Journal of Computer Science and Technology Studies
ISSN: 2709-104X
DOI: 10.32996/jcsts
Journal Homepage: www.al-kindipublisher.com/index.php/jcsts

JCSTS
AL-KINDI CENTER FOR RESEARCH AND DEVELOPMENT

| RESEARCH ARTICLE

Research on the Application of Machine Learning in Fault Diagnosis Technology

Chen Yao

College of Horticulture, Northwest A&F University, Yangling, 712100, China

| ABSTRACT

Modern industrial equipment is increasingly characterized by large scale, high complexity, and continuous operation. Traditional fault diagnosis methods exhibit significant limitations when dealing with massive volumes of data and various nonlinear relationships. Machine learning, which is capable of autonomously discovering patterns and identifying structures from data, provides a novel pathway for fault diagnosis. This paper systematically reviews the fundamental theories, commonly used algorithms, and practical application scenarios of machine learning in fault diagnosis, with particular attention to the applications of supervised learning and unsupervised learning in the fields of machinery, electric power, and transportation. It then discusses the current challenges, including poor data quality, limited model interpretability, and insufficient fault samples. Finally, future development directions are explored from the perspectives of explainable artificial intelligence, generative models, edge computing, and physics-informed data fusion. Overall, machine learning is driving fault diagnosis from a "repair-after-failure" paradigm toward "predictive maintenance." However, for robust deployment in industrial environments, further efforts are still required in algorithm transparency, data efficiency, and cross-domain adaptability.

| KEYWORDS

Machine learning; fault diagnosis; deep learning; predictive maintenance; explainable artificial intelligence

| ARTICLE INFORMATION

ACCEPTED: 20 March 2026 **PUBLISHED:** 09 April 2026 **DOI:** 10.32996/jcsts.2026.8.5.13

1. Introduction

Most modern industrial equipment operates continuously over long periods. Under such conditions, unexpected failures of critical components may not only interrupt production lines but also lead to major accidents. For example, transformer faults in substations may cause regional power outages, while engine failures in aircraft or automobiles may result in serious safety incidents. Therefore, the rapid and accurate detection of equipment faults is of critical importance.

Traditionally, fault diagnosis has mainly relied on expert experience or physical model-based inference. However, the accumulation of experience requires substantial time and often lacks rigorous scientific support. Although physical models may perform well in relatively simple systems, they become difficult to establish for modern complex equipment characterized by diverse monitoring data and varying operating conditions. In addition, the computational cost is often prohibitively high, and the resulting models may lose effectiveness when applied to different scenarios.

The current situation is markedly different. With the decreasing cost of sensors and the widespread deployment of the Internet of Things, massive volumes of operational data are continuously accumulated during equipment operation. These data contain real-time information that can be used to determine whether faults are present. Transforming such data into intelligent and actionable insights has become a mainstream research direction. Against this background, machine learning has emerged, enabling computers to perform self-analysis, discover underlying patterns, and learn to predict future states. The application of this paradigm to fault diagnosis is of great significance.

The introduction of machine learning into fault diagnosis is feasible from both theoretical and practical perspectives. On the one hand, it can automatically extract features from data and identify weak fault signals that are imperceptible to human observation.

On the other hand, it enables real-time monitoring and early warning, transforming maintenance strategies from "repair after failure" or "periodic maintenance" to "predictive maintenance," thereby extending operating time and reducing maintenance costs [1,2].

This paper focuses on the specific applications of machine learning in fault diagnosis. Section 2 introduces the fundamental theories and common algorithms. Section 3 elaborates on practical applications in machinery, electric power, and transportation. Section 4 discusses the current challenges. Section 5 presents future research directions, followed by the conclusions.

2. Fundamental Theories of Machine Learning

2.1 Basic Concepts and Paradigms

Machine learning is a branch of artificial intelligence whose core idea is straightforward: instead of relying on manually programmed rules, data are used to "train" models so that they can autonomously perform analysis and prediction. According to different training paradigms, machine learning is generally categorized into three types: supervised learning, unsupervised learning, and reinforcement learning.

Supervised learning is currently the mainstream approach. It requires labeled data, such as samples indicating normal states and fault conditions. The model learns the mapping relationship between inputs and outputs and can subsequently predict new samples. In fault diagnosis, supervised learning is commonly used for fault classification or equipment life prediction. As long as sufficient normal and fault samples are available, classifiers can be trained to identify specific fault categories.

In unsupervised learning, the data are unlabeled, and the model primarily relies on autonomous structure discovery. Common applications include clustering, dimensionality reduction, and anomaly detection. In industrial scenarios where labeled samples are scarce, unsupervised learning plays an important role by first learning the characteristics of normal data and then identifying abnormal points, which is highly valuable for early warning.

Reinforcement learning is a more intelligent paradigm in which an agent interacts with the environment and learns strategies based on reward signals. In fault diagnosis, it is mainly applied to maintenance strategy optimization. However, it is still in the exploratory stage and has not yet been widely adopted.

2.2 Commonly Used Algorithms

2.2.1 Support Vector Machine

Support Vector Machine (SVM) is a classical classification algorithm in statistical learning theory. Its core idea is to identify an optimal "decision boundary" in the feature space that separates different types of data as much as possible while maximizing the margin on both sides of the boundary. If the data are not linearly separable in the original space, kernel functions can be employed to map them into a higher-dimensional space, where linear separation becomes feasible.

SVM is highly robust when handling small-scale, high-dimensional datasets. In fault diagnosis, the extracted features of vibration signals are often of relatively high dimensionality, while the number of fault samples is limited, making SVM particularly suitable for this task. Numerous studies have demonstrated significant performance in bearing fault classification and gearbox condition recognition using SVM. However, its performance is sensitive to parameter settings, and both the selection of kernel functions and the configuration of penalty factors require careful tuning.

2.2.2 Artificial Neural Networks and Deep Learning

Artificial neural networks are inspired by the biological mechanism of neuronal signal transmission. Input data are propagated layer by layer through the network to generate outputs. A typical neural network consists of an input layer, hidden layers, and an output layer. The weights are adjusted through the backpropagation algorithm so that the output progressively approaches the true values.

Deep learning refers to neural networks with multiple hidden layers. Owing to their powerful feature learning capability, deep learning methods have had a profound impact on the field of fault diagnosis. Convolutional Neural Networks (CNNs) are particularly effective for processing images and grid-structured data such as two-dimensional time–frequency maps, enabling the automatic learning of local features. Recurrent Neural Networks (RNNs) and their variant, Long Short-Term Memory (LSTM), are suitable for time-series data and can capture temporal variation patterns in monitoring signals. More recently, Transformer models, driven by self-attention mechanisms, have demonstrated outstanding performance in sequence modeling and have gradually been introduced into fault diagnosis tasks.

Deep learning enables end-to-end fault detection, eliminating the need for labor-intensive manual feature engineering by directly extracting useful fault information from raw signals layer by layer. Multiple studies have shown that, in rotating machinery fault diagnosis, deep neural networks often achieve higher accuracy than traditional methods, especially under complex operating conditions and in scenarios involving multiple coupled faults [3,4].

2.2.3 Decision Trees and Ensemble Learning

A decision tree resembles a step-by-step decision-making process in which conclusions are reached through hierarchical judgments. This method is intuitive and easy to understand; however, a single decision tree is prone to instability and may lose accuracy when the dataset changes.

Ensemble learning addresses this limitation by combining multiple decision trees to improve robustness through collective decision-making. Random Forest is a representative example in which multiple decision trees are constructed simultaneously, with each tree trained on randomly sampled data and features, and the final result determined by voting. This approach is highly practical in industrial data processing.

Gradient Boosting Decision Tree (GBDT) and its advanced variants, such as XGBoost and LightGBM, are also widely used ensemble algorithms. These methods employ sequential training to progressively correct the errors of preceding models and have demonstrated excellent performance in numerous industrial data competitions.

2.2.4 Other Algorithms

In addition to the above methods, clustering algorithms such as K-means and DBSCAN also play an important role in unsupervised fault identification. Autoencoders, as another type of unsupervised neural network, can learn low-dimensional representations of data, and large reconstruction errors often indicate potential abnormalities. Recently, Generative Adversarial Networks (GANs) and Variational Autoencoders (VAEs) have attracted considerable attention. These models are capable of generating highly realistic fault samples and are particularly useful for addressing data imbalance problems [5].

3. Applications of Machine Learning in Fault Diagnosis

3.1 Mechanical System Fault Diagnosis

Rotating machinery is one of the most common types of industrial equipment, including bearings, gearboxes, motors, and turbines. Once faults occur, vibration signals usually exhibit corresponding changes, making vibration analysis a core technique in mechanical fault diagnosis.

Traditional vibration analysis mainly relies on signal processing for feature extraction, such as Fourier transform for spectrum analysis and wavelet analysis for time–frequency representations. With the emergence of machine learning, these extracted features can be mapped to fault categories. Taking rolling bearings as an example, researchers collect vibration signals under different conditions, including normal operation, inner-race faults, outer-race faults, and rolling element faults. Time-domain statistical features (e.g., peak value, root mean square, and kurtosis) and frequency-domain features are then extracted and fed into SVMs or neural networks for classification.

Deep learning has further automated this process. CNNs can directly process raw vibration signals or two-dimensional time–frequency maps for autonomous feature learning. The team led by Chen Xuefeng at Xi'an Jiaotong University pioneered a “sparse sensing” diagnostic method capable of capturing weak fault features from complex vibration signals under varying rotational speeds. This method has been deployed in tens of thousands of wind turbines across hundreds of wind farms nationwide. It successfully provided early warning of initial pitting in the gearbox gears of two 6.7 MW wind turbines in the Hami wind farm in Xinjiang, where severe speed fluctuations made the fault undetectable by traditional methods [6].

For more complex systems such as gearboxes and engines, fault types are diverse and signal transmission paths are complicated, making a single sensor often insufficient. In such cases, multi-sensor information fusion has become a research hotspot, where vibration, acoustic emission, temperature, and other data sources are integrated to improve diagnostic reliability.

3.2 Power System Fault Diagnosis

Power systems contain a large number of critical assets, including generators, transformers, transmission lines, and switchgear, all of which directly affect grid stability. Their monitoring data include voltage, current, temperature, partial discharge, and dissolved gases in oil, characterized by large volumes, high dimensionality, and complex relationships.

Transformer fault diagnosis is a representative application scenario. Dissolved Gas Analysis (DGA) of transformer oil can reflect internal faults, as different fault types generate different gas combinations. Traditional methods mainly rely on expert experience and ratio-based criteria, such as the IEC three-ratio method. Machine learning, however, can learn the complex relationships between historical DGA data, fault patterns, and gas concentrations. Algorithms such as neural networks and random forests often outperform traditional methods in transformer fault classification.

Machine learning can also be applied to transmission lines and power generation equipment. Faults such as generator stator winding insulation aging and rotor inter-turn short circuits can be identified through the analysis of voltage/current harmonics and vibration signals. In wind farms, anomaly detection models built on SCADA data can help identify potential issues in converters and pitch control systems in advance.

3.3 Transportation System Fault Diagnosis

Modern transportation systems are equipped with multiple sensors, providing a solid data foundation for condition monitoring and fault diagnosis. In the field of rail transit, the team led by Xu Peng at Beijing Jiaotong University developed an intelligent rail diagnostic analysis system capable of real-time monitoring of minute rail deformations. By integrating high-precision positioning and high-resolution identification technologies, the system can capture millimeter-level changes in tracks, including slight deformations caused by repeated wheel loading and displacements induced by temperature variations. Although these changes are difficult to detect in daily operations, they pose major risks to high-speed rail systems. The system can detect deformation points along 300 km of railway lines within one minute, automatically generate defect distribution maps and hierarchical warning reports, and reduce the original manual inspection time from three days to two hours. It has already been deployed in 11 railway administrations [7].

Switch machines are key devices in rail transit signaling systems, and their failures directly affect operational safety. The team at Alstom used deep learning to analyze the power signals of switch machines, enabling accurate identification of fault types such as jamming, abnormal friction, and power supply issues, with an accuracy exceeding 99.99%. This method requires only a single-channel power signal and demonstrates strong generalization and scalability [8].

In the automotive field, engine management systems, braking systems, and battery management systems all incorporate online diagnostic functions. Machine learning-based remaining useful life prediction models can estimate the safe operating duration of critical components based on historical usage data and real-time monitoring information, thereby providing support for preventive maintenance.

3.4 Other Industrial Scenarios

In process industries such as petrochemicals and metallurgy, production processes are continuous and highly complex, and faults propagate rapidly. Traditionally, multivariate statistical methods such as Principal Component Analysis (PCA) and Partial Least Squares (PLS) have been widely used for process monitoring. Machine learning can further improve the sensitivity to weak faults.

The aerospace field imposes the most stringent requirements on fault diagnosis. Failures in critical systems such as flight control systems, engines, and landing gear are often catastrophic. Therefore, diagnostic systems must be highly reliable while also remaining interpretable and easy to understand. Hybrid methods that combine physical models with data-driven approaches have become the mainstream direction in this field.

4. Advantages and Challenges

4.1 Benefits Brought by Machine Learning

Automatic feature learning. In the past, feature design largely relied on manual engineering, which required substantial domain expertise and repeated trial-and-error efforts. Machine learning models can automatically extract hierarchical features from raw data and identify weak fault patterns that are difficult to observe directly.

Capability to handle complex relationships. Industrial systems are often nonlinear, time-varying, and strongly coupled among variables, making them difficult to describe using traditional linear models. Through multilayer nonlinear transformations, deep neural networks can approximate highly complex functional relationships and adapt to the complexity of real operating conditions.

Capability for future prediction. Machine learning can not only determine the current health state but also predict potential future faults based on historical data and evolving trends, thereby transforming maintenance from “repair after failure” to “preventive intervention in advance.”

4.2 Challenges in Engineering Applications

Poor data quality. Data collected from industrial sites are often incomplete or even erroneous. Sensor failures and communication interruptions can both severely affect data quality. Therefore, data preprocessing is indispensable, although it usually incurs high time costs.

Scarcity of fault samples. In practical operation, equipment remains in a normal state most of the time, while fault samples—especially those corresponding to early-stage faults—are extremely scarce, leading to severe class imbalance in the training data. As a result, models tend to be biased toward the majority class (normal condition) and perform poorly on minority classes (fault conditions). Although generative models such as Generative Adversarial Networks can synthesize fault samples, the realism and diversity of the generated data still require further validation.

Limited model interpretability. Deep learning models are often regarded as “black boxes,” making it difficult to understand the basis of their decisions. This issue is particularly critical for safety-critical equipment. Recently, research on explainable artificial intelligence (XAI) has attracted significant attention, aiming to address this challenge either through post hoc interpretation or by directly constructing inherently interpretable models so that the basis of diagnostic results can be clearly understood [1].

Limited generalization capability. Models that achieve high accuracy in laboratory settings may fail when deployed in real industrial environments. Data distributions vary across different equipment, operating conditions, and life-cycle stages, which may cause model degradation. Transfer learning and domain adaptation technologies are being actively explored as potential solutions.

Insufficient computational resources. Deep models usually involve a large number of parameters and high computational complexity, placing strict demands on hardware resources. Therefore, model compression, lightweight design, and edge computing deployment have become current research hotspots.

5. Future Development Trends

5.1 Explainable Artificial Intelligence

As machine learning models are increasingly applied in safety-critical domains, interpretability has become indispensable. Explainable artificial intelligence mainly follows two directions. One is ante hoc interpretability, in which inherently interpretable models are constructed, such as attention mechanisms that explicitly indicate which parts of the input the model focuses on. The other is post hoc interpretability, where trained models are analyzed using methods such as gradient attribution and class activation mapping to reveal which features contribute most to the output. In the future, interpretable models integrated with physical knowledge are expected to gain wider acceptance [1].

5.2 Generative Artificial Intelligence and Data Augmentation

Generative models, including Generative Adversarial Networks, Variational Autoencoders, and diffusion models, provide new opportunities for addressing the scarcity of fault samples. These methods can learn the data distributions of both normal and fault samples and generate realistic synthetic fault samples, thereby improving the balance of training datasets. Generative models can also be used for data augmentation by transforming original data into more diverse training samples, which further enhances model generalization capability. In the future, generative AI is expected to be integrated with digital twin technology to simulate various fault modes and degradation processes in virtual environments, thereby providing rich training data for diagnostic models [5].

5.3 Edge Intelligence and Real-Time Diagnosis

With the continuous improvement of edge computing capabilities, deploying diagnostic models on edge devices close to data sources has become an important trend. This approach reduces data transmission latency and bandwidth pressure while enabling real-time online monitoring and early warning. Model lightweighting techniques, such as knowledge distillation, network pruning, and quantization, allow complex models to run efficiently on resource-constrained edge hardware. In the future, cloud–edge collaborative intelligent diagnostic architectures are expected to become increasingly common, where the edge side is responsible for rapid real-time decision-making, while the cloud side handles complex model training and global optimization.

5.4 Physics-Informed Integration

Purely data-driven methods often neglect physical mechanisms, whereas purely physics-based models struggle to cope with system complexity. Physics-Informed Neural Networks (PINNs) incorporate physical equations into the loss function as constraints, enabling models to conform not only to data characteristics but also to physical laws. This hybrid approach combines the strengths of both paradigms: physical laws ensure the fundamental rationality of the model, while data-driven learning captures the details of real operating conditions. In fault diagnosis, integrating prior knowledge such as equipment dynamic equations and fault mechanisms with machine learning models is expected to improve both diagnostic reliability and generalization capability.

5.5 Large Language Models and Knowledge Integration

Large Language Models (LLMs) have become mainstream in natural language processing and are now gradually extending into fault diagnosis applications. These models are capable of processing multimodal data, such as maintenance text records, voice reports, and sensor measurements, thereby enabling cross-modal information fusion. They can also leverage strong knowledge understanding capabilities to assist in fault reasoning and maintenance decision-making. However, LLMs still face several challenges in industrial applications, including insufficient domain knowledge, limited real-time performance, and uncertain reliability. Integrating LLMs with knowledge graphs, where structured domain knowledge guides reasoning processes, represents a promising direction worthy of further exploration.

6. Conclusion

This paper systematically discusses the fundamental theories of machine learning in fault diagnosis, including supervised learning and unsupervised learning, as well as mainstream algorithms such as Support Vector Machines, neural networks, and ensemble learning methods. Application scenarios in machinery, electric power, transportation, and process industries are further elaborated through representative practical cases. The analysis of advantages and challenges demonstrates the significant strengths of data-driven methods in improving diagnostic efficiency, accuracy, and predictive capability. At the same time, issues such as inconsistent data quality, scarcity of fault samples, limited model interpretability, weak generalization capability, and insufficient computational resources remain key directions for future breakthroughs.

Looking ahead, explainable artificial intelligence will make models more transparent and trustworthy; generative models will alleviate data imbalance; edge intelligence will enable real-time online monitoring; physics-informed integration will improve diagnostic reliability; and large language models will open new paradigms for intelligent interaction. As these frontier technologies continue to mature and converge, machine learning will play an increasingly important role in fault diagnosis, driving industrial equipment maintenance from passive response to proactive prediction, and from experience-based decisions to data intelligence, thereby providing strong support for the safe, efficient, and intelligent operation of industrial systems.

Funding: This research received no external funding.
Conflicts of Interest: The authors declare no conflict of interest.

References

[1] Li, T., Chen, J., Liu, T., Sun, C., Zhao, Z., Chen, X., & Yan, R. (2026). Explainable artificial intelligence based intelligent fault diagnosis: A systematic review from applications to insights. *Reliability Engineering & System Safety, 267*, 111935. https://doi.org/10.1016/j.ress.2025.111935
[2] *AI-driven predictive maintenance and fault diagnosis: Challenges and future directions in Industry 4.0*. (2025). Zenodo.
[3] *A review of fault diagnosis methods: From traditional machine learning to large language model fusion paradigm*. (2026). *Sensors, 26*(2), 702.
[4] *A systematic review of anomaly and fault detection using machine learning for industrial machinery*. (2026). *Algorithms, 19*(2), 108.
[5] *A comprehensive review on generative artificial intelligence models for fault diagnostics and prognostics of rotating machineries*. (2025). *Engineering Applications of Artificial Intelligence*.
[6] *Science and Technology Daily*. (2025, June 25). Chen Xuefeng: "Listening to wind power equipment through sparse sensing." *Xi'an Jiaotong University News*.
[7] *China Science Daily*. (2025, November 26). The "smart brain" capturing millimeter-level hidden hazards in railway tracks. *Beijing Jiaotong University News*.
[8] Di Santi, E., Ci, R., Lefebvre, C., Mijatovic, N., Pugnaloni, M., Brown, J., Martín, V., & Saiah, K. (2025). *Scalable, technology-agnostic diagnosis and predictive maintenance for point machine using deep learning* (arXiv:2508.11692). arXiv. https://doi.org/10.48550/arXiv.2508.11692