Proceedings of CONF-MLA 2025 Symposium: Intelligent Systems and Automation: AI Models, IoT, and Robotic Algorithms
DOI: 10.54254/2755-2721/2025.LD29657

Predictive Maintenance of Equipment Driven by Machine Learning

Yuhao Yi

College of Mechanical and Electrical Engineering, Hebei Agricultural University, Baoding, China
3444143867@qq.com

Abstract. Industry is the leading sector of the national economy, and equipment maintenance, as a key link within it, ensures the stability of industrial production, improves production efficiency, and is crucial to industrial development. This paper studies the current application status and future development of machine learning-based equipment predictive maintenance in mechanical manufacturing through literature review and case analysis. The research reveals that machine learning algorithms, such as supervised learning, unsupervised learning, and deep learning, can significantly enhance the accuracy of predictive maintenance. Case studies on medical equipment using Weibull Distribution, new energy vehicles using Convolutional Neural Networks (CNN), and concrete-filled steel tubular shear walls using Random Forest demonstrate the effectiveness of these algorithms in predicting failures and optimizing maintenance strategies. The findings suggest that integrating machine learning with predictive maintenance can reduce costs, improve efficiency, and enhance safety. Future work should focus on multi-technology integration and industry-specific customization to further advance the field.

Keywords: Machine Learning, Predictive Maintenance, Weibull Distribution, Convolutional Neural Network (CNN), Random Forest

1. Introduction

Equipment maintenance is a critical part of industrial production, directly affecting the operational duration and work efficiency of equipment. To avoid machine shutdowns during production, it is necessary to accurately predict machine lifespan and timely identify and handle equipment failures. Maintenance strategies fall into three types.

Breakdown Maintenance is a simple maintenance strategy. Equipment is shut down and parts are replaced only when a failure occurs. When a single part fails suddenly, this strategy may increase the failure rate of related parts and raise production costs.

Preventive Maintenance is a regularly implemented maintenance strategy, carried out according to preset cycles or experience. While it offers predictability, it can lead to increased costs due to over-maintenance.

Predictive Maintenance is a maintenance strategy that uses predictive tools to forecast failures and remaining useful life. This strategy requires continuous monitoring of the machine's operation

© 2026 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2755-2721/2025.LD29657

process and status, and needs to combine historical data processing, integrity analysis, statistics, and engineering reasoning. It can predict early equipment failures and provide solutions.

Traditional predictive maintenance relies on physical models, which require complex formulas and in-depth engineering background. It tends to ignore data patterns, resulting in low prediction accuracy. Driven by Industry 4.0 and 5G technology, modern industrial systems can collect massive amounts of data. Predictive maintenance systems can build and train models based on data, adjust strategies dynamically, significantly optimize maintenance costs, failure frequency, downtime, and spare parts management, and also improve production output, safety, and profits. Among them, machine learning in the field of artificial intelligence has become the core technical support for developing intelligent prediction algorithms.

This paper first reviews the existing literature to organize the current research status and identify the core controversies in the field of machine learning. It pinpoints the deficiencies in existing research. Based on this foundation, the paper carries out case analysis in conjunction with practical cases from the field of equipment maintenance. This analysis serves to verify the effectiveness of predictive maintenance methods driven by machine learning. Ultimately, the paper draws targeted conclusions and outlines future development directions. It provides references for the in-depth development of machine learning in predictive maintenance, offers insights for industrial enterprises to optimize their equipment maintenance in the future, and puts forward suggestions for potential technical issues that may arise in practical applications.

2. Classification and application characteristics of core machine learning algorithms

Supervised learning includes linear regression, support vector machines (SVM), decision trees, random forests, and neural networks.

Linear Regression establishes a linear relationship based on the historical data of equipment, uses key features of machinery to predict the remaining useful life (RUL) or failure probability of machines, and its results have high interpretability. SVM excels in processing data with high-dimensional features. It accurately distinguishes the operational status of equipment by establishing an "optimal classification hyperplane" and is highly suitable for scenarios with small sample sizes. Decision Tree adopts "if-else" tree-like logic to gradually classify equipment features. It can intuitively judge whether equipment is faulty, and understanding its decision-making process does not require relevant professional knowledge.

Random Forest is an ensemble algorithm based on multiple decision trees. It reduces the overfitting risk of a single decision tree through voting among multiple trees, handles multi-dimensional features of equipment more stably, and has both high accuracy and strong robustness. Neural Network is a computational model designed to mimic the structure of human brain neurons. It uses connections and calculations between multi-layer nodes to automatically learn patterns from data and complete relevant predictions [1].

Unsupervised learning includes clustering analysis and principal component analysis (PCA). Clustering Analysis groups a large amount of unlabeled data according to their similarity, making data within the same group as similar as possible and data from different groups as dissimilar as possible. PCA compresses high-dimensional data to low-dimensional space while preserving key original data. It simplifies the amount of computation and removes redundant information at the same time.

Deep learning encompasses several advanced algorithms, including convolutional neural networks (CNN), long short-term memory networks (LSTM), and generative adversarial networks (GAN). CNN extracts spatial or temporal features from the data of the object to be tested, so as to

DOI: 10.54254/2755-2721/2025.LD29657

realize early identification of potential failures or future risks. LSTM is a special recurrent neural network. It solves the problem that ordinary RNN cannot store information for a long time through a gating mechanism and can effectively process data that depends on sequences [2]. GAN is a deep learning model composed of a "generator" and a "discriminator". Through adversarial training between the two, the generator learns to generate simulated data highly similar to real data [3].

Combining machine learning with equipment predictive detection significantly enhances the efficiency and accuracy of analysis. Machine learning algorithms can uncover failure patterns and rules that traditional detection methods often miss, providing more scientific and reliable data for mechanical design, manufacturing, and maintenance. The strong data analysis and processing capabilities of machine learning can solve the difficulties and challenges encountered in traditional equipment detection, including strong subjectivity relying on humans, difficulty in capturing dynamic working conditions, and difficulty in discovering hidden defects. Through machine learning, the accuracy of fault detection can be improved, dependence on expert experience can be reduced, and fault early warning capabilities can be enhanced. Figure 1 illustrates the current application status of machine learning in equipment predictive maintenance by systematically summarizing existing research and practical application cases.

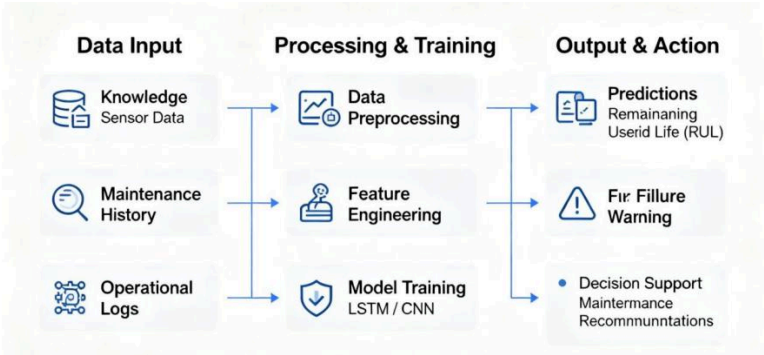


Figure 1. Application of machine learning in failure analysis [4]

3. Practical case application and analysis

3.1. Case 1: Weibull Distribution-based predictive maintenance for medical equipment

Based on the fault data of medical equipment, Weibull Distribution is used to fit the fault intervals of each medical equipment with days as the basic unit. Screening is conducted with Kolmogorov-Smirnov (K-S) test, and the probability of each medical equipment failing within a fixed number of days in the future is predicted. The maintenance cycle of each equipment can be obtained by referring to the average fault interval of each equipment.

When the failure rate of some equipment in the next 30 days is higher than a certain threshold, predictive maintenance is required, and this threshold can be adjusted by the business department according to the actual situation. Therefore, in practical application, it is necessary to extract the latest fault data of equipment from the database at a fixed time every day for recording to ensure the accuracy of predictive maintenance [5].

Solidifying the predictive model based on Weibull Distribution into the medical equipment maintenance management system is the core path to realize the transformation of equipment management from "passive response" to "active prevention", and its core value is reflected in the full-process intelligent closed-loop design [6].

DOI: 10.54254/2755-2721/2025.LD29657

In the data integration stage, this system connects multi-source information such as equipment operation data and fault records, constructs a dynamic fault prediction database, provides a real-time updated "data foundation" for the model, and ensures that the prediction basis is consistent with the actual status of the equipment.

In the model training and deployment stage, through the built-in parameter adaptive mechanism, real-time iteration and lightweight embedding of the Weibull Distribution model are realized, enabling the system to output accurate failure probability based on dynamic data and complete the leap from "offline fitting" to "online decision-making".

3.2. Case 2: fault diagnosis technology for new energy vehicles

The intelligent diagnosis technology based on machine learning theory constructs an accurate and efficient diagnosis model by learning and training the collected massive fault data. Compared with the traditional diagnosis method that relies on DTC to judge faults and is prone to missed detection or false detection, this diagnosis technology applies algorithms such as Convolutional Neural Network (CNN) in deep learning. It extracts fault information contained in signals, such as time-series signals of battery charging and discharging status, and uses methods such as transfer learning to establish fault diagnosis models for different working conditions, greatly improving the efficiency and accuracy of fault detection [7].

To enable the intelligent diagnosis system to have autonomous learning and adaptive capabilities, this technology adopts a lightweight online learning algorithm. When the vehicle status or use environment changes, the diagnosis system can also adjust adaptively to maintain a high level of diagnosis. The application of intelligent diagnosis technology can significantly improve the efficiency and reliability of fault detection for new energy vehicles and reduce the rate of false and missed alarms [8]. This technology subverts the traditional DTC fault code diagnosis mode. With machine learning as the core, it trains the diagnosis model through massive fault data to replace the rule judgment of traditional DTC, transforming from "passive matching" to "active learning". It uses Convolutional Neural Network (CNN) to extract fault information such as time-series signals during battery charging and discharging, adapts to different working conditions through transfer learning, and realizes adaptability through lightweight online learning, establishing an efficient diagnosis model that can significantly improve the efficiency and reliability of fault detection for new energy vehicles and reduce the rate of false and missed alarms.

3.3. Case 3: prediction of failure modes of concrete-filled steel tubular shear walls based on machine learning

To accurately predict the failure modes of concrete-filled steel tubular shear walls, this study collected 107 sets of test data to establish a database, used the SMOTE algorithm to improve the problem of unbalanced samples, and compared 6 types of machine learning algorithms to determine that the random forest model has the optimal performance (the average F1-score before and after SMOTE processing is 0.87 and 0.92 respectively).

Analysis through the SHAP method found that the shear-span ratio λ is the core factor affecting the failure mode: the higher the λ, the higher the probability of flexural failure; the lower the λ, the higher the probability of shear failure; when λ is moderate ($1 \leq \lambda \leq 2$), the probability of flexure-shear failure increases.

The higher the web reinforcement index, the lower the probability of flexure-shear failure; the higher the steel reinforcement index of the edge member steel tube and the axial compression ratio,

DOI: 10.54254/2755-2721/2025.LD29657

the higher the probability of flexure-shear failure. At the same time, it is clarified that the prediction of flexure-shear failure is the most difficult, and the key parameter intervals of the three failure modes are given to provide references for engineering design [9].

This study focuses on predicting the failure modes of concrete-filled steel tubular shear walls. It collects 107 sets of test samples in advance to establish a database, uses the SMOTE algorithm to improve the problem of unbalanced samples, and after comparing 6 types of machine learning algorithms, determines the random forest model as the algorithm with the optimal performance (the average F1-score before and after SMOTE processing is 0.87 and 0.92 respectively).

Through analysis with the SHAP method, it clarifies that the shear-span ratio is the core influencing factor and concludes that a high shear-span ratio tends to cause flexural failure and a low shear-span ratio tends to cause shear failure. The study also points out the difficulty in predicting flexure-shear failure and gives the key parameter intervals of the three failure modes, making the prediction of the failure modes of concrete-filled steel tubular shear walls more accurate.

4. Challenges and future development directions

The challenges faced in the development of mechanical predictive maintenance include the following aspects.

In terms of data, there is a shortage of high-quality data. In actual working scenarios, the environment and noise will interfere with sensor collection. The data formats and sampling procedures of different equipment and sensors also vary, leading to great difficulties in integration.

Due to the scarcity of fault samples, the proportion of normal data and fault data is unbalanced, making the model tend to predict towards the normal side and making it difficult to identify early fault signals.

In terms of implementation, the problems of technology landing cost and technology compatibility are prominent. Small and medium-sized enterprises lack unified data planning, and old equipment leads to weak data collection capabilities. Moreover, the system must be deployed around the existing operation and maintenance system, resulting in high initial investment in hardware and training costs.

Future development directions include integrating technology with cutting-edge technologies such as artificial intelligence, big data, the Internet of Things, and edge computing. The Internet of Things allows for interconnection of equipment, collecting massive data, and providing rich data resources. Artificial intelligence conducts intelligent diagnosis and accurate prediction, while edge computing improves real-time data processing performance.

Refinement and personalized customization are also essential for meeting the needs of different industries. For instance, the aerospace industry focuses on fault prediction of key components, while the automobile manufacturing industry focuses on production line equipment maintenance to improve efficiency and product quality.

5. Conclusion

Compared with breakdown maintenance and preventive maintenance, machine learning-driven predictive maintenance can rely on massive data to improve the accuracy of fault prediction, reduce costs, and improve efficiency. This paper presents the features of algorithms like supervised learning and unsupervised learning, validates their efficacy using three different case types—medical equipment (Weibull Distribution), new energy vehicles (CNN), and concrete-filled steel tubular

DOI: 10.54254/2755-2721/2025.LD29657

shear walls (random forest). It also highlights the need for future promotion of multi-technology integration and industry-specific customization.

Through literature review and case analysis, this paper sorts out the context of core research. However, there are shortcomings in the breadth and depth of literature coverage, focusing mainly on the achievements of domestic core journals and paying less attention to the cutting-edge trends of international top journals. Although the effectiveness of relevant methods has been initially verified, the scope of case analysis can be further expanded.

References

[1] Shanghai East Hospital (East Hospital Affiliated to Tongji University). (2024). A multi-purpose segmented wearable electronic intelligent bandage: CN202410965039.0 [Invention Patent Publication].
[2] Quan, X. W., Wang, L. L., Han, W. D., Ren, G. F., et al. (2025). Research on two-stage IoT asset identification model. Network Security and Data Governance.
[3] Li, Z. H., You, Y. L., Zhao, Z. H., et al. (2024). Application of artificial intelligence technology in the field of failure analysis. Journal of Aeronautical Materials, 44(5), 1-16.
[4] Wu, S. C., Hu, Y. N., Yang, B., et al. (2021). A review of characterization methods for defects in additive manufacturing materials and structural integrity assessment. Journal of Mechanical Engineering, 57(22), 3-34.
[5] Yang, L. Q., Wu, M. W., Su, L. N., et al. (2025). Research on predictive maintenance method of medical equipment based on Weibull distribution. Chinese Digital Medicine, 20(4), 115-120.
[6] Wang, J. F. (2024). Discussion on the application of digital technology in power communication optical cable resource management system. Consumer Electronics.
[7] Li, Y. J. (2025). Research progress of fault diagnosis technology for new energy vehicles. Automobile Maintenance Technician, (4), 27-28.
[8] Li, Z. M. (2024). Research on the application of big data analysis technology in communication network operation and maintenance. Telecom Power Technologies.
[9] Ma, G., & Wang, Y. (2022). Prediction and interpretation of failure modes of concrete-filled steel tubular shear walls based on machine learning. Earthquake Engineering and Engineering Vibration, 42(3), 143-152.