based fault diagnosis

Table 1
Differences between this paper and other similar literature.

	This paper	Other similar studies
Cross-discussion on applications and technologies.	Discusses the current state and challenges of fault diagnosis methods based on ensemble learning from both application and technical perspectives. It provides a comprehensive discussion and overview of the models used, model performance, datasets, equipment types, fault types, and fault data signals.	References (Wen and Lv, 2020), (Cai et al., 2017), and (Cen et al., 2022) provide a technical review of fault diagnosis techniques. Literatures (Ding et al., 2022), (Chen et al., 2019), (Naveen and Sugumaran, 2021), and (Zang et al., 2019) discuss the research status of various fault devices from an application perspective. These literatures mainly outline the status of fault diagnosis from a technical or application perspective, or focus on systems or models in different fields, with relatively limited analysis of model performance and fault types.
Target model and application domain.	Primarily provides an overview of ensemble learning methods used for fault diagnosis. Furthermore, it highlights on discussing and summarizing key technical challenges and issues in this context.	Reference (Xu and Yang, 2018) discusses ensemble learning from two aspects: ensemble strategies and application scenarios. However, its applications are mainly targeted at three fields, time series analysis, healthcare, and intrusion detection, rather than fault diagnosis. References (Wen and Lv, 2020), (Cai et al., 2017), and (Cen et al., 2022) review fault diagnosis technologies based on deep learning, Bayesian network, and machine learning respectively.
Performance measurement.	Uses accuracy of fault diagnosis, generalization, and imbalanced data processing ability to comprehensively evaluate the performance of reviewed ensemble learning methods.	References (Wen and Lv, 2020), (Xu and Yang, 2018), (Cai et al., 2017), (Chen et al., 2019), (Naveen and Sugumaran, 2021), (Cen et al., 2022), and (Ding et al., 2022) provide in-depth discussion on the challenges and opportunities faced by existing fault diagnosis methods from a technical or application perspective. However, relatively little attention has been paid to the discussion of performance metrics and comparative analyses between the performance of different methods.
Dataset discussion.	Summarizes the datasets used in the literature, both public and private. This paper especially organizes and discusses the data types and fault types in the reviewed literature in detail.	Reference (Wen and Lv, 2020) emphasizes the importance of data in data-driven fault diagnosis. It discusses the problems existing in data collection. However, it has less discussion on the specific data used by different methods, especially the fault types and data collection methods. Reference (Chen et al., 2019) discusses the fault types of permanent magnet synchronous motors and conducts a detailed analysis of its fault information, but their work only focused on permanent magnet synchronous motors.
Model-based reliability and intelligent maintenance discussion.	Explores the possibility of selecting a general model and dynamic matching schemes suitable for different equipment by analysing the performance of different methods for different faults.	Reference (Xu and Yang, 2018) conducts a comprehensive review of existing ensemble learning methods from the perspectives of ensemble strategies and the application of ensemble learning solutions. It also discusses the limitations of current ensemble learning algorithms.
Featured discussions include low-energy, data security and social benefits.	In addition to focusing on fault diagnosis technology itself, this paper also discusses intelligent maintenance in closely related fields such as society, economy, low-energy consumption, data security and encryption.	Reference (Zhang and Li, 2022) discusses federated transfer learning and its application in mechanical fault diagnosis, with special emphasis on its ability to provide privacy protection for data from different source domains.

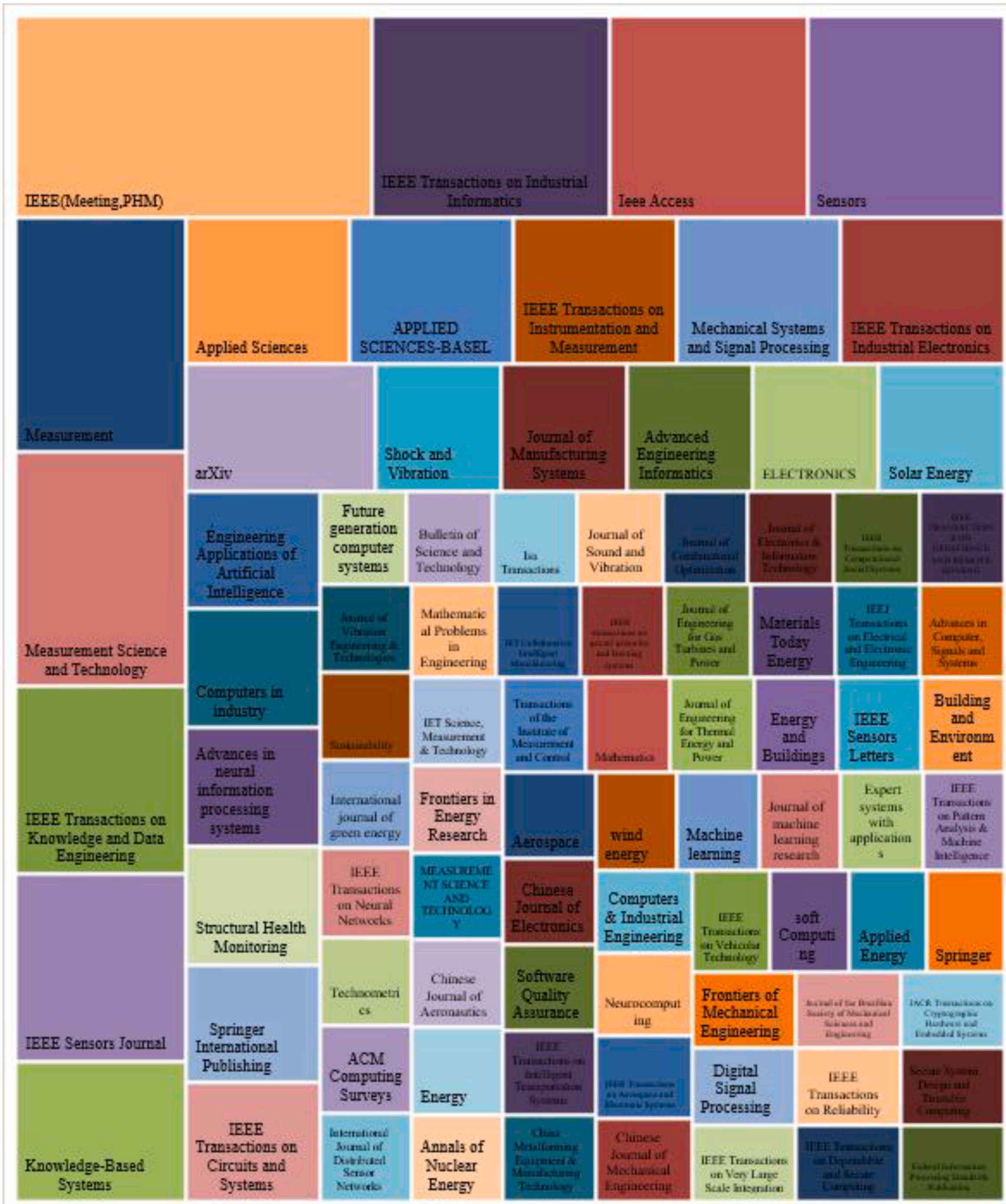


Fig. 4. The distribution of relevant journals and conferences; and the larger the rectangular area, the greater the number of references summarized from relevant journals or conferences.

methods is generally evaluated and summarized based on key dimensions including the accuracy of fault diagnosis, model methods, fault types, generalization of models and imbalance data processing capabilities.

3. From a technical perspective, this paper discusses the challenges and problems faced by fault diagnosis and proposes some future research directions, including: The influence and challenges of cross-system models on fault diagnosis research; The prospect of ensemble learning methods in the research of equipment remaining life prediction; The future development of the combination of deep learning and ensemble learning; The paper also discusses at the data level some of the factors that negatively impact fault diagnosis accuracy and fault model performance, including unbalanced data, concept drift, and fault data under speed variation.
4. This paper conducts a detailed discussion on AI and security and low-power encryption technologies that are highly related to intelligent fault diagnosis from special perspectives such as society, economy, security, and low energy consumption. These discussions enrich the ecological environment of fault diagnosis.
5. Fault data arrangement is one of the core tasks of fault diagnosis research. This paper sorts out some influential and core equipment fault data datasets in the field of fault diagnosis, including their

descriptions and download addresses, so that interested researchers can refer to them.

The rest of the paper is organized as follows. Section 2 summarizes different ensemble methods according to different equipment. Section 3 analysis the fault diagnosis data, method performance and ensemble technology based on ensemble learning. Section 4 summarizes some of challenges and directions. Section 5 summarizes lessons learned from the surveyed papers. Section 6 presents our conclusion and future work.

2. Fault diagnosis based on ensemble learning

Empirically, ensemble learning methods tend to yield better results when there is a significant diversity among the individual methods (Kuncheva and Whitaker, 2003). It might appear attractive to classify and compare the ensemble learning based fault diagnosis methods in terms of types of ensembles (Wikipedia contributors, 2023). Theoretically, the performance can be compared and evaluated based on the types of ensemble learning. However, due to the use of different application scenarios and datasets, it is difficult to make a fair and objective evaluation.

Engineers are more willing to focus on the application of ensemble learning. From the perspective of application, different enterprises have different target equipment with different operating environments and fault types. Therefore, they pay more attention to the diagnosis of similar target devices or failure types. Based on the same or similar type of equipment or failure types brings certain comparability to the comparison of diagnostic performance when using different ensemble learning methods. Thus, this section summarizes the ensemble learning based fault diagnosis methods according to different types of equipment including mechanical bearing, power transformer, gearbox, new energy system, and other equipment. Section 3 compares and summarizes the accuracy of fault diagnosis based on ensemble learning.

2.1. The fault diagnosis methods used for bearings

Bearings are an important component of mechanical power systems. Their healthy and stable operation is closely related to the production activities of mechanical equipment, and thus they have been widely concerned with fault diagnosis. In addition, because of its rich open datasets, it has become one of the main research interests in fault diagnosis.

Most bearing fault diagnosis methods based on ensemble learning use the Case Western Reserve University (CWRU) bearing dataset (Smith and Randall, 2015) for model training and verification. Fig. 5 shows the CWRU bearing data set test stand consisting of a 2 hp motor (left), a torque transducer/encoder (center), a dynamometer (right). The CWRU bearing fault data is collected under 4 classes motor loads, ball, inner race and outer race vibration data (48 khz and 12 khz). Each position contains 4 classes fault diameters. Readers are referred to (Smith and Randall, 2015) for more details on the CWRU bearing failure dataset. Tang et al. (2018) used fast spectral correlation and particle swarm optimization to improve and optimize an ensemble learning algorithm based on random forest. They trained their fault diagnosis model with CWRU bearing datasets. Finally, a random forest bearing fault diagnosis method was proposed based on improved fast spectrum correlation and particle swarm optimization. Wan et al. (2021) proposed a bearing fault diagnosis model based on Spark and improved random forest by combining Spark big data processing technology with improved random forest algorithm. Li et al. (2021a) proposed an optimal ensemble deep transmission network for unlabeled data. This method integrates the advantages of small data volume model of transfer learning and its reliability, the differential adaptability of domain adaptation (the difference of marker amount between source domain and target domain) and the advantages of strong data processing ability, high reliability, high accuracy and easy parallel computing of ensemble learning. He et al. (2020) proposed an ensemble sparse autoencoder based bearing fault diagnosis method and verified that noise has an impact on fault diagnosis results. Han et al. (Han and Jiang, 2016) proposed a fault feature extraction method based on variational mode decomposition (VMD) and autoregression (AR) model parameters, thus forming a bearing fault diagnosis model structure based on VMD-AR and random forest. The model had achieved not only good results in fault location, but also excellent diagnosis results in fault degree classification. Wang et al. (2021a) proposed a multi-round voting bearing fault diagnosis method based on similarity measurement. They verified its advantages over one-round voting and nonensemble methods. Han et al. (2019) proposed a method for long term fault diagnosis for train rolling bearing based on exponential smoothing predictive segmentation and improved ensemble learning algorithm. Zhang et al. (2021a) proposed a multi-model ensemble deep learning method based on deep convolutional neural network (DCNN). The method was verified on both bearing data and gearbox data. It was an important attempt to study the generalization of diagnostic methods. Wang et al. (2023a) proposed an imbalanced ensemble method with dense-net and evidential reasoning rule, to incorporate both human experience and machine wisdom for machinery fault diagnosis under the class-imbalance situation. Jian and Ao (2022) developed a diverse variable weighted model ensemble (DVWME) method for industrial fault diagnosis. Multiple diverse training sets were produced from the primary training set by Bootstrap sampling method. Different base models were used with different hyper parameters to build the ensemble learning method on diverse training sets, which improved the ensemble performance. Che et al. (2022) proposed an ensemble learning model (ELM) for small fault diagnosis of rolling bearings. In this research, the high-dimensional vibration signals of rolling bearings are converted into grayscale images with different fault markers as input. Considering the difference of vibration signals under different state parameters, different meta-learning models were combined using an ensemble learning framework to obtain good diagnosis results. Li et al. (2022) proposed a reinforcement ensemble deep transfer learning network (REDTLN) for multi-source domain fault diagnosis. Firstly, various novel kernel maximum mean square errors were used to construct multiple deep transfer learning networks (DTLNs) for single-source single-target domain adaptation. The differences in kernel functions and source domains can help DTLNs learn different transferable features. Secondly, a new unified metric based on kernel maximum mean square error and diversity metric was designed for unsupervised ensemble learning. Finally, with the unified metric as the reward, it explores a reinforcement learning method to generate effective combination rules for the reinforcement integration of multi-domain and multi-model. In (Sharma et al., 2017), four machine learning techniques including two ensemble techniques (rotation forest and random subspace), and two mature techniques (support vector machines and artificial neural networks) were discussed for fault

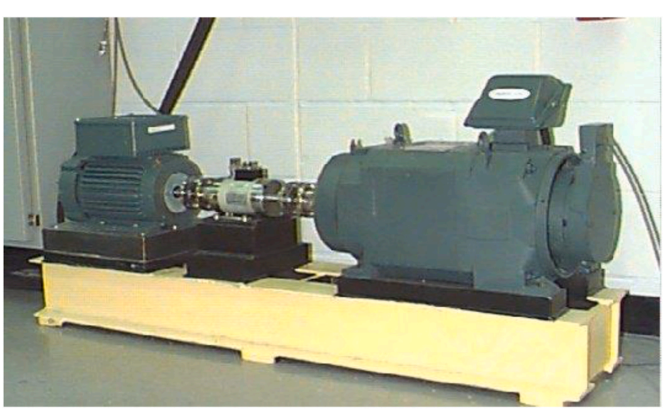


Fig. 5. The CWRU bearing dataset test stand discussed in literature (Smith and Randall, 2015).

severity classification. The results show that the ensemble technique has superior classification efficiency compared to two mature techniques in terms of computation time required. Li et al. (2017) proposed an algorithm called IDSCNN. It is a bearing fault diagnosis algorithm based on ensemble deep convolutional neural network and improved Dempster-Shafer theory-based evidence fusion. The experimental results show that the improved method has better fault diagnosis performance than the existing machine learning methods.

There are several fault diagnosis methods that have used the KAT (Hasan, 2013) public bearing datasets to train their model. Li et al. (2020) proposed an improved domain adaptation method and proposed the necessity of studying imbalanced data processing in the field of fault diagnosis. The KAT datasets were also used for model training. Moreover, there are also a number of bearing fault diagnosis methods using nonpublic datasets (Cao, 2019), (Li et al., 2020), (Wu, 2020), (Jiang, 2020), (Beretta et al., 2021), (Zhang et al., 2020a), (Pang et al., 2021), and (Yang et al., 2021). Furthermore, imbalanced data processing ability is one of critical features in improving the accuracy of fault diagnosis. Wu (2020) demonstrated the influence of imbalanced data processing on fault diagnosis results when they proposed an improved stacking rolling bearing fault diagnosis model. Cao (2019) proposed a dynamic weighted ensemble learning model based on incremental information and proposed a dynamic weighted imbalanced processing method based on sample distribution information. This method implemented an imbalanced incremental real-time processing of data. Jiang (2020) proposed a bearing noise fault classification method based on neural network filter and ensemble learning classification, where the noise signal is used instead of the vibration signal. This implies that various signals and data could be used in fault diagnosis depending on the nature of the equipment and its operation environment. Beretta et al. (2021) introduced an ensemble learning based fault diagnosis method into the fault diagnosis of the main bearing of fans. Deep learning, a technique that has attracted a lot of interest recently, has been applied in the domain of fault diagnosis and achieved promising results. Ma and Chu (2019) proposed an ensemble deep learning method for bearing fault diagnosis based on multi-objective optimization. Pang et al. (2021) proposed an ensemble algorithm based on a deep convolutional extreme learning machine model. Yang et al. (2021) combined the deep learning and ensemble learning and proposed an ensemble deep learning based bearing fault diagnosis method. Xu et al. (2019) proposed an ensemble learning method based DCNN and random forest for fault diagnosis. This method combines deep learning and ensemble learning mothed for fault diagnosis. This implies that ensemble learning is well inclusive. Chen et al. (2021) proposed a fault diagnosis model based on XG-Boost, whose high efficiency and flexibility enabled it to achieve a relatively higher accuracy of fault diagnosis. To further improve the effect of multi-sensor information fusion, reduce the difference between the actual value and the predicted value, and to improve the accuracy of rolling bearing fault diagnosis, Tong et al. (2023) proposed a multi-sensor information integration method based on ensemble learning. They verified the effectiveness and robustness of the proposed method on two different types of rolling bearing datasets. Luo et al. (2016) proposed a hybrid system HGSA-ELM for rolling bearing fault diagnosis. It uses a real-valued gravity search algorithm (RGSA) to optimize the input weights and biases of ELM and uses the binary value of GSA (BGSA) to select important features from the composite feature set. Yu (2016) developed a novel manifold learning algorithm including a global and local/non-local discriminant analysis (GLNDA), The algorithm aims to extract effective intrinsic geometric information from given vibration data. By introducing Bagging into feature selection and fisher discriminant rule (FDR) based on GLNDA scores, a new manifold ensemble method (Selective ensemble GLNDA, SE-GLNDA) for mechanical fault diagnosis is studied. The experimental results of the bearing life tester verify the effectiveness of the fault diagnosis method based on SE-GLNDA. Zhang et al. (2015) proposed an intelligent fault diagnosis method based on multi-variable ensemble incremental support vector machine (MEISVM), which was compared and verified with other methods on the benchmark of rolling bearing experiments. Experimental and engineering test results show that the proposed method is effective in fault diagnosis of rolling bearing with vibration signals. To obtain better generalization ability of fault diagnosis and multiple monitoring variables with corresponding fault patterns, Wang et al. (2018a) proposed a fault diagnosis method (particle swarm optimization based selective ensemble learning, PSOSEN). The adaptive particle swarm optimization (APSO) based on nonlinear reduction of inertia weight is used to effectively strengthen the learning process by selecting excellent individuals for ensemble. In order to ensure the accuracy of the algorithm and the instability of the rolling bearing fault prediction results caused by the random selection of input weights and the deviation of the hidden layer of the extreme learning machine, He et al. (2021) proposed an ensemble error minimization extreme learning machine (EEM-ELM). Wu et al. (2021) developed a deep ensemble dense convolutional neural network (DEDCNN) and verified the effectiveness of the proposed method through a large number of imbalanced data experiments. The experimental results show that the proposed method is superior to the existing methods in dealing with imbalanced data problems. You et al. (2023) proposed an efficient rolling bearing fault diagnosis model with a hybrid neural network with a lightweight attention mechanism. The experimental results show that the model has reliable, robust and efficient performance. Cao et al. (2022) proposed a fault diagnosis method based on ACWGAN-GP and homogeneous superposition ensemble learning to tackle the issue of small sample fault diagnosis. Firstly, the method uses argmax multi-class classification to construct multiple different training subsets. Second, these constructed training subsets are used to train multiple base learners based on ACWGAN-GP. Finally, a meta-learner based on softmax regression is used to fuse these trained base learners to obtain the diagnosis results. Wang and Vinogradov (2023) introduced a general approach for intelligent data analysis that is suitable for continuous monitoring of acoustic emission (AE) signals from sensors in rotating machinery. They also designed a novel architecture with a convolutional generative adversarial network (GAN) to extract deep information embedded in AE waveforms. To improve the robustness of the proposed GAN framework, they introduced an ensemble technique called "historical state ensemble" (HSE) paired with GAN. The proposed ensemble method can improve the robustness of GAN.

It is worth noting that vibration signals are mostly used as the source data in the public datasets for bearing fault diagnosis, whereas a combination of both vibration signals and noise signals are used in nonpublic datasets. This indicates ensemble learning method could use diverse data sources and has sound feasibility and adaptability in bearing fault diagnosis. However, it also shows the shortcomings of the ensemble method in generalization, data imbalanced process ability and noise resistance. These metrics are important criteria for evaluating whether such methods are well-behaved, and this will be discussed in detail in Section 3.

2.2. Fault diagnosis methods for power transformers

Transformers are important components in the power energy transmission system and the fault diagnosis for this equipment is also vital. The data acquisition for transformers fault diagnosis can be mainly divided into three different signals, i.e., the gas volume fraction signal generated from the dissolved gas analysis (DGA) in the transformer oil, the vibration signal, and the current signal of the transformer. Thus, in this section, we discuss the ensemble learning-based fault diagnosis methods according to its data acquisition classification respectively.

DGA is chosen by many scholars in transformer fault diagnosis because of its rich data features. Wang and Han (2020) and Zhang and Wang (2021) constructed a transformer fault diagnosis model based on the classic stacking structure in ensemble learning. Zheng (2011) proposed an improved bagging algorithm model for transformer fault diagnosis and feature prediction based on information entropy and sample entropy. Zhou (2021b) optimized the parameters of cat boost algorithm by Bayesian optimization (BO) and constructed a transformer fault diagnosis model based on BO-Cat Boost. These methods and models utilize the classical structure of ensemble learning to construct the fault diagnosis model. Wang (2015) established a new combination diagnosis model composed by sample weight least squares support vector (LS-SVM) and ensemble learning.

In addition to DGA data, vibration signal and current signal can also be used as resource data for fault diagnosis. Zhang et al. (2020a) proposed a transformer fault diagnosis method based on an internet of things (IOT) monitoring system and ensemble machine learning by using transformer vibration signal as diagnostic data. Raichura et al. (2020) pro-posed a classifier technology based on hierarchical ensemble extreme learning machine by using the current signal output from the transformer as experimental data for transformer fault diagnosis.

Xuan et al. (2022) proposed an image dataset optimization method based on seamless cloning algorithm and image cleaning. It aims to solve the problem that the accuracy of appearance fault diagnosis in intelligent detection images is limited by the number and quality of images. By constructing multiple learning models and using weighted voting strategy to fuse the model outputs into the final output, the accuracy of transformer housing fault diagnosis is improved. Fig. 6 shows the appearance of the transformer provided by (Xuan et al., 2022).

To sum up, most ensemble learning methods used for fault diagnosis of transformer use traditional ensemble methods to design their model. This might be because they are easy to implement and can achieve high accuracy. However, the fault diagnosis methods based on traditional ensemble model do not consider the adaptability of the model to unknown data. The complexity of the real environment often causes the generation of unknown data such as noise, which may cause a failed fault diagnosis of equipment. Being able to handle these features is a key factor with respect to model performance.

2.3. Fault diagnosis methods for gearboxes

The gearbox is another important component of mechanical equipment. This section summarizes ensemble learning based fault diagnosis methods for gearboxes.

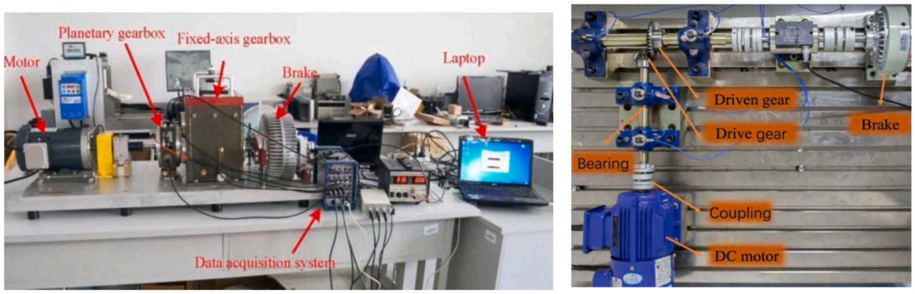
To better verify and improve the generalization of diagnosis methods, Pang et al. (2021) verified the proposed ensemble algorithm based on deep convolutional extreme learning machine of base learner on both bearing and gear-box fault data. Zhu (2017) proposed a selective neural network ensemble algorithm based on artificial bee colony algorithm for gearbox fault diagnosis. They compared the full ensemble and partial ensemble of the base learner and verified the superiority of partial ensemble of the base learner. Kundu et al. (2020) proposed a random forest regression method for predicting the remaining life of gearbox in pitting failure mode. Wang et al. (2021b) proposed a fault diagnosis method for planetary gearbox based on multi-standard fault feature selection and heterogeneous ensemble learning classification by using vibration signals collected by acceleration sensors as experimental data. Zhou et al. (2019) proposed an online gearbox condition monitoring method based on ensemble multiple fault feature metrics method. On-line diagnosis model has a high demand for real-time data processing, and the timeliness is a necessary factor to be considered in system dependability analysis. Therefore, this method can not only provide reference for on-line fault diagnosis model, but also provide theoretical and methodological support for system dependability analysis. Kang et al. (2022) introduced a new ensemble convex hull (EnCH) based ensemble model. It aims to solve the problem that geometric learning models have difficulty generating sub-classifiers with differences. Experimental results on the gearbox fault dataset show that the EnCH classification model can improve the generalization ability of the geometric learning model and has good tolerance to noise and outliers. Cheng et al. (2016) developed a planetary gear fault diagnosis method based on entropy feature fusion of ensemble empirical mode decomposition (EEMD). It is stated that the fault recognition rate of normal gear, tooth missing gear and broken gear can reach 100%. Fig. 7 shows two different experimental platform setups discussed in (Kang et al., 2022) (Fig. 7 (b)) and (Cheng et al., 2016) (Fig. 7 (a)). Fig. 7 (b) shows the composition of the mechanical system. Besides, Fig. 7 (a) also marks the data acquisition system and notebook terminal. These two gearbox fault data experimental platforms have obvious differences in the mechanical and data acquisition system, though both collect the same vibration signal data. This implies that although different methods use the same data types, their equipment and data acquisition platforms may be different, which also lead to poor comparability between different models.

In summary, it is feasible and effective to apply ensemble learning methods to gearbox fault diagnosis and monitoring. To better adapt the algorithm to a wider range of applications, however, it is particularly important to study the model generalization performance. Furthermore, with higher industrial demand, on-line real-time diagnosis is a key point in fault diagnosis research. Timeliness is also a very important performance metric in system dependability analysis. Realtime analysis, however, has high requirements on data acquisition and processing. The parallelism of ensemble learning can help in solving the problem of data computation and processing.

2.4. Fault diagnosis methods for new energy systems

A new energy system refers to an energy system that uses non-fossil fuels to provide power resources for daily production and living. The widely used new energy systems consist of nuclear power generation system, wind power generation system and photovoltaic power generation system. The ensemble learning has been applied to a few new energy systems. Li and Lin (2021) integrated multiple base learners through multiple voting method and weighted voting method respectively and proposed an ensemble learning method based intelligent fault diagnosis in nuclear power plants. Kapucu and Cubukcu (2021) proposed a fault diagnosis method for photovoltaic power generation system based on ensemble learning method. They verified the feasibility and effectiveness of real-time fault diagnosis on photovoltaic system. Wang (2021) used the XGBoost algorithm, an ensemble learning algorithm optimized by a grid search algorithm to diagnose photovoltaic array failures. This method, however, has not been proven to be well-behaved under the condition of large volume of data (tens of thousands) and multiple sensors connected to many photovoltaic system components. Xu (2020) used the sine cosine stochastic optimization

Fig. 6. Appearance diagram of the transformer discussed in the literature (Xuan et al., 2022).



(a)   (b)

Fig. 7. Different gearbox test bench discussed in (Cheng et al., 2016) (a) and (Kang et al., 2022) (b).

algorithm to optimize the smoothing factor of the probabilistic neural network (PNN) algorithm and integrated the optimized PNN through the AdaBoost model to form a fault diagnosis model. Eskandari et al. (2020) proposed an intelligent fault diagnosis method for photovoltaic system based on ensemble learning model and current-voltage (I–V) characteristics. These methods have achieved excellent results in fault diagnosis accuracy. However, there is a lack of analysis for concurrent faults, and most diagnostic methods in the references do not involve the study of concurrent problems of multiple faults. There is usually not only one kind of fault in the real environment, and various concurrent faults will cause or lead to other problems at the same time. Therefore, a robust fault diagnosis model must be capable of diagnosing concurrent faults. Badr et al. (2023) introduced a novel strategy that combines the ensemble learning concept with a self-training based semi-supervised learning method to achieve concurrent fault diagnosis. These concurrent faults include photovoltaic (PV) module radian, connecting lines, power tracking units, open circuit and partial shading. The developed ensemble learning paradigm contains multiple merged machine learning models that can improving overall diagnostic performance. Zhong and Ban (2022) proposed an ensemble learning method for fault diagnosis of rotating machinery in nuclear power plants to mitigate the negative impact of noise on field fault data and plant measurement data. Wang et al. (2019a) proposed an ensemble fault diagnosis framework for simultaneous and coupled failures. This framework solves the fuzzy boundary of complex equipment and the individual variation of invisible single or simultaneous faults. The experimental results show that this framework can detect single and simultaneous faults accurately and quickly. Zhang et al. (2018) combined random forest (RF) with extreme gradient Boosting (XGBoost) to build a data-driven fault detection framework for wind turbines. The experimental results show that the proposed ensemble classifier can prevent overfitting. In addition, it can obtain better wind turbine fault detection results than support vector machine (SVM) method when dealing with multi-dimensional data. Mellit et al. (2023) proposed a new technique to identify and classify faults in small photovoltaic (PV) systems. The fault detection rate of this method is 98.56%, and the classification rate is 96.21%. They conducted a comparative study of the proposed model with different ensemble learning algorithms, including AdaBoost, CatBoost and XGBoost, and demonstrated that the proposed method can effectively address some fault problems of small photovoltaic systems. Wang et al. (2021c) aim at the problem that the strength of the fault factor itself was ignored in the existing research. They proposed a new method for PV module fault diagnosis based on heterogeneous ensemble learning of current-voltage characteristic curves and environmental conditions. Compared with the single classifier and the ensemble model based on stacking algorithm, the proposed method has more comprehensive diagnostic ability.

In summary, the application of ensemble learning method in new energy system has achieved significant success. It shows the feasibility and effectiveness of using ensemble learning method for fault diagnosis for new energy system. However, it still needs to be further explored in terms of multiple types of concurrent faults, large amount of data and real-time diagnosis. These problems are common to other equipment fault diagnosis model and their solution will be an important improvement for fault diagnosis.

2.5. *Fault diagnosis methods for other equipment*

In the domain of engine systems, Wang et al. (2019b) proposed an intelligent diagnosis method integrated with broad learning and ensemble learning models for aircraft engine wear and faults based on oil analysis. Kowalski et al. (2017) proposed a fault diagnosis method for marine diesel engine based on diesel engine exhaust fraction data. These methods are trained and validated respectively on failure data of aircraft engines and diesel systems.

In the domain of refrigeration, Zhang et al. (2020b) modeled the failure of refrigeration systems by ensemble five machine learning models with the method of majority voting. Han et al. (2020) proposed a fault detection and diagnosis method for building energy system. They integrated three machine learning methods using a majority voting method and verified the feasibility and dependability of the modified model by using fault data training of air-conditioning chillers. Both methods adopt optimized and improved majority voting method as the main ensemble strategy. To alleviate the instability problem and improve the prediction accuracy, Wang et al. (2018b) developed an ensemble learning technique called "integrated bagging trees" (EBT). The results show that the proposed EBT model can predict the hourly electricity demand of the test building, and the mean absolute prediction error accuracy is improved by 2.97%.

In the railway system domain, we summarized several fault diagnosis models applying ensemble learning in rail-way turnouts, hydraulic pumps, transistors, and compressor system. Liu (2020) used stacking algorithm to diagnose railway turnoff faults by monitoring the status of railway turnoff current. Li et al. (2021b) proposed a fault diagnosis method based on improved ensemble empirical mode decomposition, auto-regressive spectral energy, and wavelet kernel extreme learning machine, which achieved 100% diagnostic accuracy in the fault datasets of hydraulic pump. Xia et al. (2019) proposed an open-circuit fault diagnosis method for insulation-gate bipolar transistor of three-phase pulse-width-modulated (PWM) converter based on data drive, and the diagnostic accuracy of experimental data was 99%. Zhang et al. (2020c) proposed a compressor fault diagnosis method based on convolutional deep confidence network to solve the problems of original signal noise reduction and multi-source information fusion in fault diagnosis of reciprocating compressors. This method adopted a new multi-source information fusion framework for improving the performance of fault diagnosis.

Zhong et al. (2018) developed a new framework combining feature extraction, probabilistic committee machine, and decision threshold optimization based on fair evaluation (F-measure). It was used to address the challenges of simultaneous fault diagnosis and multi-signal analysis of automotive engines. In the proposed framework, feature extraction techniques are designed by combining EEMD + SVD + DK to effectively capture individual fault components from simultaneous fault patterns. Moreover, it can diagnose single faults and simultaneous faults when only single fault data is used for training. Tang et al. (2016) pointed out that the selective existing ensemble (SEN) modeling methods based on the “input feature manipulation” from the perspective of multi-information fusion, which cannot selectively and jointly fuse the information hidden in multi-scale spectral features. They proposed a soft sensor method based on ensemble empirical mode decomposition (EEMD) and SEN. The industrial application of the EEMD-SEN method is then discussed, and a new virtual sample generation method is proposed to tackle the modeling problem based on small sample spectral data. The existing techniques rely on hand-designed features, and only use the traditional single shallow machine model as the basic classifier, which does not have the ability to self-learn meaningful features. The diagnostic performance of single-base classifiers sometimes fails to meet engineering requirements. Wang et al. (2023b) proposed a heterogeneous ensemble deep neural network (DNNs) method based on bagging strategy for multi-component fault diagnosis of hydraulic systems. Compared with methods, the proposed method can accurately diagnose hydraulic system faults.

Chowdhury et al. (2022) proposed a data-driven fault diagnosis method for 3D printer. The experimental results show that the accuracy rate of fault diagnosis is 99.75%. Zhang et al. (2023a) proposes the refrigerant charge amount (RCA) fault diagnosis strategy of variable refrigerant flow (VRF) systems based on Stacking ensemble learning, this method has a good diagnosis effect on fault level. Hou et al. (2023) studied the initial surge detection and diagnosis (ISDD) of centrifugal compressors based on bearing vibration signals. They developed a data-driven intelligent diagnosis model using adaptive feature fusion and sparse ensemble learning methods. Gao et al. (2019) studied a mechanical fault diagnosis method for high-voltage circuit breakers based on hybrid feature extraction and integrated extreme learning machine (IELM). This method can effectively diagnose mechanical faults of high-voltage circuit breakers with small samples. Zhou et al. (2021a) proposed a text-based fault diagnosis model. In this method, Word2vec is used to map text words into vector space, and then the extracted text feature vectors are input into a classifier based on stacked ensemble learning scheme. They used a real aircraft fault text dataset for validation. The results show that the fault diagnosis accuracy of the proposed method is 97.35%, which is about 2% higher than that of the suboptimal method. Gou et al. (2018) designed an ensemble ELM classifier to improve accuracy of sensor fault diagnosis, the method is stated to be robust to fluctuations in the overhead line voltage and DC-link voltage, fault severity, and changes in model parameters, speed, and load. Wang et al. (2015) proposed an improved ensemble binary tree support vector machine for multi-classification. The algorithm considers the Bagging learning technique and ensemble more improved binary tree SVMS, so that each SVM weak learning machine has its own kernel function and parameters. The proposed algorithm improves the performance of the model by effectively integrating the best features of each SVM. The fault diagnosis of blast furnace in ironmaking and smelting process has achieved good results.

In summary, ensemble learning methods have been studied widely in the field of fault diagnosis. The review shows that the application of ensemble learning in machinery fault diagnosis is feasible and effective. However, there are also several remaining problems in terms of noise resistance, generalization, imbalanced data processing ability and data adaptability of different data sources. The fault diagnosis for transformers relies mainly on traditional ensemble strategies and there is a lack of new strategies. There are real-time online diagnosis problems in fault diagnosis for gearboxes. For new energy system, there is insufficient research in terms of multiple types of concurrent failures and large amounts of data.

3. Ensemble learning based fault diagnosis methods

This section first summarizes the fault diagnosis dataset and the application of the dataset of the integrated fault diagnosis method. Then, this section analyzes the performance of summarized ensemble learning methods for fault diagnosis. This paper collated fault diagnosis methods based on ensemble learning, including accuracy, fault type, input data, and proposed methods, and conducted a simple analysis using accuracy rates. This paper using the accuracy of fault diagnosis as the main evaluation metrics supplemented by generalization and imbalanced data processing ability to evaluate the performance of those ensemble learning method. The evaluated data is obtained from references. The optimal performance of diagnosis accuracy will be chosen if there is an optimal performance discussed in the paper. The average accuracy of fault diagnosis is calculated if there are experimental results shown under various conditions in the paper. In most cases, the values of accuracy are collected directly from the reference.

3.1. Dataset for fault diagnosis

Datasets are the heart of data-driven intelligent fault diagnosis research. A device with more publicly available datasets allows more research to focus on such system devices (e.g., bearings). In order to facilitate intelligent fault diagnosis research, it is meaningful to collect and organize rich fault data. This paper summarizes the datasets from references and open platforms such as prognostics and health management (PHM), Git Hub, university websites, kaggle.com, and IEEE. These datasets include fault data for critical systems such as bearings, gears and power grids. Table 2 shows the equipment types, download links, and published dataset descriptions for the dataset. It can be seen from Table 2 that most of the failure data is for mechanical equipment (bearings, gears, etc.). The lack of diversity in fault data is a challenge to both the breadth and generality of fault diagnosis research. Therefore, fault diagnosis research still requires enriched datasets.

Table 3 shows the dataset used to apply the ensemble learning based fault diagnosis method to different equipment. As can be seen from Table 3, several bearing-based fault diagnosis methods all use public datasets. This makes it possible to compare their performance. For fault diagnosis of other system equipment, the test datasets are all based on non-public datasets, which makes performance comparison difficult. However, in either case, the accuracy of fault diagnosis is still the main metric to judge the performance of these methods. In addition, Table 3 shows that bearing-based fault diagnosis methods not only have more public test data than other equipment, but also have rich public datasets. This could be the reason bearing-based fault diagnosis has become a popular research area.

3.2. Performance comparison and analysis

Table 4 to Table 8 list the diagnostic accuracy of fault diagnosis for bearings, gearboxes, transformers, new energy systems, and other equipment with used ensemble learning methods, input data and fault types. The symbol (# or *) in the following tables indicate that the model covers generalization (#) or imbalanced data processing (*). Generalization and imbalanced data processing will be discussed in the next two subsections. The N/A symbol indicates that there is no such data, or the method does not use the accuracy of fault diagnosis as the evaluation index.

Table 4 shows the accuracy of fault diagnosis for bearing with type of input data required by the model and the classification of fault types. In Table 4, vibration signal is mostly used as input data, and the most studied bearing fault types are normal, inner-race fault, outrace fault and element fault for bearing fault diagnosis.

Combining Table 2 and Tables 4, it can be seen that the same test data were used in (Tang et al., 2018), (Wan et al., 2021), (Li et al., 2021a), (He et al., 2020), (Han and Jiang, 2016), (Wang et al., 2021a),

Table 2
Public datasets.

Equipment	Download Link	Related papers
Milling	https://phm-datasets.s3.amazonaws.com/NASA/3.+Milling.zip	Agogino and Goebel (2021)
Bearings (IMS)	https://phm-datasets.s3.amazonaws.com/NASA/4.+Bearings.zip	Lee et al. (2007)
Batteries	https://phm-datasets.s3.amazonaws.com/NASA/5.+Battery + Data + Set.zip	(Saha and Goebel)
Turbofan Engine Degradation Simulation	https://phm-datasets.s3.amazonaws.com/NASA/6.+Turbofan + Engine + Degradation + Simulation + Data + Set.zip	Celaya et al. (2009)
Prognostics Health Management 8 (PHM08) Challenge	https://phm-datasets.s3.amazonaws.com/NASA/8.+IGBT + Accelerated + Aging.zip	Saxena and Goebel (2008)
Fatigue Crack Growth in Aluminum Lap Joint	christopher.a.teubert@NASA.gov	Peng et al. (2015)
Bearings (CWRU)	https://engineering.case.edu/bearingdatacenter/download-data-file	Smith and Randall (2015)
Gearbox (Southeast University)	https://github.com/cathysiyu/Mechanical-datasets	Shao et al. (2019)
Bearings (KAT)	http://groups.uni-paderborn.de/kat/Bearing Data Center	Hasan (2013)
Bearings (XJTU-SY)	http://biaowang.tech/xjtu-sy-bearing-datasets/	Wang et al. (2018c)
Bearings (FEMTO-ST)	https://github.com/hustcxl/Rotating-machine-fault-data-set/blob/master/doc/FEMTO_ST.md	Porotsky and Bluvband (2012)
Gearbox (University of Connecticut)	https://figshare.com/articles/dataset/Gear_Fault_Data/6127874/1	Cao et al. (2018)
DIRG Bearing Data	https://www.mmnt.net/db/0/0/ftp.polito.it/people/DIRG_BearingData/	Daga et al. (2019)
Electricity Transformer Dataset	GitHub - zhouhaoyi/ETDataset: The Electricity Transformer dataset is collected to support the further investigation on the long sequence forecasting problem.	Zhou et al. (2021b)
Defect cases of the centrifugal pump	https://IEEE-dataport.org/documents/acoustic-and-vibration-data-defect-cases-centrifugal-pump#files	Kumar and Kumar (2022)
Defecting and diagnosis rotor broken bar in a three-phase induction motor	https://IEEE-dataport.org/open-access/experimental-database-detecting-and-diagnosing-rotor-broken-bar-three-phase-induction	Treml et al. (2020)
Fault Detection Dataset in Photovoltaic Farms	https://www.techscience.com/iasc/v30n2/44023	(Saha and Goebel)
Bearing	https://www.kaggle.com/datasets/pouriyaamini/intelligent-bearing-fault-diagnosis-dataset	Chegini et al. (2022)

(Han et al., 2019), (Zhang et al., 2021a), (Wang et al., 2023a), (Jian and Ao, 2022), (Che et al., 2022), (Li et al., 2022), (Sharma et al., 2017), (Li et al., 2017), (Li et al., 2020)– (Jiang, 2020), (Xu et al., 2019), (Wu, 2020), (Cao, 2019), (Tong et al., 2023), (Zhang et al., 2015), (Chen et al., 2021), (He et al., 2021), (You et al., 2023), (Badr et al., 2023) and (Cao, 2019)*. Based on this scenario, in Tables 4 and it can be seen that references (Tang et al., 2018), (Beretta et al., 2021) and (Wu, 2020)*, have the best diagnostic accuracy rate (100%). The lowest accuracy of diagnosis is identified for (Li et al., 2021a), (Han et al., 2019), (Cao, 2019)* and (Wang and Vinogradov, 2023). Among all the literatures with * (considering imbalanced data processing), (Wu, 2020)* has the highest accuracy and (Cao, 2019)* has the lowest accuracy. Note that (Wu, 2020)* not only considers imbalanced data processing, but also has a 100% diagnosis accuracy, which seems to indicate that this method has the best performance. From the perspective of generalization, references (Zhang et al., 2021a)#, (Pang et al., 2021)#, and (Wang et al., 2018a)# considered the generalization. Compared the three, the multi-model ensemble deep learning method adopted in (Zhang et al., 2021b) has higher fault diagnosis accuracy (99.48%). Note that there is no fault diagnosis accuracy rate data for (Wang et al., 2023a), (Yu, 2016), and (Cao et al., 2022) since they do not use accuracy as an evaluation index for the proposed method. For example, Wang et al. (2023a) used Macro-F1 and MG-mean evaluation models, while (Cao et al., 2022) used cosine similarity and person correlation coefficient evaluation models. The method proposed in (Yu, 2016) is used to predict the remaining useful life of bearings.

It is worth noting that reference (Li et al., 2020)#* is the only work which considered both generalization and imbalanced data processing. The fault diagnosis accuracy of (Li et al., 2020)#* is relatively high (99.4%). This implies that the method not only has higher capability for handling new, unseen and in-distribution samples but also could lead to higher diagnostic accuracy.

Table 5 summarizes a number of different ensemble learning methods used for the fault diagnosis for transformer fault types. It can be seen that (Zheng, 2011) has the lowest rate (81.9%) whereas (Raichura et al., 2020) has the highest rate (99.91%). It is worth noting that though (Zheng, 2011) has the lowest diagnosis accuracy rate, it diagnoses the most fault types. In Table 5, more than 60% of the transformer fault diagnosis methods use the gas analysis data as input data for verifying their methods. This implies that gas analysis data could be selected for similar fault types. In Table 5, it is interested that references (Wang and Han, 2020), (Zhang and Wang, 2021) and (Wang, 2015) have similarity in terms of methods, input data and fault types. Does this mean that methods discussed in (Wang, 2015) and (Wang and Han, 2020) are better than (Zhang and Wang, 2021) according to accuracy rate of 90.42%, 89.36% and 83.3% respectively? The diagnosis results of the ensemble learning methods are dependent on the learner. Factors such as the number of learners and learning ability will have a great influence on the results of the adopted ensemble method. The verification environment used by the designers of each method is not uniform and different environment designs are also important factors that affect the learning results of the ensemble model. However, it might be necessary to further explore the learning ability of (Zhang and Wang, 2021) since e.g., it has the same setting with (Wang and Han, 2020) but 6% lower in accuracy rate.

Table 6 shows the methods, input data, fault diagnosis accuracy and fault types for gearbox fault diagnosis based on ensemble learning. In Table 6, though the diagnostic accuracy of (Zhou et al., 2019) is relatively high (97.5%), its significance is limited due to the vague description of the verified fault classification. In contrast, the others classify gearbox faults more clearly while having high fault diagnosis accuracy. Note that references (Zhang et al., 2021a)#, (Pang et al., 2021)#, and (Wang et al., 2018a)# investigated both the gearbox and bearing fault diagnosis (see Table 4). This implies that (Zhang et al., 2021a)#, (Pang et al., 2021)#, and (Wang et al., 2018a)# may be suitable for scenarios where there are fault diagnosis requirements for both bearings and gearboxes. It is also worth noting that (Zhang et al., 2021a)#, (Pang et al., 2021)#, and (Wang et al., 2018a)# not only consider generalization but also produce relatively high accuracy rate. This implies that the adopted methods might have better performance from a generalization perspective. The N/A for (Kundu et al., 2020) shown in Table 6 means that there is no data for fault diagnosis

Table 3
Dataset used for ensemble learning based fault diagnosis.

References	Equipment	Data set	Public
(Tang et al., 2018), (Wan et al., 2021), (Li et al., 2021a), (He et al., 2020), (Han and Jiang, 2016), (Wang et al., 2021a), (Han et al., 2019), (Zhang et al., 2021a), (Wang et al., 2023a), (Jian and Ao, 2022), (Che et al., 2022), (Li et al., 2022), (Sharma et al., 2017), (Li et al., 2017), (Wu, 2020), (Cao, 2019), (Jiang, 2020), (Beretta et al., 2021), (Zhang et al., 2020a), (Pang et al., 2021), (Yang et al., 2021), (Ma and Chu, 2019), (Xu et al., 2019), (Chen et al., 2021), (Tong et al., 2023), (Luo et al., 2016), (Yu, 2016), (Zhang et al., 2015), (Wang et al., 2018a), (He et al., 2021), (Wu et al., 2021), (You et al., 2023), (Cao et al., 2022), (Wang and Vinogradov, 2023)	Bearing	CWRU	YES
Li et al. (2020)	Bearing	KAT	YES
(Wu, 2020), (Cao, 2019), (Jiang, 2020), (Beretta et al., 2021), (Zhang et al., 2020a), (Pang et al., 2021), (Yang et al., 2021), (Ma and Chu, 2019), (Xu et al., 2019), (Chen et al., 2021), (Tong et al., 2023), (Luo et al., 2016), (Yu, 2016), (Zhang et al., 2015), (Wang et al., 2018a), (He et al., 2021), (Wu et al., 2021), (You et al., 2023), (Cao et al., 2022), (Wang and Vinogradov, 2023)	Bearing	Self-collection	NO
(Zhang et al., 2020a), (Wang and Han, 2020), (Zhang and Wang, 2021), (Zheng, 2011), (Zhou, 2021b), (Wang, 2015), (Raichura et al., 2020), (Xuan et al., 2022)	Transformer	Self-collection	NO
(Zhang et al., 2021a), (Pang et al., 2021), (Wang et al., 2018a), (Kundu et al., 2020), (Wang et al., 2021b), (Zhou et al., 2019), (Kang et al., 2022), (Cheng et al., 2016)	Gearbox	Self-collection	NO
(Li and Lin, 2021), (Kapucu and Cubukcu, 2021), (Wang, 2021), (Xu, 2020), (Eskandari et al., 2020), (Badr et al., 2023), (Zhong and Ban, 2022), (Wang et al., 2019a), (Zhang et al., 2018), (Mellit et al., 2023), (Wang et al., 2021c)	New energy system	Self-collection	NO
(Wang et al., 2019b), (Kowalski et al., 2017), (Zhang et al., 2020b), (Han et al., 2020), (Wang et al., 2018b), (Liu, 2020), (Li et al., 2021b), (Xia et al., 2019), (Zhang et al., 2020c), (Zhong et al., 2018), (Tang et al., 2016), (Wang et al., 2023b), (Chowdhury et al., 2022), (Zhang et al., 2023a), (Hou et al., 2023), (Gao et al., 2019), (Zhou et al., 2021a), (Gou et al., 2018), (Wang et al., 2015)	Other equipment	Self-collection	NO

accuracy. This work mainly studied the remaining useful life of gearboxes rather than fault classification.

Table 7 shows the use of ensemble learning based fault diagnosis methods in the field of new energy systems. In Table 7, it can be seen that the fault diagnosis accuracy rate for new energy system is higher than 85%. The highest fault diagnosis accuracy rate reaches 100% (Xu, 2020). It seems that the application of ensemble learning based fault diagnosis method is relatively successful in the field of new energy system. In addition, from the perspective of fault types, some methods could only diagnose and identify fewer fault types. Fewer fault types mean simpler nonlinear fault relationship fitting requirements, which may lead to higher fault diagnosis accurate easier. Fewer fault types may also imply incomplete fault diagnosis, which may not be suitable for fault diagnosis of equipment with multiple and complex fault types. References (Li and Lin, 2021), (Zhong and Ban, 2022) and (Wang et al., 2019a) seem to be more suitable for this situation where multiple fault types are considered. The N/A shown in Table 7 for (Zhang et al., 2018) means that there is no data for fault diagnosis accuracy, but it uses the hit rate (also known as recall) evaluation model.

Table 8 enumerates the application of ensemble learning based fault diagnosis in other different equipment, including aeroengines, marine diesel engines, refrigeration system, building energy system, railway switch, hydraulic pump, IGBT open-circuit, and reciprocating compressor et al. This indicates that ensemble learning-based fault diagnosis methods could achieve relatively high diagnostic accuracy in a wide range of equipment with different fault types and input data.

It is worth noting that reference (Li et al., 2021b) reported a 100% of fault diagnosis accuracy. However, it seems that (Li et al., 2021b) studied fewer fault types, and this may lead to a higher probability in the success of fault diagnosis. Furthermore, though different systems are described in (Zhang et al., 2020b) and (Han et al., 2020), they target the same fault types and input data. Compared with (Zhang et al., 2020b) and (Han et al., 2020), (Han et al., 2020) yields better results. This is probably because (Han et al., 2020) integrates member optimization into its ensemble learning method, which produces better member learners. This implies that optimization of the learners could impact on the learning of the ensemble model. The N/A shown in Table 8 for (Wang et al., 2018b), (Zhong et al., 2018), and (Tang et al., 2016) means that there is no data for fault diagnosis accuracy. However, they did use other evaluation metrics, i.e., the Mean Absolute Percentage Error (MAPE) (Wang et al., 2018b), F-measure (Zhong et al., 2018) and the relative prediction stability (RPS) (Tang et al., 2016) to evaluate the model.

From Table 4 to Tables 8 and it is hard to fairly evaluate the performance of those ensemble learning methods. There could be many reasons for this. Firstly, there are differences in data acquisition and processing methods among different learning method models. In addition, the differences in data acquisition could result in differences in using fault diagnosis methods. For example, in mechanical equipment such as bearings and gearboxes, most of the fault data are collected from the vibration signals of the equipment during operation. To obtain more complete data features, one can obtain more abundant fault features from various signal types. For example, literatures (Glowacz, 2015), (Glowacz, 2016a) and (Glowacz, 2016b) discussed the use of acoustic signals as the acquisition data for motor rotor fault diagnosis. In addition, images have been used in literature (Xuan et al., 2022) and text data have been used in literature (Zhou et al., 2021a). Secondly, different methods have different performance requirements other than accuracy. For example, the diagnostic accuracy in paper (Cao, 2019) is 90.54%, which is lower than the average accuracy. This is because it considers the factor of unbalanced data processing capacity, however, some other studies rarely discuss this issue.

It is worth noting that literatures (Wang et al., 2023a), (Cao et al., 2022), (Wang et al., 2018b), (Zhong et al., 2018), (Tang et al., 2016), and (Li et al., 2019) do not use accuracy as a metric to evaluate their model's performance, but this doesn't mean these methods have worse diagnostic performance than others. The use of recall and F-measure is better than that of accuracy in terms of fault diagnosis especially in the case of imbalanced dataset. This will be discussed in more detail later.

3.3. Technology comparison and analysis

The ensemble learning based fault diagnosis method can obtain higher fault diagnosis accuracy than single learner. The main structure of ensemble learning is learners and ensemble strategies, both of which are summarized in this section.

Generally, the types of learners can be divided into machine learning-based learners and neural networks and deep learning-based learners. Fig. 8 (a) shows the statistics of learner types used in the

Table 4
The accuracy of fault diagnosis for bearing with used method, input data type and classification of fault types.

References	Method	Accuracy (%)	Input data	Fault types
Tang et al. (2018)	Improved fast spectral correlation and optimized random forest.	100	Vibration signals.	The normal, inner-race fault, out-race fault, element fault.
Wan et al. (2021)	Spark and improved random forest algorithm.	98.12	Vibration signals.	
Li et al. (2021a)	Optimal ensemble deep transfer network.	84.27	Vibration signals.	
He et al. (2020)	An ensemble sparse autoencoder.	99.71	Vibration signals.	
Han and Jiang (2016)	VMD-AR model and random forest classifier.	98.13	Vibration signals.	
Wang et al. (2021a)	A multi-round voting method based on similarity measurement.	97.5	Vibration signals.	
Han et al. (2019)	Exponential smoothing predictive segmentation and improved ensemble learning algorithm	90	Vibration signals.	
(Zhang et al., 2021a)#	Multi-model ensemble deep learning method based on DCNN.	99.48	Vibration signals.	
(Li et al., 2020)#*	An improved domain adaptation method.	99.4	Vibration signals.	
(Wu, 2020)*	An improved Stacking model.	100	Vibration signals.	
(Cao, 2019)*	A dynamic weighted ensemble learning model.	90.54	Vibration signals.	
(Jiang, 2020)*	NN and XG-Boost.	97.5	Vibration signals.	
Wang et al. (2023a)	IEMD-ER	N/A	Vibration signals.	
Jian and Ao (2022)	DVWME	97.2	Vibration signals.	
Badr et al. (2023)	Crack fault diagnosis of rotating machine in nuclear power plant based on ensemble learning.	92.85	Vibration signals.	
Che et al. (2022)	ELM.OVO	96	Vibration signals.	
Li et al. (2022)	REDTLN	99.70	Vibration signals.	
Tong et al. (2023)	MSCNN	99.1	Vibration signals.	
Xu et al. (2019)	DCNN (Deep convolutional neural network) and random forest ensemble learning.	97.38	Vibration signals.	
Chen et al. (2021)	The XG-Boost integrated algorithm.	92.89	Vibration signals.	
He et al. (2021)	EEM-ELM	98.21	Vibration signals.	
Sharma et al. (2017)	Novel ensemble techniques for classification of rolling element bearing faults.	91.7	Vibration signals.	
(Pang et al., 2021)#	Cross-domain negative Correlated ensemble algorithm.	94.73	Vibration signals.	
Yu (2016)	Machinery fault diagnosis using joint global and local/nonlocal discriminant analysis with selective ensemble learning.	N/A	Vibration signals.	
Zhang et al. (2015)	MEISVM	98.5	Vibration signals.	
Beretta et al. (2021)	An ensemble of neural network and an isolation forest algorithm.	100	Temp, and rotor speed.	All fault and normal status classifications come from a ticket log. There is no detailed description in the text.
Ma and Chu (2019)	An ensemble deep learning diagnosis method based on multi-objective optimization.	98.09	Vibration signals.	The misalignment fault, the unbalance fault, the displacement caused by loosening of the base, the rub-impact fault, fatigue crack, the outer race, inner race, cage and one of the rollers.
Yang et al. (2021)	Non-equivalent costs of fault severities for rolling bearing diagnostics.	99.28	Vibration signals.	The normal, inner-race fault, out-race fault and element fault. Each fault location is divided into three types: mild fault, moderate fault and severe fault.
Luo et al. (2016)	HGSA-ELM	98.93	Vibration signals.	Six different fault conditions (the fault in inner race (IR), outer race (OR) and ball (B)) with defect diameters (0.007 in. and 0.021 in.).
(Wang et al., 2018a)#	PSOSEN	95.34	Vibration signals.	Normal, Servo valve fault, Hydraulic motor fault, Rotary actuator fault, Shaft fault.
Wu et al. (2021)	DEDCNN	98.42	Vibration signals.	Normal state, slight outer race fault, severe outer race fault, ball fault, inner race fault and four compound faults.
Cao et al. (2022)	A Homogeneous Stacking Ensemble Learning Model for Fault Diagnosis of Rotating Machinery with Small Samples	N/A	Vibration signals.	Good, chipped, ball, inner race, outer race, imbalance, keyway sheared.
Wang and Vinogradov (2023)	Improving the Performance of Convolutional GAN Using History-State Ensemble for Unsupervised Early Fault Detection with Acoustic Emission Signals	87.32	Acoustic emission (AE)	No damage, 0.5 mm crack, 1 mm crack, 1.5 mm crack, 2 mm crack.

Table 5
The accuracy of fault diagnosis for Transformer with used method, input data type and classification of fault types.

References	Method	Accuracy (%)	Input data	fault types
Wang and Han (2020)	Stacking ensemble learning.	89.36	Gas volume fraction	Normal, medium and low-temperature overheating, high-temperature overheating, low energy discharge, high energy discharge and partial discharge.
Zhang and Wang (2021)	Stacking-Ensemble Meta-Algorithms.	83.3	Gas volume fraction	Normal, low energy discharge, high energy discharge, medium and low temperature overheating and high temperature overheating.
Zheng (2011)	A bagging model Improved algorithm	81.9	Gas volume fraction	According to the transformer function, transformer faults was divided into magnetic circuit (iron core), circuit (winding, tap changer, etc.), insulation, sealing, cooling, detection and protection, and mechanical structure and other subsystems faults. According to the parts, transformer faults was divided into faults in windings, iron cores, tap changers, bushings, insulating oil, cooling systems, protection devices, measuring systems, oil tanks.
Zhou (2021b)	BO-Cat Boost.	96	Gas volume fraction	Low temperature overheating, medium temperature overheating, high temperature overheating, partial discharge, low energy discharge, high energy discharge.
Wang (2015)	Combination LS-SVM and ensemble learning.	90.42	Gas volume fraction	Normal, medium and low-temperature overheating, high-temperature overheating, low energy discharge, high energy discharge.
Zhang et al. (2020a)	Using IoT based monitoring system and ensemble machine learning.	99.38	Vibration signals.	The transformer faults are divided into one normal operation condition, winding deformation, winding overlapping, winding losseness, core deformation, winding deformation & core deformation, winding overlapping & core deformation, winding losseness &core deformation.
Raichura et al. (2020)	Using hierarchical ensemble extreme learning machine technique.	99.91	Current signals.	Internal fault conditions, external fault, external fault with CT saturation conditions, and various inrush conditions.
Xuan et al. (2022)	Appearance Fault Diagnosis of a Transformer Based on Data Set Optimization and Heterogeneous Model Ensemble	96	Image	Rust and normal

ensemble learning based fault diagnosis methods from Table 4 to Table 8. Although there are more (53.33%) learning methods based on machine learning (Tang et al., 2018), (Wan et al., 2021), (Yu, 2016), (Zhang et al., 2015), (Wang and Han, 2020), (Zhang et al., 2021), (Zheng, 2011), (Zhou, 2021b), (Wang, 2015), (Zhu, 2017), (Kundu et al., 2020), (Wang et al., 2021b), (Han and Jiang, 2016), (Zhou et al., 2019), (Kang et al., 2022), (Cheng et al., 2016), (Kapucu and Cubukcu, 2021)– (Eskandari et al., 2020), (Zhong and Ban, 2022), (Wang, 2021), (Xu, 2020), (Zhang et al., 2018), (Han et al., 2019), (Wang et al., 2019a), (Mellit et al., 2023), (Wang et al., 2019b), (Zhang et al., 2020b), (Han et al., 2020), (Wang et al., 2018b), (Liu, 2020), (Li et al., 2021b), (Zhong et al., 2018), (Tang et al., 2016), (Wang et al., 2015), (Wang et al., 2023a), (Sharma et al., 2017), (Wu, 2020), (Cao, 2019), (Jiang, 2020), and (Chen et al., 2021), it can be seen from Table 4 to Table 8 that the fault diagnosis accuracy of the ensemble model integrating neural network and deep learning learners (Li et al., 2021a), (He et al., 2020), (Yang et al., 2021), (Ma and Chu, 2019), (Xu et al., 2019), (Tong et al., 2023), (Luo et al., 2016), (Wang et al., 2018a), (He et al., 2021), (Wu et al., 2021), (You et al., 2023), (Cao et al., 2022), (Zhang et al., 2021a), (Wang and Vinogradov, 2023), (Raichura et al., 2020), (Xuan et al., 2022), (Li and Lin, 2021), (Badr et al., 2023), (Wang et al., 2021c), (Kowalski et al., 2017), (Xia et al., 2019), (Zhang et al., 2020c), (Wang et al., 2023b), (Che et al., 2022), (Chowdhury et al., 2022), (Zhang et al., 2023a), (Hou et al., 2023), (Gao et al., 2019), (Zhou et al., 2021a), (Li et al., 2022), (Li et al., 2017), (Li et al., 2020), (Beretta et al., 2021), (Zhang et al., 2020a), (Pang et al., 2021) is higher. In particular, the fault diagnosis accuracy of ensemble deep learning is close or equal to

Table 6
The accuracy of fault diagnosis for gearbox with used method, input data type and classification of fault types.

References	Method	Accuracy (%)	Input data	fault types
(Zhang et al., 2021a)#	Multi-model ensemble deep learning method based on DCNN	98.75%	Vibration signals.	880 samples were collected under 10 operating working conditions, and each kind of crack severity had 220 samples.
(Pang et al., 2021)#	Cross-domain negative correlated ensemble algorithm.	94.5	Vibration signals.	Broken teeth, missing teeth, tooth sur-face wear, tooth root crack.
Kundu et al. (2020)	A random forest regression method.	N/A	Vibration signals.	Remaining useful life prediction of spur gears under natural pitting progression.
Wang et al. (2021b)	Multi-criteria feature Selection and heterogeneous ensemble learning classification.	99.62	Vibration signals.	Wear fault, crack fault, gnash fault and tooth mission fault.
Zhou et al. (2019)	An ensemble multi-fault features indexing approach.	97.5	Vibration signals.	Normal and five fault condition (No detailed description in the text).
(Wang et al., 2018a)#	PSOSEN	95.34	Vibration signals.	Normal, Servo valve fault, Hydraulic motor fault, Rotary actuator fault, Shaft fault.
Kang et al. (2022)	EnCH	98.01	Vibration signals.	Healthy, Crack of drive gear.
Cheng et al. (2016)	EEMD	91.75	Vibration signals.	A normal gear, a gear with one missing tooth, a broken gear, a gear with wear and a gear with a tooth root crack.

100%, such as (Wang et al., 2019b) and (Li et al., 2021b). This is because traditional machine learning methods need to manually select faulty features. Manual feature extraction is difficult and limited by expert experience, which can easily lead to incomplete feature extraction. The deep learning method replaces the manual feature extraction process through the complex neural network structure. This process reduces the influence of expert experience and makes feature extraction more accurate. Therefore, from the perspective of learners, ensemble deep learning is an effective way to improve the accuracy of ensemble models.

Fig. 8 (b) shows the statistics of the ensemble strategy used in the references from Table 4 to Table 8. The ensemble strategy can be mainly classified into three categories: bagging (Tang et al., 2018), (Wan et al., 2021), (Pang et al., 2021), (Ma and Chu, 2019), (Xu et al., 2019), (Luo et al., 2016), (Yu, 2016), (Zheng, 2011), (Raichura et al., 2020), (Xuan et al., 2022), (Kundu et al., 2020), (Zhou et al., 2019), (Li et al., 2021a), (Li and Lin, 2021), (Kapucu and Cubukcu, 2021), (Eskandari et al., 2020), (Zhang et al., 2018), (Wang et al., 2019b), (Zhang et al., 2020b), (Han et al., 2020), (Wang et al., 2018b), (Li et al., 2021b), (Xia et al., 2019), (He et al., 2020), (Zhang et al., 2020c), (Zhong et al., 2018), (Wang et al., 2023b), (Chowdhury et al., 2022), (Han and Jiang, 2016), (Wang et al., 2021a), (Sharma et al., 2017), (Li et al., 2020), (Cao, 2019), (Beretta et al., 2021), boosting (Han et al., 2019), (Jiang, 2020),

Table 7
The accuracy of fault diagnosis for new energy systems with used method, input data type and classification of fault types.

References	Method	Accuracy (%)	Input data	fault types
Li and Lin (2021)	Ensemble learning with diversified base models.	99.29	Personal Computer Transient Analysis simulator (PCTRAN).	Normal condition, Steady-state conditions at 100% power, turbine trip, steam pipe rupture in containment, steam pipe rupture outside containment, steam generator heat transfer pipe rupture, main steam isolation valve closed, rod lifting accidents, rod insertion accidents and steam turbine load shedding etc.
Kapucu and Cubukcu (2021)	A supervised ensemble learning method.	99.67	Electrical and climatic data.	Normal and faulty operating states (partial shading and short-circuit).
Wang (2021)	An ensemble learning algorithm optimized by a grid search algorithm.	95.2	Voltage sensors.	Short circuit, open circuit and aging fault.
Xu (2020)	The probabilistic neural networks is as a learner optimized by the sine- cosine stochastic optimization algorithm.	100	Electrical data.	Open circuit, abnormal going, blocking shadow and short circuit.
Eskandari et al. (2020)	An ensemble learning model based on the probabilistic strategy.	99.5	Current-Voltage.	109 line-line faults and 92 normal conditions.
Badr et al. (2023)	Crack fault diagnosis of rotating machine in nuclear power plant based on ensemble learning.	94.16	Vibration signals.	Health, Sun tooth crack, Sun tooth crack, Sun tooth crack, Sun tooth broken, Planet tooth crack, Planet tooth crack, Planet tooth crack, Planet tooth broken.
Zhong and Ban (2022)	Intelligent fault identification strategy of photovoltaic array based on ensemble self-training learning.	90.48	Voltage	Physical faults, environmental faults, and electric faults.
Wang et al. (2019a)	Signal processing and simultaneous-fault diagnostic method for wind turbine.	85.51	Vibration signals.	Normal, Chipped tooth, Looseness, Gear crack, Wear of cage and rolling elements of bearing, Wear of outer race of

(continued on next page)

Table 7 (*continued*)

References	Method	Accuracy (%)	Input data	fault types
				bearing, Gear tooth broken, Mechanical misalignment, Unbalance, Gear tooth broken.
Zhang et al. (2018)	Fault Detection of Wind Turbines Using Random Forests and XGBoost.	N/A	Vibration signals.	Scaling, offset, stuck, slow dynamic change, abrupt dynamic change.
Mellit et al. (2023)	Fault Diagnosis Approach for Photovoltaic Systems Based on Stacking-Based Ensemble Learning Methods.	96.21	Current-voltage (I–V)	Dust deposit on PV modules, partial shading effect, open circuit diode with dust accumulation, partial shading with dust accumulation, and shunted diode with shading.
Wang et al. (2021c)	Heterogeneous Ensemble Learning in Fault Diagnosis of Photovoltaic Modules.	97.78	Current-voltage (I–V)	Specific fault classifications can be found in the paper.

(Zhang et al., 2018), (Tang et al., 2016), (Chowdhury et al., 2022), (Gou et al., 2018), (Yang et al., 2021), (Chen et al., 2021), (Zhang et al., 2015), (Wu et al., 2021), (Zhou, 2021b), (Wang, 2015), (Wang, 2021), (Xu, 2020) and stacking (Zhang et al., 2021a), (Wu, 2020), (Cao et al., 2022), (Wang and Han, 2020), (Zhang and Wang, 2021), (Mellit et al., 2023), (Liu, 2020), (Zhou et al., 2021a). There are also some other innovative ensemble strategies. This paper classifies these innovative ensemble strategies as others (Jian and Ao, 2022), (Che et al., 2022), (Wang et al., 2021b), (Kang et al., 2022), (Cheng et al., 2016), (Badr et al., 2023), (Zhong et al., 2022), (Wang et al., 2019a), (Wang et al., 2021c), (Kowalski et al., 2017), (Zhang et al., 2023a), (Hou et al., 2023), (Li et al., 2022), (Gao et al., 2019), (Li et al., 2017), (Zhang et al., 2020a), (Tong et al., 2023), (Wang et al., 2018a), (He et al., 2021), (You et al., 2023), (Zhu, 2017).

Bagging (Breiman, 1996) trains multiple learners in parallel, and the result is a vote or weighted calculation of the results of multiple learners. The learner consists of the same or different fault diagnosis algorithms. The general framework of the ensemble fault diagnosis based on bagging is shown in Fig. 9. Since bagging trains learners individually, it is easier to integrate the results of each learner through voting and weighting, it is more widely used in fault diagnosis research than other strategies.

Fig. 10 shows the general structure of the boosting ensemble model (Schapire, 1990). The boosting algorithm obtains the result through a linear combination of multiple learners. The calculation result of each learner in the boosting algorithm is weighted with the original data as the input of the next learner. In addition, the sample with the higher recognition error rate of the previous learner has a greater weight in the next learner learning sample.

The fault diagnosis algorithm based on boosting is different from the fault diagnosis algorithm based on bagging: the bagging algorithm trains different learners by sampling different subsets of data, while in boosting the samples of each learner are the same; The bagging algorithm trains multiple learners in parallel to obtain the result through voting or meaning, while the boosting algorithm is a linear combination of multiple learners. From structural analysis, the boosting algorithm is more complex than bagging, but the boosting algorithm is simpler than bagging in data processing. The two algorithms have their own advantages, they can be selected according to actual needs.

Fig. 11 shows the general structure of the stacking ensemble model (Wolpert, 1992). The structure of the stacking model can be seen as a combination of bagging and boosting. The first-level structure of the stacking model is parallel learning similar to bagging, and the relationship between the second layer and the first-layer model is similar to boosting.

In our survey, the stacking strategy has the least application compared to bagging and boosting. An important reason may be that this strategy uses the training results of multiple learners as feature data to train the second stage learners and complicates the implementation of the strategy. Moreover, the structure of stacking is similar to a neural network, which makes the overall structure of ensemble learning more complex. This may limit its application as well. Table 9 summarizes the advantages and disadvantages of different ensemble learning strategies.

Other methods (Zhang et al., 2020a), (Tong et al., 2023), (Wang et al., 2018a), (Zhu, 2017), (Wang et al., 2021b), (Wang et al., 2021c), (Kowalski et al., 2017): Literature (Zhang et al., 2020a) proposes an ensemble strategy based on threshold setting and weight redistribution to solve the problem that each member of a voting policy has the same weight. In transformer fault diagnosis, 99.38% fault diagnosis accuracy was achieved. Literature (Zhu, 2017) utilis the combination of artificial bee colony algorithm and the ensemble neural network algorithm to achieve good results in the fault diagnosis of wind tur-bine gears. Literature (Wang et al., 2021b) uses Dezert-Smarandache theory (DSmT) (Smarandache and Dezert, 2006) to perform decision-level fusion of the output of the base classifier to improve fault diagnosis performance. Reference (Kowalski et al., 2017) uses Error-Correcting Output Codes (ECOC) (Allwein et al., 2000) combiners and one-vs-one (OVO) (Liu et al., 2008) solutions together to efficiently select the final class to predict, while providing an efficient way to reduce errors that can occur at the individual classifier level.

Fig. 12 shows the line chart of performance for the four types of ensemble methods reviewed in terms of fault diagnosis accuracy. If the number of references reviewed is ignored, bagging and stacking ensemble fault diagnosis has a greater range of accuracy, which means that their performance may be unstable. The range of accuracy for boosting and other ensemble methods is relatively small, which means that their performance may be relatively stable. Clearly, there are relatively few studies involved in the comparison, this conclusion does not necessarily apply at any time or space. In fact, if some small samples (outliers (Grubbs, 1969)) are ignored, there is no excessive gap in the stability of these ensemble methods.

In summary, ensemble method can achieve better performance than a single model. However, given constraints such as learners, learning environments, and diagnostic scenarios, it remains a challenge to find optimal ensemble strategies. These challenges are little discussed in ensemble learning based fault diagnosis and should serve as meaningful future work.

4. Challenges and directions

The core goal of ensemble learning based fault diagnosis is to improve the accuracy of fault diagnosis. The previous overview and discussion of this paper show that the fault diagnosis method based on ensemble learning can obtain higher diagnostic accuracy than a single learner. However, the application research of fault diagnosis based on ensemble learning is still in the development stage and still faces many challenges and problems. This section will mainly discuss some of the challenges and difficulties that hinder the application of fault diagnosis methods from a technical perspective. In addition, some potential research directions for ensemble learning based fault diagnosis are discussed.

Table 8
The accuracy of fault diagnosis for different equipment with used method, input data type and classification of fault types.

References	Equipment	Method	Accuracy (%)	Input data	fault types
Wang et al. (2019b)	Aeroengines	Integrated with BLS and ensemble learning model.	99.88	Oil data.	Wear fault diagnosis of aeroengines.
Kowalski et al. (2017)	Marine 4-stroke diesel engines	Using a one-vs-one Extreme learning ensemble.	98	Exhaust gas from marine 4- stroke diesel engines.	Normal condition, the throttling of the exhaust gas duct, the throttling of the air inlet duct, the shift of the fuel pump cam on the camshaft, which causes a delay in the fuel injection, the leakage of the air inlet valve, the leakage of the exhaust gas valve, the decrease in the opening pressure of the fuel injector, the increase in the opening pressure of the fuel injector, the chocked fuel injector, the discalibrated fuel injector, the leakage of the fuel injection pump.
Zhang et al. (2020b)	Refrigeration System	Ensemble learning Of multiple models.	99.58	Fault simulation data of centrifugal chiller.	Normal condition, insufficient refrigerant, scaling of the condenser, insufficient water flow on the condenser side, non-condensable gas mixed in the refrigerant, insufficient water flow on the evaporator side, excessive refrigerant and excessive lubricating oil.
Han et al. (2020)	Building energy system	Ensemble learning With member optimization.	99.88	Fault simulation data of centrifugal chiller.	
Wang et al. (2018b)	Building Energy	EBT	N/A	Meteorological, occupancy, and temporal.	Out-of-bag errors including Temp, Dew, Hum, Press, Wind, Solar, Occ, Time etc.
Liu (2020)	Railway Switch	Stacking algorithm.	97.83	Current signals of rail-way switch.	Normal condition, line disconnection, external line short circuit, relay failure, turnout switch turn short circuit, switch close contact force is too large, not unlocked, electronic rotor disconnection, motor rotor disconnection, switch not locked, automatic switch is not flexible, the extension and close contact with the turnout are not well pasted, and the conversion resistance increases.
Li et al. (2021b)	Hydraulic pump	Ensemble empirical mode decomposition and wavelet kernel extreme learning machine methods.	100	Vibration signals of a hydraulic pump.	Single slipper wear, single slipper loosing and center spring wear type.
Xia et al. (2019)	IGBT open circuit	A data-driven method based on hybrid ensemble learning and sliding-window classification.	99	Three-phase current signals of insulated gate bipolar transistor (IGBT).	Classic three-phase full-bridge circuit, which consists of 6 insulated gate bipolar transistor (IGBT) with corresponding antiparallel connected diodes. There are 6 and 15 types, respectively for single IGBT and double IGBTs open-circuit fault. Considering both normal working conditions, there are 22 labels totally.
Zhang et al. (2020c)	Reciprocating compressor	An ensemble Empirical mode decomposition- convolutional deep belief network.	91.89	Displacement signals, velocity signals, vibrati on signals	Four typical faults of the reciprocating compressor, which are gas valve failure, piston breaking, cylinder scraping and bearing shell wearing.
Zhong et al. (2018)	Car engines	PCSBELM	N/A	Oxygen sensor, ignition pickup, sound.	High resistance in secondary circuit, Retarded ignition timing, Misfire due to extremely lean mixture, partially broken spark-plug cable, Narrow spark plug gap, Rich mixture, Carbon fouled in spark plug, Engine knock, Wide spark plug gap, Lean mixture.
Tang et al. (2016)	Ball mills	Ensemble learning approach to support building energy use prediction.	N/A	Vibration and sound signals.	Detailed data are available from the literature.
Chowdhury et al. (2022)	3D printers	3-D Printer fault diagnosis With Explainable Ensemble AI.	99.75	Accelerometer.	Healthy condition, bed failure, and arm failure.
Zhang et al. (2023a)	Variable refrigerant flow (VRF) systems	VRF systems based on stacking ensemble learning.	95.55	Pressure and temperature.	10 labels, detailed data can be consulted from the literature.
Hou et al. (2023)	Centrifugal compressors	ISDD	95.8	Vibration, pressure sensor.	Surge, Normal condition.
Gao et al. (2019)	High-Voltage Circuit Breaker	IELM	98.61	Piezoelectric accelerometer for vibration signal (LC0102T model).	The normal state, the base-loosen state, the C-phase leading-closing state, and the C-phase hysteresis-closing state
Zhou et al. (2021a)	Sensor	A Text-Driven Aircraft Fault Diagnosis Model.	97.35	Real-world aircraft failure text dataset.	Sensor faults, equipment aging, equipment ablation, human error, circuit faults or mechanical faults.
Wang et al. (2015)	Blast furnace faults and the Tennessee Eastman process (TEP)	A Process Industry Fault Diagnosis Algorithm Based on Ensemble Improved Binary-Tree SVM.	90.2	Cross temperature at blast furnace top, pressure in blast furnace, material feeding velocity, CO2 percentage, venting quality, volume of blast, blast pressure and temperature, etc.	Ten kinds of states: one normal and nine kinds of abnormal states, such as pipe fracture, cooler, warmer, slip, low stockline, brim gas flow, center gas flow, hanging, moving handicap.

4.1. Imbalanced data processing

Imbalanced data processing is an important challenge for fault classification. Imbalanced data generally refers to the imbalance of data categories. Data imbalance in the field of fault diagnosis is manifested by the fact that the number of samples of system health data is much larger than that of fault data. Ensemble learning models and non-ensemble learning models usually assume that the data distribution is balanced, which easily leads to model learning results tending to multi-sample categories and ignoring few-sample categories (Kaur et al., 2019), (Patel et al., 2020).

The most commonly used imbalance processing methods are sampling (He and Garcia, 2009), (Guo et al., 2017) and data augmentation (He et al., 2008). In addition, there is an extreme state of imbalanced data (Yang et al., 2022), known as the long tail distribution (Anderson, 2008). Therefore, the long-tail problem tacking methods can be used to solve the data imbalance problem. Common long-tail problem solving methods include classical methods (Chen et al., 2022a), hard negative mining (Lin et al., 2020), meta-learning (Wang et al., 2017), deep learning methods (Zhang et al., 2023b) and contrastive self-supervised

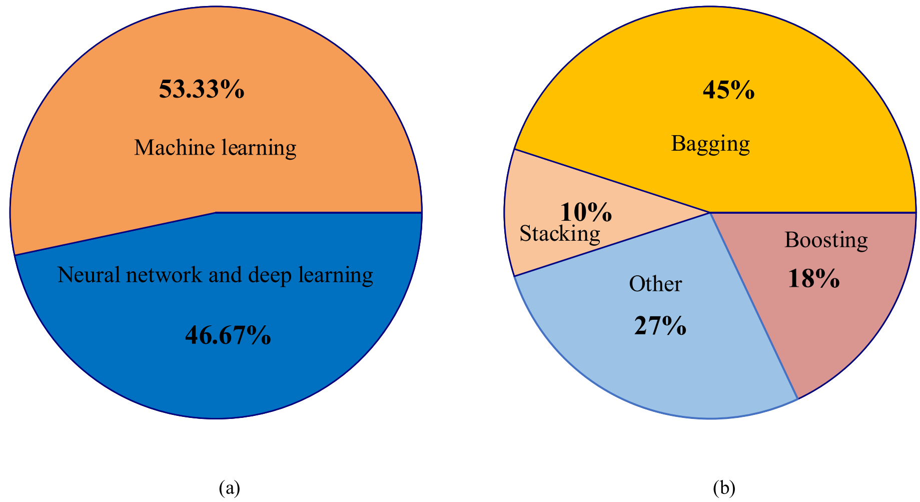


Fig. 8. Statistics of (a) the learner types, and (b) the ensemble strategies used in the ensemble learning based fault diagnosis methods.

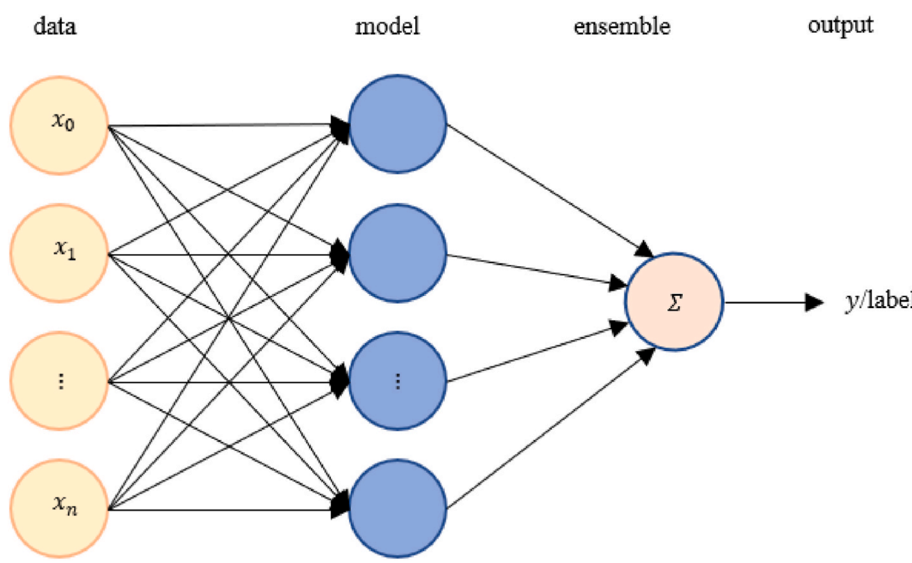


Fig. 9. The general framework of the ensemble fault diagnosis algorithm based on bagging.

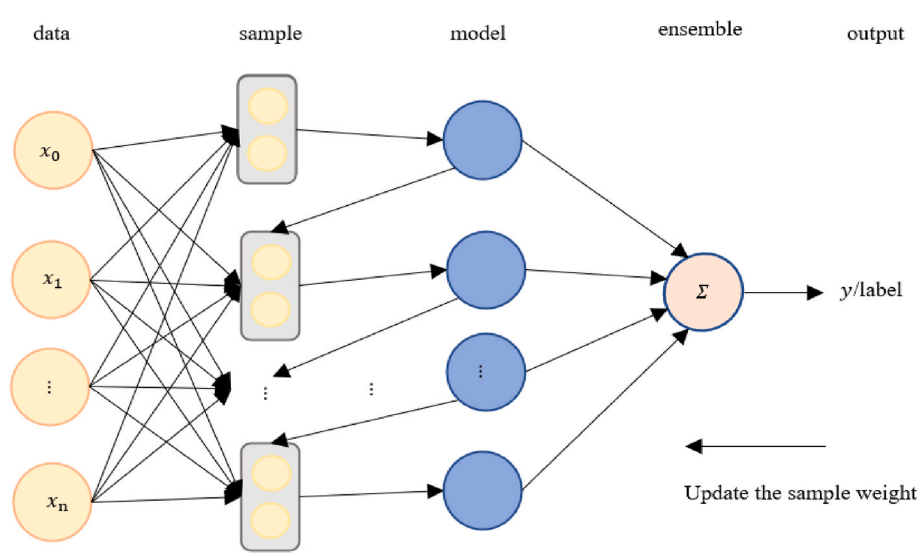


Fig. 10. The general framework of the ensemble fault diagnosis algorithm based on boosting.

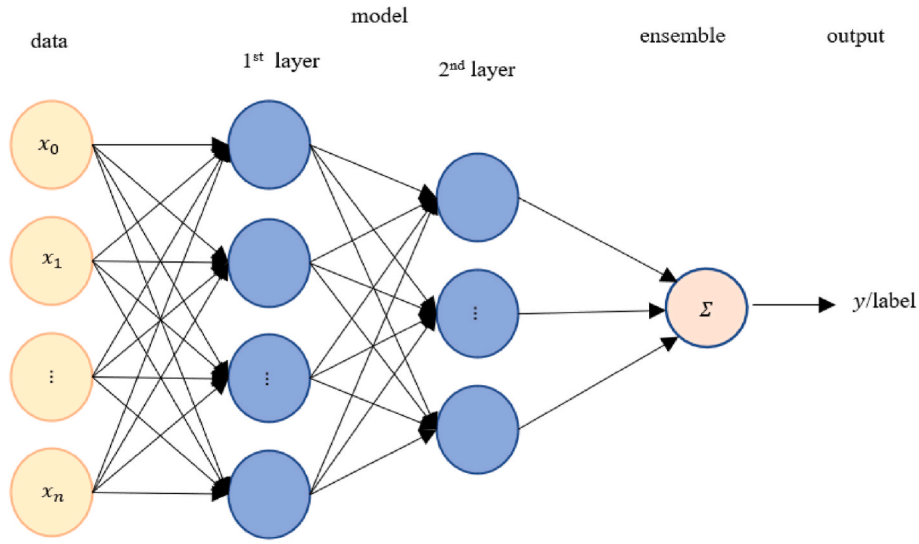


Fig. 11. The general framework of the ensemble fault diagnosis algorithm based on stacking.

learning (Grill et al., 2020). Although data imbalance has been studied in many fields, there is no quantitative and standard method for defining data imbalance. It is necessary to construct methods for defining thresholds of balance and imbalance (Ren et al., 2022).

4.2. The concept drift (CD) of fault diagnosis

Most of the existing diagnostic methods assume that the data distribution is static, and that all data is available during the training, while in real applications, the data become available as data streams (Hosseinpoor et al., 2020). Concept drift (CD) (Sato et al., 2021) means that the target variable i.e., the fault data in the data stream change over time in unforeseen ways. Concept drift could easily lead to performance degradation of fault detection and diagnosis (FDD) systems (Ardakani et al., 2018). Therefore, solving concept drift is one of the key tasks to ensure accuracy.

Supervised learning has been a common method for solving conceptual drift problems in fault diagnostic research (Hosseinpoor et al., 2020), (Ardakani et al., 2018), (Zenisek et al., 2019) but complex manual labeling of samples is required (Hosseinpoor et al., 2020). The cost of obtaining true labels could be expensive. Screening out instructive significance samples is still a challenge and a difficulty in concept drift research. In addition, adaptive models and ensemble techniques have played an increasingly important role in recent concept drift adaptation developments (Lu et al., 2019). Adaptive learning models alleviate the conceptual drift problem by automatically correcting the structure or parameters of the model. The lack of adaptive model structure or parameter correction is then compensated by the group decision making of the ensemble learning model. This finally leads to a relatively good adaptive ensemble model.

4.3. Cross system model and generalization

In this paper, the cross-system models and the generalization of fault diagnosis models are considered as close concepts. Generalization means the ability to accurately infer new unseen and in-distribution samples (Wang et al., 2020). Generalization of fault diagnosis refers to the predictive ability of the fault diagnosis algorithm model to the unknown system datasets (training, test, and validation datasets). Verifying the diagnostic accuracy of the model through unknown system data can not only prove the cross-system generalization of the model, but also reduce the scientific research cost of designing different model methods for different devices.

Some studies have investigated cross-system generalizations. Zhang et al. (2021a) propose a multi-model ensemble deep learning method based on deep convolutional neural network (DCNN) to accomplish fault recognition of high dimensional samples. The proposed method, which is validated using both bearing and gearbox dataset, has high recognition accuracy. Li et al. (2020) propose an improved domain adaptive network to obtain an effective fault diagnosis model from data collected from different devices or under different working conditions. They regard the problem of detecting the failure of different equipment as an important direction of expansion. To solve the problem of performance deterioration of deep diagnostic models due to changes in working conditions, Pang et al. (2021) proposed a new ensemble algorithm that employs deep convolutional extreme learning machine (DCELM) as base learners. Different equipment datasets are used for the analysis including gear, bearing, and rotor datasets. On average, its accuracy is 1.9% higher than other ensemble methods, and it achieves 3.6% enhancement compared with its base learner.

In summary, the fault data for different equipment and working conditions will affect the performance of the fault diagnosis algorithm. Therefore, it is useful to study cross system research of fault diagnosis algorithms. The reviewed work such as references (Zhang et al., 2021a), (Li et al., 2020), and (Pang et al., 2021) have demonstrated that ensemble learning based fault diagnosis methods could contribute to cross-system research. However, it is not clear why those ensemble methods help these algorithms in cross system diagnostic capabilities. It is valuable to study whether the ensemble strategies are useful in improving cross-system capability.

4.4. Fault diagnosis under speed variation

Fault diagnosis of rotating machines is crucial for their reliable operation and continuous availability (Randall, 2010). However, most diagnostic methods use a constant speed regime. In practice, almost all industrial machines experience different levels of speed variations during operation (Sun et al., 2010).

Some studies have discussed fault diagnosis at variable speeds. Han et al. (2021) proposed a deep learning-based fault diagnosis method for dealing with the speed fluctuation problem, aiming at the problem that existing methods have shortcomings in computational efficiency and diagnostic errors. Their experiments show that the proposed method can solve the influence of speed fluctuation and achieve accurate identification of different fault types. It also obtains a higher accuracy than other methods. Chen et al. (2022b) proposed a Multi-expert Attention

Table 9
The differences, advantages and disadvantages of different ensemble learning strategies.

	Bagging	Boosting	Stacking
Structural complexity	The learner learns independently and outputs the result by voting or averaging.	Boosting constructs multiple base models sequentially, with each model relying on the performance of the previous one.	Stacking combines multiple layers of models including base models and a meta-model. It allows freedom in selecting models for each layer.
Learner	It is suitable for learners with low bias.	It can combine weak models into a strong model.	model combinations can be freely selected.
Parallel computing	Each learner can be generated in parallel.	In theory, it can only be produced sequentially, because each model needs the results of the previous model.	Learn weak learners in parallel and then combine them by training a metamodel.
The extent of wide application	The literature surveyed in this paper shows that the bagging ensemble method is the most widely used, up to 45%.	The boosting ensemble method is less commonly used than bagging, accounting for 18% of the total.	The stacking ensemble method is the least used, accounting for only 10% of the total.
Training data	Samples are randomly selected from the original dataset.	Assign a weight to each data sample and adjust the selection of data samples in each iteration based on the distribution of weights.	No data sampling is involved. All training data is used for each learner training.
Advantages	The variance of the model is reduced by averaging the predictions of multiple learners, which enhances the stability and generalization ability of the model. Training each learner individually reduces the risk of overfitting and is well suited for parallel training, making the learner highly adaptable to large-scale data.	It can significantly improve model performance and reduce model bias. Multiple weak models can be combined into a strong model, so that there are no strict requirements for the selection of the base model. It can handle imbalanced data sets and is more friendly to minority class sample data.	The hierarchy of model combinations and stacking can be freely chosen according to the needs of the problem. Stacking can provide learner predictions, thereby enhancing the interpretability of model results.
Disadvantages	The goal of bagging is to reduce the variance of the model, thus models with high bias are not suitable for bagging. Bagging involves random sampling of data and is not highly tolerant to outliers.	Boosting assigns higher weights to incorrectly predicted samples, which can lead to overfitting. It is also sensitive to outliers and noise. Since the sample weights need to be adjusted, the training time of the model will be relatively long.	The complexity of stacking is relatively high because it requires training multiple learners and meta-models, resulting in high computational costs. Stacking also demands larger datasets, and its performance may not be optimal with small sample sizes.

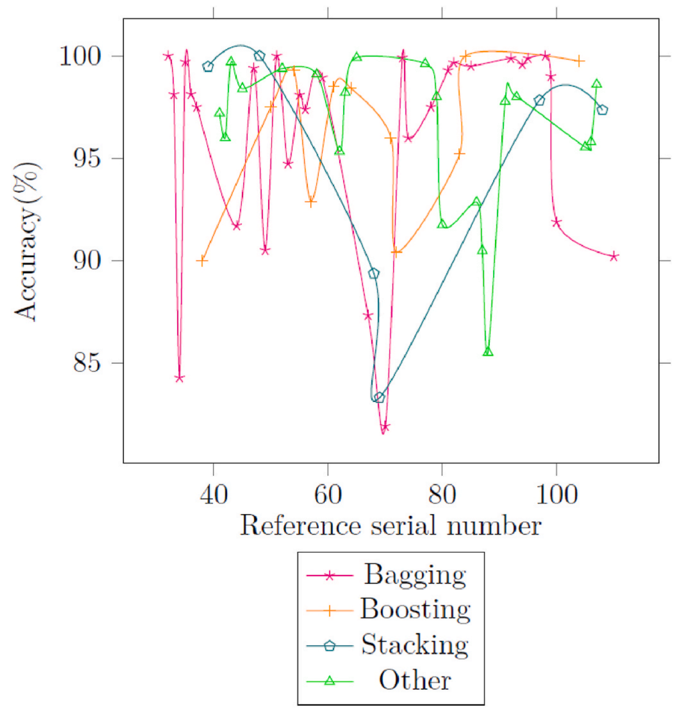


Fig. 12. The line chart of performance for the four types of ensemble methods reviewed in terms of fault diagnosis accuracy.

Network with Unsupervised Aggregation (UA-MAN) to solve the problem that the fault signal is difficult to obtain, leading to the long-tailed distribution of data and domain shift caused by speed variation, which further deteriorates the reliability of the model. Their experimental results show that the recognition accuracy of the proposed method under different imbalance class distributions exceeds 98%, which has a good inhibitory effect on the domain shift caused by speed variation.

4.5. Remaining useful life (RUL) prediction

System RUL prediction is a popular topic in the field of fault diagnosis. Systematic RUL prediction can be divided into physical model-based prediction method, data-driven prediction method and hybrid model method. Physical model-based prediction methods require the development of a physical model that fully describes the degradation process and the set of phenomena of a system or subsystem (Saidi and Benbouzid, 2021). They use statistical physics or stochastic processes (Leao et al., 2008; Batzel and Swanson, 2009; Saidi et al., 2017). They are affected by the complexity of the system and are also difficult to implement (Yi et al., 2022). In contrast, data-driven methods are not affected by system complexity but require models to be trained on historical data (Hu et al., 2019). Data-driven RUL prediction is currently the most studied and mature method (Ren et al., 2020; Liu et al., 2022a; Wang et al., 2022a; Zhang et al., 2021b; Muneer et al., 2021). A comprehensive review of various data-driven algorithms has been carried out by Nam-Ho et al. in (Kim et al., 2017). Due to the strong dependence of data-driven RUL prediction on historical data, the difficulty of obtaining high precision historical data has become a technical problem that needs to be solved in the process of achieving accurate evaluation.

The hybrid model is the fusion of the two methods (physical model-based prediction method, data-driven prediction method), which reduces the data dependence to a certain extent. Deng et al. (2020) proposed a hybrid GRU-PF methods integrating the data-driven model and the physical model into the particle filter (PF) network to achieve the remaining useful life prediction of the ball screws. Experiments show that this method is more sensitive to fault problems.

The ensemble learning is a group decision-making behavior, which similar to the fusion model. Some studies discussed RUL prediction based on ensemble learning and demonstrated that ensemble learning can improve the accuracy of RUL prediction. Kundu et al. (2020) proposed a random forest regression method based on ensemble decision tree to predict the remaining useful life of spur gears in pitting failure mode. Gungor et al. (2021) proposed an optimal weighted ensemble learner for RUL prediction. It has good adaptability to different data sets and underlying system parameters. Yao et al. (2019) proposed a new deep learning method for RUL estimation by using temporal empirical Mode Decomposition (EMD) and convolutional neural Network (CNN). Ture et al. (2023) used deep learning algorithm to develop a prediction model for estimating the remaining useful life of the turbofan engine on the NASA turbofan engine degradation simulation data set. They proved that the stacked ensemble learning method could obtain an accuracy of 95.72%. Li et al. (2019) proposed an ensemble learning algorithm to predict the RUL of aircraft engines, which was shown to be robust and superior to other prediction methods in the literature. Yang et al. (2016) proposed a hybrid prediction approach that can predict the RUL of degraded Li-ion batteries using both physical laws and data-driven modeling. Xia et al. (2020) proposed an ensemble learning framework based on multi-time window convolutional bidirectional long short-term memory to accurately predict RUL in the case of length inconsistency of condition monitoring data. Zeng and Cheng (2020) introduced a RUL prediction method for aircraft turbine engines based on ensemble learning and Euclidean distance weighting. The weight of each member algorithm is assigned based on the Euclidean distance between the RUL predicted by each member algorithm and the true RUL calculated from the training dataset. Wang et al. (2022b) proposed a long short-term memory neural network with transfer learning and ensemble learning for RUL prediction. The model demonstrates that it has good performance on small sample data sets. Cheng et al. (2020) proposed an ensemble long short-term memory neural network model. The experimental results show that it has improved the generalization of RUL prediction methods. Zhang et al. (2016) proposed a multi-objective deep belief network ensemble method for RUL estimation with superior performance. The similarities between ensemble learning and hybrid models imply that using ensemble learning methods may reduce data dependence.

4.6. Ensemble deep learning

In recent years, the ensemble learning based fault diagnosis learners have widely used neural networks and deep learning models. References (Zhang et al., 2021a), (Jiang, 2020), (Ma and Chu, 2019), and (Xu et al., 2019) discussed the combination of deep learning and ensemble learning in fault diagnosis methods. The literature shows that the fault diagnosis accuracy of the ensemble deep learning model is higher than that of the deep learning model (under the condition of the same model depth). However, these methods only study the application of shallow depth models with fewer network layers in fault diagnosis. These shallow models are relatively weak in generalization and accuracy. This implies that models with stronger generalization ability and deeper network level would be promising research for ensemble learning based fault diagnosis.

4.7. AI and security

Data security has consistently been one of the significant challenges facing AI. Although traditional privacy protection schemes can effectively protect data security, with the advent of quantum computers, traditional public key cryptography (PKC) algorithms face significant challenges in security (Sanal et al., 2021). One potential solution is to use post-quantum cryptography (PQC) (Bernstein, 2009). If PQC tends to replace traditional encryption techniques, every AI-based security application including fault diagnosis will be affected. Furthermore, PQC can enhance security and reliability in the field of fault diagnosis, but this does not mean that PQC is completely secure and invulnerable. It also faces threats of malicious attacks. PQC and its threats and attacks are one of the topics worth discussing in the field of artificial intelligence. When looking to implement a secure and private intelligent model, understanding and preventing potential attacks is critical. These attacks include, e.g., injecting malicious or unexpected faults into the system encryption process, or even maliciously targeting redundant information within the system. This can decrease system reliability, resulting in potential economic losses (Ali et al., 2016).

To tackle this problem, several fault detection schemes based on the Advanced Encryption Standard (AES) have been proposed (Daemen and Rijmen, 2001). Moreover, the Sub Bytes (S-boxes) is an important hardware architecture in AES. To prevent potential damage due to internal faults or intrusions by attackers, Kermani and Masoleh (2011) presented a high-speed architecture for the S-boxes constructed using mixed bases to counteract these internal or malicious faults. They conducted ASIC synthesis using a 65-nm CMOS standard technology for the proposed concurrent fault detection architectures and their counterparts. Compared to approaches with similar error coverage, the proposed approach is the most efficient one, achieving an efficiency of 5.02 Mbps/μm^2 while maintaining a throughput of 5 Gbps. Jalali et al. (2017) proposed a method to enhance the super singular isogeny Diffie-Hellman (SIDH) key exchange protocol. This method significantly reduced the size of keys and made the method more suitable for bandwidth-constrained communication applications. Additionally, Niasar et al. (2021) studied the key exchange mechanism of the algebraic lattice cipher suite Kyber. This provides a viable research direction for lattice based PQC. However, there is still ample room for improvement in terms of resisting side-channel attacks (SCA) and devising strategies to counter such attacks. Canto et al. (2023a) discussed the security issues of the PQC. Similarly, they pointed out that the PQC algorithm was vulnerable to SCA attacks. Therefore, developing effective SCA defense strategies is crucial and necessary for safeguarding emerging PQC systems. Sarker et al. (2020) studied the error detection schemes for two phases of ring learning with errors encryption operations and implemented them on application-specific integrated circuits (ASICs). Their proposed architecture can be customized based on constraints from different usage models and applied to PQC schemes.

Dubrova et al. (2023) presented a method for deep learning-based message recovery attacks on the 5-order masked implementations of CRYSTALS-Kyber in ARM Cortex-M4 CPU. This method allows them to train neural networks that can recover a message bit with the probability above 99% from high-order masked implementations. Berzati et al. (2023) presented a new profiling side-channel attack on CRYSTALS-Dilithium. They designed a template attack and demonstrated that it could recover part of the secret key that is sufficient to produce universal forgeries. They argued that the use of expensive masking for protection remains a matter of debate.

An important dimension is that security is expected to be built in low-energy contexts. Kaur et al. (2023) conducted a comprehensive discussion on the development of lightweight cryptography standards. They placed special emphasis on the potential application of ASCON in low-cost fault diagnosis. Canto et al. (2023b) introduced a novel approach that utilizes the GPT-4 model to implement the NIST Lightweight Cryptography (LWC) standard named ASCON. The process of implementing ASCON using ChatGPT, along with practical Python implementations, holds significant value in understanding the application of advanced AI language models in the field of cryptography. Aghaie et al. (2016) proposed a fault diagnosis method based on a lightweight block cipher called Midori. This method not only enhances the reliability of the new, energy-efficient lightweight block cipher but also can be customized based on reliability- and cost-tolerance objectives. Sarmadi et al. (2013) addressed the limitations of traditional cryptographic algorithms in nodes with constrained finite field arithmetic operations by proposing a low-complexity super-serial multiplier structure. Compared to conventional serial multipliers, the hardware complexity of the proposed super-serial multiplier is significantly lower. Koziel et al. (2015) presented a co-processor designed for area-constrained devices, which leverages state-of-the-art mixed-point addition and doubling on binary Edwards curve equations. Compared with existing research, the proposed co-processor reduces the required clock cycles for point multiplication by approximately 50% under similar silicon area constraints. Jalali et al. (2019) designed a set of constant-time and highly optimized field and group algorithm implementations based on ARM components. They also provided secure CSIDH software to defend against Simple Power Analysis (SPA) and Differential Power Analysis (DPA) attacks. Benefiting from CSIDH's fast key verification and small key size, this solution is suitable for a wide range of application scenarios that require static keys and limited bandwidth.

5. Lessons learned from the surveyed papers

The survey has identified the following areas as important to consider in ensemble learning.

1. Ensemble learning based fault diagnosis methods can effectively improve the diagnosis accuracy. The summary shows that using neural networks and deep learning-based learners can achieve higher diagnostic accuracy. In addition, considering multiple evaluation indicators makes the ensemble learning models perform better in terms of fault diagnosis accuracy and applicability. The most used evaluation metric in summarized fault diagnosis research is accuracy. The accuracy is obtained by using the confusion matrix shown in Table 10 and the evaluation index shown in Table 11 (Ben and Zhang, 2006). Table 11 also shows some other commonly used evaluation metrics including precision, recall and F-score. For general classification problems, some of these evaluation indexes may be used to evaluate the performance of a method. However, by using only accuracy as a metric may not give robust estimations. For example, in fault diagnosis domain the fault data is usually extremally unbalanced. This means that the amount of fault data will be much less than that of the normal data. In this situation, measuring the performance of a model with accuracy alone is meaningless. Because there are relatively few positive examples, no matter how inaccurate the diagnosis is, the accuracy value is very high. Therefore, one need to consider other more stable and effective evaluation methods, such as by using Receiver Operating Characteristic (ROC) and Area Under the Curve (AUC) (Fawcett, 2006).

2. Among ensemble strategies used, bagging is the most widely used, possibly because it trains learners individually and it is easier to integrate the results of each learner through voting and weighting. But this does not mean that it is the optimal integration strategy. Given constraints such as learners, learning environments, and diagnostic scenarios, it remains a challenge to find optimal ensemble strategies. The soft sensor method based on ensemble Empirical Mode Decomposition (EEMD) and SEN provided in literature (Tang et al., 2016) can be taken as a reference. Another method could be using automatic machine learning (Kowalski et al., 2017; Liu et al., 2022b), where the ensemble model can automatically generate the ensemble strategy and select the appropriate learner for ensemble. The ensemble strategy obtained, however, may be uninterpretable i.e., the details of the ensemble strategy obtained may not be proved, explained, and verified.
3. Note that the experimental environments of surveyed models including methods, test dataset, input data and fault types are different. Even if the same diagnostic methods and equipment is used, there may be differences in fault types and verification data. This means that the performance of the fault diagnosis models is not easily comparable. From the perspective of practical application, however, listing the methods with its fault diagnosis accuracy, fault data objects and failure types will be valuable and useful since they can be used at least as a reference to potential researcher and users to judge and select appropriate models based on equipment or dataset and input data similarities and their expected diagnostic accuracy.
4. A big challenge in fault diagnosis research is the lack of high-quality public datasets. The lack of diversity in fault data is a challenge to both the breadth and generality of fault diagnosis research. Therefore, fault diagnosis research still requires enriched datasets. Table 1 shows some public datasets collated. However, it has been found that these data sets are partial and not complete. In addition, most fault diagnosis is based on a type of data, e.g., images (Xuan et al., 2022) or texts (Zhou et al., 2021a). For a fault, if one can collect various data or signals implying the fault, such as images, texts, or vibration signals for a comprehensive diagnosis, then the diagnosis results will be mutually confirmed, and the diagnostic performance could also be improved. There is still a lack of data sets, regardless of the richness of data or the diversity of data types. The collation of fault data sets is still a key work in the field of fault diagnosis.
5. The quality of data (labels) is vital for concept drift and data imbalance problems. Unsupervised learning, semi-supervised learning, and comparative self-supervised learning model (Grill et al., 2020) can effectively reduce the impact of labeled data quality. Fault diagnosis based on semi-supervised learning, unsupervised learning and comparative self-supervised learning model would be a meaningful research direction to address data label quality.
6. Traditional fault diagnosis research usually has a single application of fault diagnosis model and fault data, resulting in a great waste of resources. Cross-system research can be used for studying general fault diagnosis models, and it is also an effective way to make full use

Table 11
Evaluation index of accuracy, precision, recall and F-score (Ben and Zhang, 2006).

Expression	Meaning
$Accuracy = \frac{TP+TN}{TP+TN+FP+FN}$	The proportion of correctly classified samples over the total number of samples.
$Precision = \frac{TP}{TP+FP}$	The proportion of samples that were predicted to be positive that were actually positive.
$Recall = \frac{TP}{TP+FN}$	The prediction result is the proportion of the actual number of positive samples in the positive samples to the positive samples in the whole sample.
$F-score = \frac{(\alpha^2+1)*Precision*Recall}{\alpha^2*(Precision+Recall)}$	A weighted average of precision and recall. When $\alpha = 1$, it is F1, the most common evaluation metric for binary classification.

Table 10
The confusion matrix (Ben and Zhang, 2006).

Confusion		Predicting labels	
		Negative	Positive
Real label	Negative	TN (True Negative): A negative example that was correctly predicted.	FP (False Positive): A positive example that was incorrectly predicted.
	Positive	FN (False Negative): A negative example that was incorrectly predicted.	TP (True Positive): The number of positive examples that were correctly predicted.

of research and development resources for fault diagnosis models. Similarly, one could use the ensemble fault diagnosis framework proposed in (Wang et al., 2019a) for simultaneous and coupled failures. This framework has an accuracy rate of 85.51% in the concurrent multi-fault diagnosis of wind turbines.

7. Literatures (Zhu, 2017), (Gungor et al., 2021), (Yao et al., 2019), (Ture et al., 2023), (Yang et al., 2016), (Xia et al., 2020), (Zeng and Cheng, 2020), (Wang et al., 2022b), (Cheng et al., 2020), (Zhang et al., 2016), (Ben and Zhang, 2006), (Fawcett, 2006), (Liu et al., 2022b), and (Gou et al., 2020) studied the use of ensemble learning for RUL prediction. These studies have demonstrated the effectiveness of using ensemble learning modes in improving the accuracy of RUL prediction. Ensemble learning can not only improve the accuracy of RUL prediction, but also could ease the model's dependence on data. In the traditional RUL prediction method, the model has a strong dependence on data (Kim et al., 2017). The hybrid RUL prediction model (Deng et al., 2020) has been shown capable to reduce the dependence on data. Ensemble learning can fuse models and adjust the weights between different models like hybrid models. This similarity between them implies that using ensemble learning methods would help reduce data dependency in RUL prediction methods.

6. Conclusion and future work

This paper reviews the latest research on ensemble learning based on 87 journals and 209 papers in the web of science and other academic resources. It summarizes 78 different ensemble learning based fault diagnosis methods, involving 18 public datasets and more than 20 different equipment systems. The paper particularly discusses the use of ensemble learning based methods for fault diagnosis from both technical and field application perspectives.

This paper first summarizes the application of ensemble learning based fault diagnosis methods in some critical equipment including bearings, gears, transformers, and new energy systems. Although this paper focuses on the fault diagnosis of these devices, this does not mean that the scope of application of ensemble learning is limited. The fault diagnosis method based on ensemble learning could be widely used in fault diagnosis of other systems and equipment. The performance of those ensemble learning methods is discussed from the perspective of multiple evaluation indicators including accuracy, generalization, and imbalanced data processing capabilities. The accuracy of ensemble learning based models shows that ensemble learning is feasible and effective in fault diagnosis. The performance of the ensemble model is generally better than that of the single learner. The paper then, discusses the core technologies used in ensemble learning, including learners and ensemble strategies, from a technical point of view. In addition, this paper discusses some open issues and research challenges for ensemble learning based fault diagnosis. These open issues include imbalanced data processing, concept drift, cross-system model and generalization, speed variation, remaining useful life prediction, and ensemble deep learning models.

The key findings and lessons learned from the surveyed papers are summarized. An important finding is that most fault diagnosis uses a single accuracy index to measure the performance of the model. This is flawed, especially in the field of fault diagnosis. When the fault data and normal data are extremely unbalanced, the accuracy is of little significance. Therefore, a comprehensive evaluation criterion including recall, F-score, ROC and AUC are necessary. In addition, this paper attempts to identify the best fault diagnosis models through a review. However, since each model uses different data sets, fault types, and evaluation indicators, it is difficult to make an objective and effective evaluation of various models. This implies that finding an optimal model is difficult. The study of cross-system models could help match diagnostic models to the best applicable systems. It is also helpful to study the generality of the fault diagnosis models. It would be valuable if one could obtain a comprehensive system diagnostic model with optimal performance by integrating the optimal models for different components or subsystems within that system. This also implies that dynamic fault diagnosis techniques would be needed for different equipment under different operation environments.

Future work will consider:

1. Comparing the performance of different ensemble strategies under specific application environments. This paper attempts to identify the best fault diagnosis models through a review. However, since each model uses different data sets, fault types, and evaluation indicators, it is difficult to make an objective and effective evaluation of various models. This implies that finding an optimal model is difficult. But this does not mean that one must lower the requirements for performance evaluation. More performance evaluation indexes related to the model, including e.g., F-score, AUC, should be carried out and evaluated in the future. In addition, the performance of ensemble learning could be better than that of non-ensemble learning models, but this does not mean that the more integrated the learner, the better the performance. Future work will also consider comparing the performance between ensemble learning and non-ensemble learning models.
2. Although there have been some discussions about imbalanced data related to fault diagnosis, there is still a lack of quantitative and standard method to define data imbalance (Ren et al., 2022). A valuable future work could be defining criteria for quantitatively representing and measuring data imbalance. For this work, it is necessary to construct methods for defining the threshold of balance and imbalance.
3. Explore research on transfer learning in fault diagnosis. Knowledge transfer can reduce reliance on labeled data. It can also improve the generalization ability of models built on data from different distributions. The data dependence problem discussed in RUL (Kim et al., 2017) can also be alleviated by applying transfer learning. In addition, knowledge transfer can improve the generalization ability of the model. It could be used as an effective method for general faults and concurrent fault diagnosis models.
4. Multiple client fault diagnosis issues. In many studies the data is essentially independent and identically distributed data from the same client. In real industrial scenarios, there will be many different clients containing differently distributed data. If different fault diagnosis models are developed for different clients, a significant number of resources will be required. Federated learning (Zhang and Li, 2022) can provide a reliable solution to this problem. In (Zhang and Li, 2022), a fault diagnosis method combining transfer learning and federated learning is proposed, which not only solves the multiple client fault problem but also addresses data distribution differences. The presence of multiple client faults and non-identically distributed fault data is a real challenge in the field of fault diagnosis. These studies are important to the development of fault diagnosis. Federated learning is one of the methods, and other effective methods can be explored in the future.
5. Investigating ensemble adaptive models. Fault data in a real production environment is dynamic and changeable. Dynamically changing fault data leads to concept drift, which becomes a challenge in fault diagnosis research. Adaptive models can independently adjust model parameters and structure, alleviating the concept drift problem. Ensemble learning can compensate for the shortcomings of single adaptive model structure and parameter correction through multiple adaptive models. Similarly, automatic machine learning (AutoML or AML) is one of the most promising research areas in machine learning (Krzywanski et al., 2023). AML enables high-quality transfer learning through

automatic algorithm selection, hyperparameter tuning, and model optimization (Liu et al., 2022b). Automatic machine learning has been shown to cope with variable data patterns. Other study e.g. (Sosnowski et al., 2021), has discussed method for modeling of complex systems. Modeling also facilitates automation.

6. Investigating ensemble deep learning models. Shallow models are relatively weak in generalization and accuracy. If the network structure can be optimized and increased, the fault diagnosis of the ensemble deep learning model could be further improved. However, with the increase of the depth of the network, the network is often accompanied by the problem of network degradation. A solution to such problem is use the emergence of deep residual network (He et al., 2016). The research of residual network with ensemble learning based fault diagnosis is relatively few. Therefore, discussing the ensemble deep residual network can be a valuable research direction in terms of ensemble deep learning. In addition, some novel network structures in deep learning are rarely discussed in fault diagnosis such as the generative adversarial networks (Goodfellow et al., 2020). This can be studied with fault diagnosis in the future work.
7. The speed of equipment operation in real production environments is variable. However, most diagnostic methods use a constant speed regime. The variable speed state is a challenge to the accuracy and confidence of existing models. This problem could be tackled by using various base learners (models trained on failure data at different speeds) (Han et al., 2021), (Chen et al., 2022b).
8. Ensemble learning based fault diagnosis can effectively improve the accuracy of fault diagnosis. Accurately discovering and solving system failures can effectively reduce operating and maintenance costs, increase economic income, and save energy. These studies have high social value. However, in our study it has been found that most studies are limited to experimental results, lacking discussion of actual socio-economic impacts. Future work will increase the research on these contents and realize the discussion on the social contribution of intelligent diagnosis.
9. Vibration signal data collected from contact sensors is widely used in fault diagnosis. However, there are cases where the contact sensor cannot be installed at the appropriate location, thus affecting the effectiveness of fault diagnosis. In (Li et al., 2023), the literature analyzed some non-contact sensors and introduced a new fault data acquisition method. For increasingly complex equipment environments, selecting appropriate data acquisition systems is crucial. Providing data acquisition systems that can meet different needs according to the equipment environment will be a valuable research direction.
10. The field of intelligent fault diagnosis faces significant challenges in privacy protection and low power consumption. More secure and efficient encryption methods such as PQC and LWC are bound to bring important opportunities to artificial intelligence security. This also means that combining these studies with intelligent fault diagnosis will be a valuable research direction. The study of low-power and more secure fault diagnosis methods is in line with the current international society's development vision of green and energy-saving. This will help build a more secure, sustainable, and efficient intelligent fault diagnosis ecosystem.

CRediT authorship contribution statement

Zhibao Mian: Conceptualization, Supervision, Formal analysis, Project administration, Funding acquisition, Writing – original draft. **Xiaofei Deng:** Methodology, Investigation, Writing – original draft, Visualization. **Xiaohui Dong:** Validation, Writing – review & editing. **Yuzhu Tian:** Funding acquisition, Writing – review & editing. **Tianya Cao:** Writing – review & editing. **Kairan Chen:** Writing – review & editing. **Tareq Al Jaber:** Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

No data was used for the research described in the article.

Acknowledgements

This work was supported in part by the Industry Support Plan Project of Gansu Province (grant #2021CYZC-06), the NSFC (grant #12061065), Ningxia Key Research and Development Plan Project (grant #2018BEE03019), the Young Doctor Research Funding grant of Gansu Province (grant #2021QB-022), and the Science and Technology Program of Gansu Province (grant #21JR7RA135).

We thank the ESTC panel and all reviewers for their valuable comments. We would like to especially thank emeritus Dr. Leonardo Bottaci for critical comments and proofreading.

References

Aghaie, A., Kermani, M., Azarderakhsh, R., 2016. Fault diagnosis schemes for low-energy block cipher Midori benchmarked on FPGA. IEEE Trans. Very Large Scale Integr. Syst. 25 (4), 1528–1536. https://doi.org/10.1109/TVLSI.2016.2633412.

Agogino, A., Goebel, K., 2021. Milling Data Set." NASA Ames Prognostics Data Repository. BEST lab, UC Berkeley [Online]. Available: https://data.nasa.gov/download/vjv9-9f3x/application%2Fzip.

Ali, S., Guo, X., Karri, R., et al., 2016. "Fault attacks on AES and their countermeasures. In: Secure System Design and Trustable Computing. Springer International Publishing. https://doi.org/10.1007/978-3-319-14971-4_5.

Allwein, E., Schapire, R., Singer, Y., 2000. Reducing multiclass to binary: a unifying approach for margin classifiers. J. Mach. Learn. Res. 1 (Dec), 113–141.

Alotaibi, B., 2020. Transportation mode detection by embedded sensors based on ensemble learning. IEEE Access 8, 145552–145563. https://doi.org/10.1109/access.2020.3014901.

Anderson, C., 2008. The Long Tail: Why the Future of Business Is Selling Less of More," *Hyperion - Acquired Assets*, Revised edition.

Ansari, L., Ji, S., Chen, Q., et al., 2023. Ensemble hybrid learning methods for automated depression detection. IEEE Trans. Comput. Soc. Syst. 10 (1), 211–219. https://doi.org/10.1109/TCSS.2022.3154442.

Ardakani, M., Shokry, A., Escudero, G., et al., 2018. Online quantification of the concept drift using incremental learned classifier and non-automatic clustering. Comput. Aided Chem. Eng. 43, 1069–1074. https://doi.org/10.1016/B978-0-444-64235-6.50187-X. Elsevier.

Badr, M., Abdel-Khalik, A., Hamad, M., et al., 2023. Intelligent fault identification strategy of photovoltaic array based on ensemble self-training learning. Sol. Energy 249, 122–138. https://doi.org/10.1016/j.solener.2022.11.017.

Batzel, T., Swanson, D., 2009. Prognostic health management of aircraft power generators. IEEE Trans. Aero. Electron. Syst. 45 (2), 473–482. https://doi.org/10.1109/TAES.2009.5089535.

Ben, K., Zhang, Y., 2006. Artificial Intelligence. Tsinghua University Press, China.

Beretta, M., Vidal, Y., Sepulveda, J., et al., 2021. Improved ensemble learning for wind turbine main bearing fault diagnosis. Appl. Sci. 11 (16) https://doi.org/10.3390/APP11167523. Aug.

Bernstein, D., 2009. Introduction to post-quantum cryptography. In: Post-Quantum Cryptography. Springer, pp. 1–14. https://doi.org/10.1007/978-3-540-88702-7_1.

Berzati, A., Viera, A., Chartouny, M., et al., 2023. Exploiting intermediate value leakage in dilithium: a template-based approach. IACR Trans. Cryptogr. Hardw. Embed. Syst. 2023 (4), 188–210. https://doi.org/10.46586/tches.v2023.i4.188-210.

Breiman, L., 1996. Bagging predictors. Mach. Learn. 24 (2), 123–140. https://doi.org/10.1007/BF00058655.

Cai, B., Huang, L., Xie, M., 2017. Bayesian networks in fault diagnosis. IEEE Trans. Ind. Inf. 13 (5), 2227–2240. https://doi.org/10.1109/tii.2017.2695583.

Canto, A., Kaur, J., Kermani, M., et al., 2023a. Algorithmic Security Is Insufficient: A Comprehensive Survey on Implementation Attacks Haunting Post-Quantum Security. arXiv Prepr.

Canto, A., Kaur, J., Kermani, M., et al., 2023b. ChatGPT vs. Lightweight Security: First Work Implementing the NIST Cryptographic Standard ASCON. arXiv Prepr.

Cao, X., 2019. Research on Equipment Fault Diagnosis Methods Based on Incremental Ensemble Learning. Hebei University of Technology.

Cao, P., Zhang, S., Tang, J., 2018. Preprocessing-free gear fault diagnosis using small datasets with deep convolutional neural network-based transfer learning. IEEE Access 6, 26241–26253. https://doi.org/10.1109/ACCESS.2018.2837621.

Cao, Z., Li, Z., Zhang, J., et al., 2022. A homogeneous stacking ensemble learning model for fault diagnosis of rotating machinery with small samples. IEEE Sensor. J. 22 (9), 8944–8959. https://doi.org/10.1109/JSEN.2022.3163760.

Celaya, J., Wysocki, P., Goebel, K., et al., 2009. IGBT Accelerated Aging Data Set," *NASA Ames Prognostics Data Repository*. NASA Ames Research Center Moffett Field, CA [Online]. Available: https://data.nasa.gov/download/7wwx-fk77/application%2Fzip.

Cen, J., Yang, Z., Liu, X., et al., 2022. A review of data-driven machinery fault diagnosis using machine learning algorithms. J. Vib. Eng. Technol. 10, 2481–2507. https://doi.org/10.1007/s42417-022-00498-9.

Che, C., Wang, H., Xiong, M., et al., 2022. Few-shot fault diagnosis of rolling bearing under variable working conditions based on ensemble meta-learning. Digit. Signal Process. 131, 103777 https://doi.org/10.1016/j.dsp.2022.103777.

Chegini, S., Amini, P., Ahmadi, B., et al., 2022. Intelligent bearing fault diagnosis using swarm decomposition method andnew hybrid particle swarm optimization algorithm. Soft Comput. 26 (3), 1475–1497. https://doi.org/10.1007/s00500-021-06307-x.

Chen, Y., Liang, S., Li, W., et al., 2019. Faults and diagnosis methods of permanent magnet synchronous motors: a review. Appl. Sci. 9 (10), 2116. https://doi.org/10.3390/app9102116.

Chen, Y., Chen, H., Ye, P., et al., 2020. Acting as a decision maker: traffic-condition-aware ensemble learning for traffic flow prediction. IEEE Trans. Intell. Transport. Syst. 23 (4), 3190–3200. https://doi.org/10.1109/tits.2020.3032758.

Chen, T., Chen, D., Lv, R., 2021. Research on bearing fault diagnosis methods based on ensemble methods. Bull. Sci. Technol. https://doi.org/10.13774/j.cnki.kjtb.2021.04.011.

Chen, Z., Chen, J., Xie, Z., et al., 2022a. Multi-expert Attention Network with Unsupervised Aggregation for long-tailed fault diagnosis under speed variation. Knowl. Base Syst. 252, 109393 https://doi.org/10.1016/j.knosys.2022.109393.

Chen, Z., Chen, J., Xie, Z., et al., 2022b. Multi-expert Attention Network with Unsupervised Aggregation for long-tailed fault diagnosis under speed variation. Knowl. Base Syst. 252, 109393 https://doi.org/10.1016/j.knosys.2022.109393.

Cheng, G., Chen, X., Li, H., et al., 2016. Study on planetary gear fault diagnosis based on entropy feature fusion of ensemble empirical mode decomposition. Measurement 91, 140–154. https://doi.org/10.1016/j.measurement.2016.05.

Cheng, Y., Wu, J., Zhu, H., et al., 2020. Remaining useful life prognosis based on ensemble long short-term memory neural network. IEEE Trans. Instrum. Meas. 70, 1–12. https://doi.org/10.1109/TIM.2020.3031113.

Chowdhury, D., Sinha, A., Das, D., 2022. XAI-3DP: diagnosis and understanding faults of 3-D printer with explainable ensemble AI. IEEE Sens. Lett. 7 (1), 1–4. https://doi.org/10.1109/LSENS.2022.3228327.

Daemen, J., Rijmen, V., 2001. Announcing the advanced encryption standard (AES). Fed. Inf. Process. Stand. Publ. 197, 1–51.

Daga, A., Fasana, A., Marchesiello, S., et al., 2019. The Politecnico di Torino rolling bearing test rig: description and analysis of open access data. Mech. Syst. Signal Process. 120, 252–273. https://doi.org/10.1016/j.ymssp.2018.10.010.

Deng, Y., Du, S., Jia, S., et al., 2020. Prognostic study of ball screws by ensemble data-driven particle filters. J. Manuf. Syst. 56, 359–372. https://doi.org/10.1016/j.jmsy.2020.06.009.

Ding, K., Yao, C., Li, Y., et al., 2022. A review on fault diagnosis technology of key components in cold ironing system. Sustainability 14 (10), 6197. https://doi.org/10.3390/SU14106197.

Dong, S., 2022. Virtual currency price prediction based on segmented integrated learning. In: 2022 IEEE 2nd International Conference on Power, Electronics and Computer Applications (ICPECA), pp. 549–552. https://doi.org/10.1109/ICPECA53709.2022.9719070.

Dubrova, E., Ngo, K., Gärtner, J., et al., 2023. "Breaking a fifth-order masked implementation of crystals-kyber by copy-paste. In: Proceedings of the 10th ACM Asia Public-Key Cryptography Workshop," in Proceedings of the 10th ACM Asia Public-Key Cryptography Workshop. Association for Computing Machinery, New York, NY, USA, pp. 10–20. https://doi.org/10.1145/3591866.3593072.

Eskandari, A., Milimonfared, J., Aghaei, M., 2020. Line-line fault detection and classification for photovoltaic systems using ensemble learning model based on IV characteristics. Sol. Energy 211, 354–365. https://doi.org/10.1016/j.solener.2020.09.071.

Fang, L., Luo, Y., Feng, K., et al., 2023. A knowledge-enriched ensemble method for word embedding and multi-sense embedding. IEEE Trans. Knowl. Data Eng. 35 (6), 5534–5549. https://doi.org/10.1109/TKDE.2022.3159539.

Fawcett, T., 2006. An introduction to ROC analysis. Pattern Recogn. Lett. 27 (8), 861–874. https://doi.org/10.1016/j.patrec.2005.10.010.

Gao, W., Wai, R., Qiao, S., et al., 2019. Mechanical faults diagnosis of high-voltage circuit breaker via hybrid features and integrated extreme learning machine. IEEE Access 7, 60091–60103. https://doi.org/10.1109/ACCESS.2019.2915252.

Glowacz, A., 2015. Recognition of acoustic signals of synchronous motors with the use of MoFS and selected classifiers. Meas. Sci. Rev. 15 (4), 167–175. https://doi.org/10.1515/msr-2015-0024.

Glowacz, A., 2016a. Diagnostics of rotor damages of three-phase induction motors using acoustic signals and SMOFS-20-EXPANDED. Arch. Acoust. Q. 41 (3), 507–515. https://doi.org/10.1515/aoa-2016-0049.

Glowacz, A., 2016b. Fault diagnostics of acoustic signals of loaded synchronous motor using SMOFS-25-EXPANDED and selected classifiers. Teh. Vjesn. 23 (5), 1365–1372. https://doi.org/10.17559/TV-20150328135652.

Goodfellow, I., Abadie, J., Mirza, M., et al., 2020. Generative adversarial networks. Commun. ACM 63 (11), 139–144. https://doi.org/10.1145/3422622.

Gou, B., Xu, Y., Xia, Y., et al., 2018. An intelligent time-adaptive data-driven method for sensor fault diagnosis in induction motor drive system. IEEE Trans. Ind. Electron. 66 (12), 9817–9827. https://doi.org/10.1109/TIE.2018.2880719.

Gou, B., Xu, Y., Feng, X., 2020. State-of-health estimation and remaining-useful-life prediction for lithium-ion battery using a hybrid data-driven method. IEEE Trans. Veh. Technol. 69 (10), 10854–10867. https://doi.org/10.1109/TVT.2020.3014932.

Grill, J., Strub, F., Altché, F., et al., 2020. Bootstrap your own latent a new approach to self-supervised learning. Adv. Neural Inf. Process. Syst. 1786, 21271–21284. Red Hook, NY, USA: Curran Associates, Inc.

Grubbs, F., 1969. Procedures for detecting outlying observations in samples. Technometrics 11 (1), 1–21. https://doi.org/10.1080/00401706.1969.10490657.

Gungor, O., Rosing, T., Aksanli, B., 2021. OPELRUL: OPtimally weighted ensemble learner for remaining useful life prediction. In: 2021 IEEE International Conference on Prognostics and Health Management (ICPHM), pp. 1–8. https://doi.org/10.1109/ICPHM51084.2021.9486535.

Guo, H., Li, Y., Jennifer, S., et al., 2017. Learning from class-imbalanced data: review of methods and applications. Expert Syst. Appl. 73, 220–239. https://doi.org/10.1016/j.eswa.2016.12.035.

Han, T., Jiang, D., 2016. Rolling bearing fault diagnostic method based on VMD-AR model and random forest classifier. Shock Vib. 2016 (6), 1–11. https://doi.org/10.1155/2016/5132046.

Han, L., Yu, C., Liu, C., et al., 2019. Fault diagnosis of rolling bearings in rail train based on exponential smoothing predictive segmentation and improved ensemble learning algorithm. Appl. Sci. 9 (15), 3143. https://doi.org/10.3390/app9153143.

Han, H., Zhang, Z., Cui, X., et al., 2020. Ensemble learning with member optimization for fault diagnosis of a building energy system. Energy Build. 226, 110351 https://doi.org/10.1016/j.enbuild.2020.110351.

Han, B., Ji, S., Wang, J., et al., 2021. An intelligent diagnosis framework for roller bearing fault under speed fluctuation condition. Neurocomputing 420, 171–180. https://doi.org/10.1016/j.neucom.2020.09.022.

Hasan, M., 2013. Current Based Condition Monitoring of Electromechanical Systems. Model-free Drive System Current Monitoring: Faults Detection and Diagnosis through Statistical Features Extraction and Support Vector Machines Classification. University of Bradford [Online]. Available: http://hdl.handle.net/10454/5732.

He, H., Garcia, E., 2009. Learning from imbalanced data. IEEE Trans. Knowl. Data Eng. 21 (9), 1263–1284. https://doi.org/10.1109/TKDE.2008.239.

He, H., Bai, Y., Garcia, E., et al., 2008. ADASYN: adaptive synthetic sampling approach for imbalanced learning. In: 2008 IEEE International Joint Conference on Neural Networks. IEEE World Congress on Computational Intelligence), pp. 1322–1328. https://doi.org/10.1109/IJCNN.2008.4633969.

He, K., Zhang, X., Ren, S., et al., 2016. Deep residual learning for image recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 770–778.

He, J., Ouyang, M., Yong, C., et al., 2020. A novel intelligent fault diagnosis method for rolling bearing based on integrated weight strategy features learning. Sensors 20 (6), 1774. https://doi.org/10.3390/s20061774.

He, C., Liu, C., Wu, T., et al., 2021. Medical rolling bearing fault prognostics based on improved extreme learning machine. J. Combin. Optim. 42, 700–721. https://doi.org/10.1007/s10878-019-00494-y.

Hosseinpoor, Z., Arefi, M., Razavi-Far, R., et al., 2020. Virtual sensors for fault diagnosis: a case of induction motor broken rotor bar. IEEE Sensor. J. 21 (4), 5044–5051. https://doi.org/10.1109/JSEN.2020.3033754.

Hou, Y., Wang, Y., Pan, Y., et al., 2023. Vibration-based incipient surge detection and diagnosis of the centrifugal compressor using adaptive feature fusion and sparse ensemble learning approach. Adv. Eng. Inf. 56, 101947 https://doi.org/10.1016/j.aei.2023.101947.

Hu, Y., Liu, S., Lu, H., et al., 2019. Remaining useful life model and assessment of mechanical products: a brief review and a note on the state space model method. Chin. J. Mech. Eng. 32, 1–20. https://doi.org/10.1186/s10033-019-0317-y.

Jalali, A., Azarderakhsh, R., Kermani, M., et al., 2017. Supersingular isogeny Diffie-Hellman key exchange on 64-bit ARM. IEEE Trans. Dependable Secure Comput. 16 (5), 902–912. https://doi.org/10.1109/TDSC.2017.2723891.

Jalali, A., Azarderakhsh, R., Kermani, M., et al., 2019. Towards optimized and constant-time CSIDH on embedded devices. In: Polian, I., Stöttinger, M. (Eds.), Constructive Side-Channel Analysis and Secure Design. Springer International Publishing, Cham, pp. 215–231. https://doi.org/10.1007/978-3-030-16350-1_12.

Janet, B., Kumar, R., Anisha, M., et al., 2022. Education data analysis using ensemble models. In: 2022 4th International Conference on Smart Systems and Inventive Technology (ICSSIT). IEEE, pp. 1612–1617. https://doi.org/10.1109/ICSSIT53264.2022.9716453.

Jian, C., Ao, Y., 2022. Industrial fault diagnosis based on diverse variable weighted ensemble learning. J. Manuf. Syst. 62, 718–735. https://doi.org/10.1016/j.jmsy.2022.02.006.

Jiang, H., 2020. Research on Fault Classification Algorithm of Rolling Bearing for Rotating Equipment Based on NN and XGBoost. China University of Mining and Technology, Jiangsu).

Kamble, K., Sengupta, J., 2022. Ensemble machine learning-based affective computing for emotion recognition using dual-decomposed EEG signals. IEEE Sensor. J. 22 (3), 2496–2507. https://doi.org/10.1109/JSEN.2021.3135953.

Kang, X., Cheng, J., Wang, P., et al., 2022. A novel ensemble convex hull-based classification model for bevel gearbox fault diagnosis. Meas. Sci. Technol. 34 (3), 35017 https://doi.org/10.1088/1361-6501/ACA8C1.

Kapucu, C., Cubukcu, M., 2021. A supervised ensemble learning method for fault diagnosis in photovoltaic strings. Energy 227, 120463. https://doi.org/10.1016/j.energy.2021.120463.

Kaur, H., Pannu, H., Malhi, A., 2019. A systematic review on imbalanced data challenges in machine learning: applications and solutions. ACM Comput. Surv. 52 (4), 1–36. https://doi.org/10.1145/3343440.

Kaur, J., Canto, A., Kermani, M., et al., 2023. A Comprehensive Survey on the Implementations, Attacks, and Countermeasures of the Current NIST Lightweight Cryptography Standard. arXiv Prepr.

Kermani, M., Masoleh, A., 2011. A high-performance fault diagnosis approach for the AES SubBytes utilizing mixed bases. In: 2011 Workshop on Fault Diagnosis and Tolerance in Cryptography, pp. 80–87. https://doi.org/10.1109/FDTC.2011.11.

Kim, N., An, D., Choi, J., 2017. Prognostics and Health Management of Engineering Systems. Springer International Publishing, Switzerland.

Kowalski, J., Krawczyk, B., Woźniak, M., 2017. Fault diagnosis of marine 4-stroke diesel engines using a one-vs-one extreme learning ensemble. Eng. Appl. Artif. Intell. 57, 134–141. https://doi.org/10.1016/j.engappai.2016.10.015.

Koziel, B., Azarderakhsh, R., Kermani, M., 2015. Low-resource and fast binary edwards curves cryptography. In: Biryukov, A., Goyal, V. (Eds.), Progress in Cryptology–INDOCRYPT 2015. Springer International Publishing, Cham, pp. 347–369. https://doi.org/10.1007/978-3-319-26617-6_19.

Krzywanski, J., Skrobek, D., Zylka, A., et al., 2023. Heat and mass transfer prediction in fluidized beds of cooling and desalination systems by AI approach. Appl. Therm. Eng. 225, 120200 https://doi.org/10.1016/j.applthermaleng.2023.120200.

Kumar, A., Kumar, R., 2022. Acoustic and vibration data for defect cases of the centrifugal pump. IEEE Dataport [Online]. Available: https://dx.doi.org/10.21227/8cqr-jr43.

Kuncheva, L., Whitaker, C., 2003. Measures of diversity in classifier ensembles and their relationship with the ensemble accuracy. Mach. Learn. 51 (2), 181. https://doi.org/10.1023/A:1022859003006.

Kundu, P., Darpe, A., Kulkarni, M., 2020. An ensemble decision tree methodology for remaining useful life prediction of spur gears under natural pitting progression. Struct. Health Monit. 19 (3), 854–872. https://doi.org/10.1177/1475921719865718.

Leao, B., Yoneyama, T., Rocha, G., et al., 2008. Prognostics performance metrics and their relation to requirements, design, verification and cost-benefit. In: 2008 International Conference on Prognostics and Health Management, pp. 1–8. https://doi.org/10.1109/PHM.2008.4711429.

Lee, J., Qiu, H., Yu, G., et al., 2007. Bearing Data Set," *NASA Ames, Moffett Field, CA*. NASA Ames Research Center, Moffett Field, CA [Online]. Available: http://ti.arc.nasa.gov/project/prognostic-data-repository).

Lei, Y., Jia, F., Lin, J., et al., 2016. An intelligent fault diagnosis method using unsupervised feature learning towards mechanical big data. IEEE Trans. Ind. Electron. 63 (5), 3137–3147. https://doi.org/10.1109/tie.2016.2519325.

Li, J., Lin, M., 2021. Novel application of ensemble learning method based on the plurality voting rule in fault diagnosis of nuclear power plants. In: Progress Report on Nuclear Science and Technology in China (Volume 7) Proceedings of the 2021 Annual Conference of the Chinese Nuclear Society, Volume 2 (Nuclear Power Volume). Chinese Nuclear Society, Beijing,China.

Li, S., Liu, G., Tang, X., et al., 2017. An ensemble deep convolutional neural network model with improved DS evidence fusion for bearing fault diagnosis. Sensors 17 (8), 1729. https://doi.org/10.3390/s17081729.

Li, Z., Goebel, K., Wu, D., 2019. Degradation modeling and remaining useful life prediction of aircraft engines using ensemble learning. J. Eng. Gas Turbines Power 141 (4), 041008. https://doi.org/10.1115/1.4041674.

Li, Y., Song, Y., Jia, L., et al., 2020. Intelligent fault diagnosis by fusing domain adversarial training and maximum mean discrepancy via ensemble learning. IEEE Trans. Ind. Inf. 17 (4), 2833–2841. https://doi.org/10.1109/tii.2020.3008010.

Li, X., Jiang, H., Wang, R., et al., 2021a. Rolling bearing fault diagnosis using optimal ensemble deep transfer network. Knowl. Base Syst. 213, 106695 https://doi.org/10.1016/j.knosys.2020.106695.

Li, Z., Jiang, W., Zhang, S., et al., 2021b. A hydraulic pump fault diagnosis method based on the modified ensemble empirical mode decomposition and wavelet kernel extreme learning machine methods. Sensors 21 (8), 2599. https://doi.org/10.3390/s21082599.

Li, X., Jiang, H., Xie, M., et al., 2022. A reinforcement ensemble deep transfer learning network for rolling bearing fault diagnosis with multi-source domains. Adv. Eng. Inf. 51, 101480 https://doi.org/10.1016/j.aei.2021.101480.

Li, X., Yu, S., Lei, Y., et al., 2023. Intelligent machinery fault diagnosis with event-based camera. IEEE Trans. Ind. Inf. 1–10. https://doi.org/10.1109/TII.2023.3262854.

Lin, T., Goyal, P., Girshick, R., et al., 2020. Focal loss for dense object detection. IEEE Trans. Pattern Anal. Mach. Intell. 42 (2), 318–327. https://doi.org/10.1109/TPAMI.2018.2858826.

Liu, J., 2020. Research on Railway Switch Fault Diagnosis System Based on Ensemble Learning. Shijiazhuang Tiedao University.

Liu, B., Hao, Z., Tsang, E., 2008. Nesting one-against-one algorithm based on SVMs for pattern classification. IEEE Trans. Neural Network. 19 (12), 2044–2052. https://doi.org/10.1109/TNN.2008.2003298.

Liu, C., Zhang, Y., Sun, J., et al., 2022a. Stacked bidirectional LSTM RNN to evaluate the remaining useful life of supercapacitor. Int. J. Energy Res. 46 (3), 3034–3043. https://doi.org/10.1002/er.7360.

Liu, G., Shen, W., Gao, L., et al., 2022b. Knowledge transfer in fault diagnosis of rotary machines. IET Collab. Intell. Manuf. 4 (1), 17–34. https://doi.org/10.1049/cim2.12047.

Lu, J., Liu, A., Dong, F., et al., 2019. Learning under concept drift: a review. IEEE Trans. Knowl. Data Eng. 31 (12), 2346–2363. https://doi.org/10.1109/TKDE.2018.2876857.

Luo, M., Li, C., Zhang, X., et al., 2016. Compound feature selection and parameter optimization of ELM for fault diagnosis of rolling element bearings. ISA Trans. 65, 556–566. https://doi.org/10.1016/j.isatra.2016.08.022.

Ma, S., Chu, F., 2019. Ensemble deep learning-based fault diagnosis of rotor bearing systems. Comput. Ind. 105, 143–152. https://doi.org/10.1016/j.compind.2018.12.012.

Mehta, S., Darji, M., Joshi, H., 2022. Threat prediction using ensemble learning algorithm to provide end-point security. In: 2022 International Conference on Electronics and Renewable Systems. ICEARS), pp. 1427–1434. https://doi.org/10.1109/ICEARS53579.2022.9751963.

Mellit, A., Zayane, C., Boubaker, S., et al., 2023. A sustainable fault diagnosis approach for photovoltaic systems based on stacking-based ensemble learning methods. Mathematics 11 (4), 936. https://doi.org/10.3390/MATH11040936.

Mian, Z., Jia, S., Shi, X., et al., 2020. A model-based RCM analysis method. In: 2020 IEEE 20th International Conference on Software Quality, Reliability and Security Companion. QRS-C), pp. 301–307. https://doi.org/10.1109/QRS-C51114.2020.00059.

Moubray, J., 2001. Reliability-centered Maintenance. Industrial Press Inc.

Muneer, A., Taib, S., Naseer, S., et al., 2021. Data-driven deep learning-based attention mechanism for remaining useful life prediction: case study application to turbofan engine analysis. Electronics 10 (20), 2453. https://doi.org/10.3390/electronics10202453.

Naveen, S., Sugumaran, V., 2021. Fault diagnosis of visual faults in photovoltaic modules: a Review. Int. J. Green Energy 18 (1), 37–50. https://doi.org/10.1080/15435075.2020.182544.

Niasar, M., Azarderakhsh, R., Kermani, M., 2021. Instruction-set accelerated implementation of CRYSTALS-Kyber. IEEE Trans. Circuits Syst. I Regul. Pap. 68 (11), 4648–4659. https://doi.org/10.1109/TCSI.2021.3106639.

Pang, S., Wang, J., Yang, X., et al., 2021. Detecting rotating machinery faults under different working conditions with cross-domain negative correlated ensemble algorithm. Measurement 184, 109951. https://doi.org/10.1016/j.measurement.2021.109951.

Patel, H., Rajput, D., Reddy, G., et al., 2020. A review on classification of imbalanced data for wireless sensor networks. Int. J. Distributed Sens. Netw. 16 (4) https://doi.org/10.1177/1550147720916404.

Peng, T., He, J., Xiang, Y., et al., 2015. Probabilistic fatigue damage prognosis of lap joint using Bayesian updating. J. Intell. Mater. Syst. Struct. 26 (8), 965–979. https://doi.org/10.1177/1045389X14538328.

Peng, D., Wang, H., Liu, Z., et al., 2020. Multibranch and multiscale CNN for fault diagnosis of wheelset bearings under strong noise and variable load condition. IEEE Trans. Ind. Inf. 16 (7), 4949–4960. https://doi.org/10.1109/tii.2020.2967557.

Porotsky, S., Bluvband, Z., 2012. Remaining useful life estimation for systems with non-trendability behaviour. In: 2012 Ieee Conference on Prognostics and Health Management, pp. 1–6. https://doi.org/10.1109/ICPHM.2012.6299544.

Raichura, M., Chothani, N., Patel, D., 2020. Identification of internal fault against external abnormalities in power transformer using hierarchical ensemble extreme learning machine technique. IET Sci. Meas. Technol. 14 (1), 111–121. https://doi.org/10.1049/iet-smt.2019.0102.

Randall, R., 2010. Vibration-based Condition Monitoring: Industrial, Automotive and Aerospace Applications. John Wiley & Sons. https://doi.org/10.1002/9780470977668.

Ren, J., Lin, X., Liu, J., et al., 2020. Engineering early prediction of supercapacitors cycle life using neural networks. Mater. Today Energy 18, 100537. https://doi.org/10.1016/j.mtener.2020.100537.

Ren, Z., Su, W., Lin, T., et al., 2022. Gradient harmonized loss: improving the performance of intelligent diagnosis models in large imbalance scenarios. In: 2022 IEEE International Conference on Prognostics and Health Management (ICPHM), pp. 34–39. https://doi.org/10.1109/ICPHM53196.2022.9815705.

Rong, Z., 2021. Study on several important design variables of control technology for energy-saving high-speed hydraulic press. China Met. Equip. Manuf. Technol. 56 (5), 29–33. https://doi.org/10.16316/j.issn.1672-0121.2021.05.005.

Saha, B., Goebel, K. "Battery data set," NASA prognostics data repository, NASA ames research center, moffett field, CA [Online]. Available: https://phm-datasets.s3.amazonaws.com/NASA/5.+Battery+Data+Set.zip.

Saidi, L., Benbouzid, M., 2021. Prognostics and health management of renewable energy systems: state ofthe art review, challenges, and trends. Electronics 10 (22). https://doi.org/10.3390/electronics10222732.

Saidi, L., Ben Ali, J., Bechhoefer, E., et al., 2017. Particle filter-based prognostic approach for high-speed shaft bearing wind turbine progressive degradations. In: IECON 2017-43rd Annual Conference of the IEEE Industrial Electronics Society, pp. 8099–8104. https://doi.org/10.1109/IECON.2017.8217421.

Sanal, P., Karagoz, E., Seo, H., et al., 2021. "Kyber on ARM64: compact implementations of kyber on 64-bit ARM cortex-A processors. In: Security and Privacy in Communication Networks. Springer International Publishing. https://doi.org/10.1007/978-3-030-90022-9_23.

Sarker, A., Kermani, M., Azarderakhsh, R., 2020. Error detection architectures for ring polynomial multiplication and modular reduction of ring-LWE in $\frac{¢/p¢[x]}{x^n+1}$ benchmarked on ASIC. IEEE Trans. Reliab. 70 (1), 362–370. https://doi.org/10.1109/TR.2020.2991671.

Sarmadi, S., Kermani, M., Azarderakhsh, R., et al., 2013. Dual-basis superserial multipliers for secure applications and lightweight cryptographic architectures. IEEE Trans. Circuits Syst. II Express Briefs 61 (2), 125–129. https://doi.org/10.1109/TCSII.2013.2291075.

Sato, D., Freitas, S., Barddal, J., et al., 2021. A survey on concept drift in process mining. ACM Comput. Surv. 54 (9), 1–38. https://doi.org/10.1145/3472752.

Saxena, A., Goebel, K., 2008. Turbofan engine degradation simulation data set. In: NASA Ames Prognostics Data Repository, vol. 18. NASA Ames Research Center Moffett Field, CA, USA [Online]. Available: https://www.nasa.gov/intelligent-systems-division/discovery-and-systems-health/pcoe/pcoe-data-set-repository/.

Schapire, R., 1990. The strength of weak learnability. Mach. Learn. 5 (2), 197–227. https://doi.org/10.1007/BF00116037.

Shao, H., Jiang, H., Zhang, X., et al., 2015. Rolling bearing fault diagnosis using an optimization deep belief network. Meas. Sci. Technol. 26 (11), 115002 https://doi.org/10.1088/0957-0233/26/11/115002.

Shao, S., McAleer, S., Yan, R., et al., 2019. Highly accurate machine fault diagnosis using deep transfer learning. IEEE Trans. Ind. Inf. 15 (4), 2446–2455. https://doi.org/10.1109/TII.2018.2864759.

Sharma, A., Amarnath, M., Kankar, P., 2017. Novel ensemble techniques for classification of rolling element bearing faults. J. Brazilian Soc. Mech. Sci. Eng. 39, 709–724. https://doi.org/10.1007/s40430-016-0540-8.

Sharvia, S., Kabir, S., Walker, M., et al., 2016. Model-based dependability analysis: state-of-the-art, challenges, and future outlook. Softw. Qual. Assur. 251–278. https://doi.org/10.1016/b978-0-12-802301-3.00012-0.

Shi, Q., Zhang, H., 2020. Fault diagnosis of an autonomous vehicle with an improved SVM algorithm subject to unbalanced datasets. IEEE Trans. Ind. Electron. 68 (7), 6248–6256. https://doi.org/10.1109/tie.2020.2994868.

Smarandache, F., Dezert, J., 2006. Advances and Applications of DSmT for Information Fusion (Collected Works), second. American Research Press, United States.

Smith, W., Randall, R., 2015. Rolling element bearing diagnostics using the Case Western Reserve University data: a benchmark study. Mech. Syst. Signal Process. 64, 100–131. https://doi.org/10.1016/j.ymssp.2015.04.021.

Sosnowski, M., Krzywanski, J., Ščurek, R., 2021. Artificial intelligence and computational methods in the modeling of complex systems. Entropy 23 (5), 586. https://doi.org/10.3390/e23050586. MDPI.

Sun, J., Zuo, H., Yang, H., et al., 2010. Study of ensemble learning-based fusion prognostics. In: 2010 Prognostics and System Health Management Conference, pp. 1–7. https://doi.org/10.1109/PHM.2010.5414582.

Tang, J., Chai, T., Yu, W., et al., 2016. A comparative study that measures ball mill load parameters through different single-scale and multiscale frequency spectra-based approaches. IEEE Trans. Ind. Inf. 12 (6), 2008–2019. https://doi.org/10.1109/TII.2016.2586419.

Tang, G., Pang, B., Tian, T., et al., 2018. Fault diagnosis of rolling bearings based on improved fast SpectralCorrelation and optimized random forest. Appl. Sci. 8 (10) https://doi.org/10.3390/app8101859. Oct.

Tang, Y., Wei, A., Tong, T., et al., 2022. A Review of Vibration Monitoring Technology for Fault Diagnosis of Planetary Gearboxes of Wind Turbines, vol. 8. Wind energy, pp. 92–95.

Tong, J., Liu, C., Bao, J., et al., 2023. A novel ensemble learning based multi-sensor information fusion method for rolling bearing fault diagnosis. IEEE Trans. Instrum. Meas. 72, 1–12. https://doi.org/10.1109/TIM.2022.3225910.

Treml, A., Flauzino, R., Suetake, M., et al., 2020. Experimental database for detecting and diagnosing rotor broken bar in a three-phase induction motor. IEEE Dataport. IEEE, Sep. 15 https://doi.org/10.21227/fmnm-bn95.

Ture, B., Akbulut, A., Zaim, A., et al., 2023. Stacking-based ensemble learning for remaining useful life estimation. Soft Comput. 1–13. https://doi.org/10.1007/s00500-023-08322-6.

Wan, L., Gong, K., Zhang, G., et al., 2021. An efficient rolling bearing fault diagnosis method based on spark and improved random forest algorithm. IEEE Access 9, 37866–37882. https://doi.org/10.1109/access.2021.3063929.

Wang, S., 2015. Study on Power Transformer Fault Diagnosis Based on Ensemble Learning. Chongqing University.

Wang, H., 2021. Research on Fault Diagnosis Algorithm of Photovolttaic Array Based on Ensemble Learning. North China Electric Power University, Beijing).

Wang, X., Han, T., 2020. Transformer fault diagnosis based on stacking ensemble learning. IEEJ Trans. Electr. Electron. Eng. 15 (12), 1734–1739. https://doi.org/10.1002/tee.23247.

Wang, Y., Vinogradov, A., 2023. Improving the performance of convolutional GAN using history-state ensemble for unsupervised early fault detection with acoustic emission signals. Appl. Sci. 13 (5), 3136. https://doi.org/10.3390/APP13053136.

Wang, A., Sha, M., Liu, L., et al., 2015. A new process industry fault diagnosis algorithm based on ensemble improved binary-tree SVM. Chin. J. Electron. 24 (2), 258–262. https://doi.org/10.1049/cje.2015.04.006.

Wang, Y., Ramanan, D., Hebert, M., 2017. Learning to model the tail. In: Advances in Neural Information Processing Systems30. NIPS, 2017.

Wang, Z., Lu, C., Zhou, B., 2018a. Fault diagnosis for rotary machinery with selective ensemble neural networks. Mech. Syst. Signal Process. 113, 112–130. https://doi.org/10.1016/j.ymssp.2017.03.05.

Wang, Z., Wang, Y., Srinivasan, R., 2018b. A novel ensemble learning approach to support building energy use prediction. Energy Build. 159, 109–122. https://doi.org/10.1016/j.enbuild.2017.10.085.

Wang, B., Lei, Y., Li, N., et al., 2018c. A hybrid prognostics approach for estimating remaining useful life of rolling element bearings. IEEE Trans. Reliab. 69 (1), 401–412. https://doi.org/10.1109/TR.2018.2882682.

Wang, X., Miao, P., Zhang, K., et al., 2019a. Study on novel signal processing and simultaneous-fault diagnostic method for wind turbine. Trans. Inst. Meas. Control 41 (14), 4100–4113. https://doi.org/10.1177/014233121984926.

Wang, M., Ge, Q., Jiang, H., et al., 2019b. Wear fault diagnosis of aeroengines based on broad learning system and ensemble learning. Energies 12 (24), 4750. https://doi.org/10.3390/en12244750.

Wang, H., Wu, X., Huang, Z., et al., 2020. High-Frequency component helps explain the generalization of convolutional neural networks. In: 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 8684–8694.

Wang, Z., Wang, C., Li, N., 2021a. Bearing fault diagnosis method based on similarity measure and ensemble learning. Meas. Sci. Technol. 32 (5), 055005 https://doi.org/10.1088/1361-6501/abda97.

Wang, Z., Huang, H., Wang, Y., 2021b. Fault diagnosis of planetary gearbox using multi-criteria feature selection and heterogeneous ensemble learning classification. Measurement 173, 108654. https://doi.org/10.1016/j.measurement.2020.10.

Wang, J., Wang, L., Qu, J., et al., 2021c. Novel application of heterogeneous ensemble learning in fault diagnosis of photovoltaic modules. In: 2021 International Conference on Smart-Green Technology in Electrical and Information Systems. ICSGTEIS), pp. 118–124. https://doi.org/10.1109/ICSGTEIS53426.2021.9650390.

Wang, C., Xiong, R., Tian, J., et al., 2022a. Rapid ultracapacitor life prediction with a convolutional neural network. Appl. Energy 305, 117819. https://doi.org/10.1016/j.apenergy.2021.117819.

Wang, L., Liu, H., Pan, Z., et al., 2022b. Long short-term memory neural network with transfer learning and ensemble learning for remaining useful life prediction. Sensors 22 (15), 5744. https://doi.org/10.3390/s22155744.

Wang, G., Zhang, Y., Zhang, F., et al., 2023a. An ensemble method with DenseNet and evidential reasoning rule for machinery fault diagnosis under imbalanced condition. Measurement 214, 112806. https://doi.org/10.1016/j.measurement.2023.112806.

Wang, J., Xu, H., Liu, J., et al., 2023b. A bagging-strategy based heterogeneous ensemble deep neural networks approach for the multiple components fault diagnosis of hydraulic systems. Meas. Sci. Technol. 34 (6), 65007 https://doi.org/10.1088/1361-6501/acbc3b.

Wen, C., Lv, F., 2020. Review on deep learning based fault diagnosis. J. Electron. Inf. Technol. 42 (1), 234–248. https://doi.org/10.11999/JEIT190715.

Wikipedia Contributors, "Ensemble Learning — Wikipedia, the Free Encyclopedia," *Wikipedia, the Free Encyclopedia*, 2023 [Online]. Available: https://en.wikipedia.org/w/index.php?title=Ensemble_learning&oldid=1172129024.

Wolpert, D., 1992. Stacked generalization. Neural Network. 5 (2), 241–259. https://doi.org/10.1016/S0893-6080(05)80023-1.

Wu, Q., 2020. Research on Fault Diagnosis of Engineering Machinery Based on Ensemble Learning. Southwest Jiaotong University.

Wu, Z., Jiang, H., Liu, S., et al., 2021. A deep ensemble dense convolutional neural network for rolling bearing fault diagnosis. Meas. Sci. Technol. 32 (10), 104014 https://doi.org/10.1088/1361-6501/AC05F5.

Xia, Y., Xu, Y., Gou, B., 2019. A data-driven method for IGBT open-circuit fault diagnosis based on hybrid ensemble learning and sliding-window classification. IEEE Trans. Ind. Inf. 16 (8), 5223–5233. https://doi.org/10.1109/tii.2019.2949344.

Xia, T., Song, Y., Zheng, Y., et al., 2020. An ensemble framework based on convolutional bi-directional LSTM with multiple time windows for remaining useful life estimation. Comput. Ind. 115, 103182 https://doi.org/10.1016/j.compind.2019.103182.

Xiong, J., Zhang, Q., Sun, G., et al., 2015. An information fusion fault diagnosis method based on dimensionless indicators with static discounting factor and KNN. IEEE Sensor. J. 16 (7), 2060–2069. https://doi.org/10.1109/jsen.2015.2497545.

Xu, L., 2020. Research on Photovoltaic Array Fault Diagnosis Method Based on Integrated Learning. Nanchang University.

Xu, J., Yang, Y., 2018. A survey of ensemble learning approaches. J. Yunnan Univ. Sci. 40 (6), 11. https://doi.org/10.7540/j.ynu.20180455.

Xu, G., Liu, M., Jiang, Z., et al., 2019. Bearing Fault diagnosis method based on deep convolutional neural network and random forest ensemble learning. Sensors 19 (5), 1088. https://doi.org/10.3390/s19051088.

Xuan, Y., Fan, L., Tu, Y., et al., 2022. Appearance Fault diagnosis of a transformer based on data set optimization and heterogeneous model ensemble. Front. Energy Res. 10, 648. https://doi.org/10.3389/FENRG.2022.902892.

Xue, Z., Zhang, Y., Zhang, L., et al., 2022. Ensemble learning embedded with Gaussian process regression for soil moisture estimation: a case study of the continental us. IEEE Trans. Geosci. Rem. Sens. 60, 1–17. https://doi.org/10.1109/TGRS.2022.3166777.

Yang, W., Xiao, M., Zhou, W., et al., 2016. A hybrid prognostic approach for remaining useful life prediction of lithium-ion batteries. Shock Vib. 2016. https://doi.org/10.1155/2016/3838765.

Yang, H., Li, W., Hu, K., et al., 2021. Deep ensemble learning with non-equivalent costs of fault severities for rolling bearing diagnostics. J. Manuf. Syst. 61, 249–264. https://doi.org/10.1016/j.jmsy.2021.09.009.

Yang, L., Jiang, H., Song, Q., et al., 2022. A survey on long-tailed visual recognition. Int. J. Comput. Vis. 130 (7), 1837–1872. https://doi.org/10.1007/s11263-022-01622-8. Jul.

Yao, Q., Yang, T., Liu, Z., et al., 2019. Remaining useful life estimation by empirical mode decomposition and ensemble deep convolution neural networks. In: 2019 IEEE International Conference on Prognostics and Health Management (ICPHM), pp. 1–6. https://doi.org/10.1109/ICPHM.2019.8819373.

Yi, Z., Zhao, K., Sun, J., et al., 2022. Prediction of the remaining useful life of supercapacitors. Math. Probl Eng. 2022, 1–8. https://doi.org/10.1155/2022/7620382.

You, K., Qiu, G., Gu, Y., 2023. An efficient lightweight neural network using BiLSTM-SCN-CBAM with PCA-ICEEMDAN for diagnosing rolling bearing faults. Meas. Sci. Technol. 34 (9) https://doi.org/10.1088/1361-6501/ACD5EF.

Yu, J., 2016. Machinery fault diagnosis using joint global and local/nonlocal discriminant analysis with selective ensemble learning. J. Sound Vib. 382, 340–356. https://doi.org/10.1016/j.jsv.2016.06.046.

Zang, Y., Shangguan, W., Cai, B., et al., 2019. Methods for fault diagnosis of high-speed railways: a review. Proc. Inst. Mech. Eng. part O J. risk Reliab. 233 (5), 908–922. https://doi.org/10.1177/1748006X18823932.

Zeng, J., Cheng, Y., 2020. An ensemble learning-based remaining useful life prediction method for aircraft turbine engine. IFAC-PapersOnLine 53 (3), 48–53. https://doi.org/10.1016/j.ifacol.2020.11.009.

Zenisek, J., Holzinger, F., Affenzeller, M., 2019. Machine learning based concept drift detection for predictive maintenance. Comput. Ind. Eng. 137, 106031 https://doi.org/10.1016/j.cie.2019.106031.

Zhang, W., Li, X., 2022. Data privacy preserving federated transfer learning in machinery fault diagnostics using prior distributions. Struct. Health Monit. 21 (4), 1329–1344. https://doi.org/10.1177/14759217211029201.

Zhang ang, L., Wang, Y., 2021. Transformer Fault diagnosis based on stacking-ensemble meta-algorithms. Adv. Comput. Signals Syst. 5 (1), 42–47. https://doi.org/10.1002/tee.23247.

Zhang, X., Wang, B., Chen, X., 2015. Intelligent fault diagnosis of roller bearings with multivariable ensemble-based incremental support vector machine. Knowl. Base Syst. 89, 56–85. https://doi.org/10.1016/j.knosys.2015.06.017.

Zhang, C., Lim, P., Qin, A., et al., 2016. Multiobjective deep belief networks ensemble for remaining useful life estimation in prognostics. IEEE Trans. neural networks Learn. Syst. 28 (10), 2306–2318. https://doi.org/10.1109/TNNLS.2016.2582798.

Zhang, D., Qian, L., Mao, B., et al., 2018. A data-driven design for fault detection of wind turbines using random forests and XGboost. IEEE Access 6, 21020–21031. https://doi.org/10.1109/access.2018.2818678.

Zhang, C., He, Y., Du, B., et al., 2020a. Transformer fault diagnosis method using IoT based monitoring system and ensemble machine learning. Future Generat. Comput. Syst. 108, 533–545. https://doi.org/10.1016/j.future.2020.03.008.

Zhang, Z., Han, H., Cui, X., et al., 2020b. Fault diagnosis of refrigeration system based on integrated learning of multiple models. J. Eng. Therm. Energy Power 35 (5), 153–162. https://doi.org/10.16146/j.cnki.rndlgc.2020.05.022.

Zhang, Y., Ji, J., Ma, B., 2020c. Fault diagnosis of reciprocating compressor using a novel ensemble empirical mode decomposition-convolutional deep belief network. Measurement 156, 107619. https://doi.org/10.1016/j.measurement.2020.10.

Zhang, X., Huang, T., Wu, B., et al., 2021a. Multi-model ensemble deep learning method for intelligent fault diagnosis with high-dimensional samples. Front. Mech. Eng. 16 (2), 340–352. https://doi.org/10.1007/s11465-021-0629-3.

Zhang, W., Wei, L., Yan, M., et al., 2021b. Research on life prediction method of vehicle supercapacitors module under service conditions. Proc. CSEE 41 (S01), 8. https://doi.org/10.1016/j.rser.2021.111408.

Zhang, L., Cheng, Y., Zhang, J., et al., 2023a. Refrigerant charge fault diagnosis strategy for VRF systems based on stacking ensemble learning. Build. Environ. 234, 110209 https://doi.org/10.1016/j.buildenv.2023.110209.

Zhang, Y., Kang, B., Hooi, B., et al., 2023b. Deep long-tailed learning: a survey. IEEE Trans. Pattern Anal. Mach. Intell. 45 (9), 10795–10816. https://doi.org/10.1109/TPAMI.2023.3268118.

Zheng, Y., 2011. Research on Ensemble Learning of Fault Diagnosis and Prediction and Maintence Decision-Making Models for Transformers. Chongqing University.

Zhong, X., Ban, H., 2022. Crack Fault Diagnosis of Rotating Machine in Nuclear Power Plant Based on Ensemble Learning. Elsevier. https://doi.org/10.1016/j.anucene.2021.108909.

Zhong, J., Wong, P., Yang, Z., 2018. Fault diagnosis of rotating machinery based on multiple probabilistic classifiers. Mech. Syst. Signal Process. 108, 99–114. https://doi.org/10.1016/j.ymssp.2018.02.009.

Zhong, G., Yin, T., Li, L., et al., 2020. Bus travel time prediction based on ensemble learning methods. IEEE Intell. Transp. Syst. Mag. 14 (2), 174–189. https://doi.org/10.1109/mits.2020.2990175.

Zhou, Z., 2021a. Ensemble learning. In: Machine Learning. Springer, Singapore, pp. 181–210. https://doi.org/10.1007/978-981-15-1967-3_8.

Zhou, H., 2021b. A Transformer Fault Diagnosis Method Based on Integrated Learning and Dissolved Gas Analysis in Oi. Kunming University of Science and Technology. https://doi.org/10.27200/d.cnki.gkmlu.2021.000861.

Zhou, M., Wang, K., Wang, Y., et al., 2019. Online condition diagnosis for a two-stage gearbox machinery of an aerospace utilization system using an ensemble multi-fault features indexing approach. Chin. J. Aeronaut. 32 (5), 1100–1110. https://doi.org/10.1016/j.cja.2019.02.013.

Zhou, S., Wei, C., Li, P., et al., 2021a. A text-driven aircraft fault diagnosis model based on Word2vec and stacking ensemble learning. Aerospace 8 (12), 357. https://doi.org/10.3390/aerospace8120357.

Zhou, H., Zhang, S., Peng, J., et al., 2021b. Informer: beyond efficient transformer for long sequence time-series forecasting. In: Proceedings of the AAAI Conference on Artificial Intelligence, pp. 11106–11115. https://doi.org/10.1609/aaai.v35i12.17325.

Zhu, J., 2017. Research on Fault Diagnosis and Condition Monitoring of Wind Gearbox Based on Improved Ensemble Learning Algorithm. Shanghai DianJi University.