Research on Equipment Failure Prediction Based on Machine Learning Models

Dong Yu[a]
Jinan Foreign Language School, Jinan, Shandong, 250000, China

Keywords: Equipment Failure Prediction, Machine Learning, AMIRA Model.

Abstract: In the field of industrial production, the failure prediction of instruments is very important. The defect of the device effectively prevents problems such as the stay of the product caused by the failure and the decline in efficiency, to improve the stability of the product. On the other hand, broken equipment can eliminate potential accident risk, reduce maintenance costs, and prevent product expansion from being maintained. This article summarizes several new ideas for error prediction of devices, including deep learning-based techniques. The Bible learns from massive data and conducts error prediction through a deep learning model, comparing the predicted moral values with the true moral values. As such, it can accurately predict errors and monitor the device in real time using electronics on the Internet. After collecting the data, we conduct data analysis through various websites to obtain the predicted results. In addition, the interpretation methods of multitime crush data are reviewed to make error prediction. Using the decision tree, the relationship between the theme and the result is verified. This article explains in detail the content of each method and the specific applications or benefits of various techniques in industrial production.

1 INTRODUCTION

Equipment failures can significantly affect the reliability and productivity of plants, as they are an essential part of industrial production. Instead of predicting a malfunction, choose the effects that the device has, avoid broken components, and know that the devices are not enough, let it do a timely repair to reduce losses and further reduce maintenance costs. In addition, equipment failure can detect possible safety risks in time, manage possible safety accidents early and efficiently, prevent the deaths of employees, and at the same time reduce losses for companies. Collecting predictable data on operational failures would help companies develop science-based maintenance plans, equip them with the right maintenance resources, improve equipment and asset management, and make companies more competitive.

Early traditional time series models, such as the AMIRA model, showed their advantage in predicting equipment failure. Chen & Xing (2024) used a time series prediction model to prepare, test, and evaluate the M6000-8s Ethernet router, and found that most of the actual data would be within the confidence interval, so the predictive accuracy of the ARIMA model was high. A comparison of the compatibility value and the true data value revealed that the model had a small error in the predictive value. The estimated projected RMSE is 211.69%, and the MAPE is approximately 10.58%. This situation shows the reliability of predicting a network device failure.

However, there are certain limitations in predicting a shortage of conventional equipment. The failure model often only applies to specific types of equipment and operating environments and lacks universality. The ability to predict new tools is significantly reduced. In addition, traditional algorithms require more manual intervention, longer calculation periods, and tend to make mistakes. The quality of the data and the accuracy of the model are not high, resulting in significant reductions in accuracy. Therefore, it is important to have the right mindset and method to predict failure.

In recent years, a large number of researchers have adopted new machine learning techniques to predict tool failures, and better results have been achieved. In a study conducted by Yu (2025), the

[a] https://orcid.org/0009-0004-9316-1280

Yu, D.
Research on Equipment Failure Prediction Based on Machine Learning Models.
DOI: 10.5220/0013822000004708

In *Proceedings of the 2nd International Conference on Innovations in Applied Mathematics, Physics, and Astronomy (IAMPA 2025)*, pages 199-204
ISBN: 978-989-758-774-0

performance of three different deep learning models on the error detection task was compared, and the results showed that the 1D convolutional neural network had the best performance, with an accuracy of 98.3% and a training time of 24 minutes. The results show that the deep learning approach has a higher accuracy than ARIMA when performing this task and effectively provides early prediction of failures.

This article introduces the traditional methods of predicting equipment failures, namely the time series algorithm, and summarizes three new methods based on deep learning models, Internet of Things technology, and multi-time scale trend attention convolutional neural networks, and explores the diversification of equipment failure prediction forms.

2 CONSTRUCTION OF THE EQUIPMENT FAILURE PREDICTION MODEL

2.1 Based on the Time Series Algorithm

This is the traditional method for predicting equipment failures. Time series data are arranged in chronological order, and the overall trend is cyclical. The autoregressive Differential Moving Average (ARIMA) model can effectively handle time series data and accurately predict the failures of network devices. The autoregressive (AR) model can discover the relationship between perception and the value of past moments.

Chen & Xing(2024) created a network fault prediction model based on a time series algorithm using the ARIMA model. The ARIMA model consists of an autoregressive model (AR) and a moving average model (MA), and differential operations are added to ensure the stability of the data. Plot graphs, autocorrelation graphs, and partial autocorrelation graphs from the time series data for observation, and determine the appropriate autoregressive model. Residual analysis is used to determine the applicability effect and prediction accuracy of the model (residual plot, recovery correlation plot, residual distribution test, fluctuation of repeated sequences). To evaluate the advantages and disadvantages of the ARIMA model, Mean quadratic Error (MSE) and Mean Absolute Percentage Error (MAPE) were used.

A time series-based algorithm is a method that uses historical data to predict future failures. They are only applicable to short-term and medium-term prediction scenarios, as well as non-continuous regular data or historical data. Obviously, traditional prediction methods have certain limitations.

In the actual case, Chen & Xing (2024) conducted model training, testing, and evaluation on the ZTE M6000-8s Ethernet router. The data of one interface of the ZTE M6000-8s switch within the time range from December 1, 2022, to February 15, 2023, was obtained as the dataset, and the training set was the data with an average resampling of these data within 5 minutes. In the distribution map of the obtained original dataset, it can be seen that in the time series information, the difference order is 1, the mean of the difference series is 1993.22, the standard deviation of the difference series is 23,384,579, the number of observations per track is 21,000, and the number of observations after each difference track is 20,580. The above information is processed by ARIMA (9991, 1). After fitting, predict the future 1000 data points. It can be seen from the recorded data that most of the true values fall within the confidence interval (for example, 22 on December 1, 2022). At 10:00 on December 2, 2022, the true data value was 1307.916MB, the fitting value of the ARIMA model was 1250.714MB, and the residual was only -57.175. The true data value was 1017.228MB, the fitting value of the ARIMA model was 1024.858MB, and the residual was also only 7.630. The difference between the true values and the fitted values of the two is relatively small, indicating that the predictive performance of the model is excellent. It can be obtained through calculation that the RMSE of the predicted value is 211.69, and the MAPE is approximately 10.58%, with the error within a reasonable range. Therefore, the ARIMA model can be used for efficient feature extraction of time series data, thereby achieving more accurate equipment failure prediction.

2.2 Based on Deep Learning Models

2.2.1 Data Preprocessing and Model Selection

Deep learning is one of the implementation methods of artificial intelligence. It can be analogous to the way humans think, thereby possessing powerful feature extraction and modeling capabilities. Deep learning learns features in large amounts of data through models with multi-layer neural networks and extracts data features in an end-to-end manner, possessing powerful feature extraction and modeling capabilities.

When the device is in operation, it generates a large amount of data, including audio data, image data, and numerical data. It is necessary to clean the operation data of the device, remove outliers, map the data to the preset range, adjust the size and number of channels of the image data to meet the input requirements of the deep learning model, and maintain the stability of the data (Yu, 2025).

Yu (2025) mentioned four deep learning models - Convolutional Neural Network (CNN), Recurrent Neural Network (RNN), Long Short-Term Memory Network (LSTM), and Gated Recurrent Unit (GRU). When dealing with spatial structure data, convolutional neural networks are usually chosen. The feature extraction ability of the convolutional layer can help the model extract local features, and the downsampling ability of the pooling layer can help the model reduce dimensions and avoid overfitting to the data. Time-frequency graphs can also have such advantages. Recurrent neural networks are used to process sequential data (such as text and speech) and obtain the temporal dependencies of the data through recurrent connections. Long short-term memory networks are suitable for processing time series data, showing the changes of device status over time, and using memory units to record this information for later data processing. The gated loop unit is an improved version of RNN and is used in long-distance dependency problems.

The preprocessed data should be divided into the training set, the test set, and the validation set. During training, the model's performance is evaluated on the validation set. Deep learning models learn the features in the data, and the learned features expand successively from the low level to the high level. Convolutional neural networks can extract the detailed parts in data images, and high-level convolutional networks can learn the abstract features directly related to fault prediction. The Long Short-Term Memory network summarizes which key features will affect the operating status of the equipment by extracting the variation patterns and dependencies of the equipment's operating status at different times, thereby inferring the factors influencing equipment failures. (Wang, 2024).

2.2.2 Model Training, Feature Extraction, and Performance Evaluation

Deep learning models propagate their prediction results forward. The input data is passed layer by layer through the neural network. The data undergoes a linear transformation at each layer of the neural network until the prediction result is obtained at the output layer. Subsequently, the deep learning model will compare the differences between its own predicted results and the actual measured values, and calculate the difference between the predicted results and the actual results. Finally, starting from the output layer, calculate the gradient of the loss model to the model parameters layer by layer in reverse. Update the model parameters with the optimizer to gradually reduce the loss function, that is, continuously reduce the difference between the predicted situation and the true value (Wang, 2024).

The accuracy of the model trained by deep learning is greatly improved. It can be used to predict faults of devices that have not been touched before, output the predicted types of equipment faults, the probability of equipment failure occurrence, and the severity of occurrence, and ensure that the error is within the allowable range, greatly improving the efficiency and quality of equipment fault prediction.

The advantages of deep learning are mainly reflected in the extraction of equipment failure features. In traditional feature extraction methods, expert knowledge and manual design are indispensable. Deep learning models can learn independently and reduce manual intervention. For example, Yu (2025) summarized that when using stacked autoencoders for feature extraction in the case of bearing datasets, the accuracy rate was as high as 95.6%. Similarly, when using convolutional autoencoders and long short-term memory networks to extract features from the gear dataset and the electric shock dataset, respectively, the accuracy rates were 97.2% and 93.8%. Thus, it can be seen that the accuracy rate of feature extraction by deep learning models is very high. The commonly used deep learning fault diagnosis models also have extremely high accuracy rates. Yu (2025) mentioned that when one-dimensional convolutional neural networks, long short-term memory networks, and deep belief networks were diagnosed with fan vibration data, generator temperature data, and pump pressure data respectively, the accuracy rates were 98.3%, 96.5% and 94.7% respectively, and the accuracy rates were still very high.

2.2.3 Application Cases of Deep Learning in Fault Diagnosis

Wang (2024) obtained the prediction situations of LSTM, CNN and ARIMA models when conducting fault prediction for the power system in a certain area according to the above-mentioned methods. After analyzing and comparing the experimental results, it was obtained that the accuracy rates of the three were

85.2%, 82.6% and 75.8% respectively, the precision rates were 83.6%, 81.2% and 76.5% respectively, the recall rates were 87.1%, 84.5% and 73.2% respectively, and the F1-score were 85.3%, 82.8% and 74.8% respectively. In the comparison of the experimental results of the baseline model, the accuracy rate of the feature extraction model was 88.4%, the precision rate was 87.2%, the recall rate was 89.6%, and the F1-score was 88.3%. The accuracy rate of the attention mechanism model was 91.2%, the precision rate was 90.5%, the recall rate was 92%, and the F1-score was 91.2%. Experiments show that the feature extraction ability and attention mechanism of the deep learning model also have good stability based on ensuring accuracy in the fault prediction and analysis of power equipment, indicating that the deep learning algorithm can better analyze complex data.

In the fault prediction of data-driven machine tools, deep learning models analyze the data collected by the sensors on the machine tools. At this point, the convolutional neural network analyzes the time-frequency graph of the equipment vibration, extracts images to predict the faults of the equipment's main shaft, and conducts timely maintenance to reduce downtime. Zhang et al. (2023) collected data from several processing devices and set up three fault models for each device, namely the fault of the front end bearing roller of the spindle, the fault of the inner ring of the front end bearing of the spindle, and the fault of the rear end bearing roller of the spindle. The CYT9200 integrated vibration sensor was placed near the bearing housing of the equipment to obtain the time-domain graphs of the communication signals of the three devices under various faults. Through data preprocessing and normalization, and by applying the activation function and loss function, the schematic diagram of data iteration times - accuracy rate was obtained. The analysis revealed that the average accuracy rate of the support vector machine was 90.54%, the average accuracy rate of one-dimensional CNN was 95.68%, and the average accuracy rate of the proposed method was 98.9%. It can be seen from this that the feature extraction and classification ability of convolutional neural networks in deep learning models is stronger than that of other models.

In transformer fault prediction, the partial discharge, temperature, and vibration data of the voltage transformer are monitored first. After the deep learning model extracts the features and combines them with the vector machine, the potential faults of the transformer are diagnosed, and the fault hazards are dealt with in a timely manner (Zhang et al., 2025).

2.3 Based on Internet of Things Technology

2.3.1 Architecture and Implementation of IoT-Based Fault Prediction

The application of Internet of Things technology can achieve real-time monitoring of devices. The prediction results are real-time, more flexible, and efficient. Meanwhile, the application of Internet of Things technology for equipment failure prediction can enable real-time management of equipment distributed in different locations, thereby enhancing management efficiency.

The Internet of Things system is divided into the perception layer, the network layer, and the application layer. Among them, the perception layer collects the operation data of the equipment in real time through sensors, including temperature, vibration, pressure, and other sensors, to obtain data such as the temperature, pressure, and vibration frequency of the equipment. The network layer uses wireless communication technologies such as 4G, 5G, and Wi-Fi to upload the data obtained from the perception layer to the local server or the cloud, and stably transmit the data. In the application layer, data processing, analysis, and modeling are mainly carried out, and equipment failures are predicted through fault prediction models. (Zhang et al, 2025).

Through the research on equipment failure prediction, it is found that by using Internet of Things technology to collect data from equipment and transmit the data through the transportation module, the research on equipment failure prediction can be carried out in the intelligent network. The architecture and implementation method of using equipment failure prediction in intelligent networks were proposed. This scheme takes sensor technology as the basic means. It carefully selects the type and specific model of sensors based on the operating status of the equipment to be monitored and the monitoring parameters to be collected. The selected sensors are required to have good stability and anti-interference ability, and be able to stably collect data under complex electromagnetic environments and drastic changes in temperature and pressure. (Sun & Guo, 2024).

Next comes data transmission and communication technology. During short-distance transportation, wireless communication technology can be applied to form a local area network among multiple device nodes. Data is collected by sensors and transmitted to the local gateway, thereby achieving efficient, low-latency, and stable data transmission between sensors and between sensors and the local gateway. To reduce maintenance costs, low-power-consumption communication technology can be adopted. In sensor networks, there are a large number of battery-powered sensor nodes. Therefore, it is necessary to collect data efficiently and transmit it to the data center or the cloud for data computing. For multi-device connections and application scenarios with high real-time requirements, networks with lower latency, higher bandwidth, and larger connection capacity need to be adopted (Liu,2024).

2.3.2 Practical Applications of IoT in Fault Diagnosis

When predicting faults, deep learning models are combined, and convolutional neural networks are used to process image and video data. It can automatically extract the key features in images and videos, and process the data layer by layer in the convolutional layer, pooling layer, and fully connected layer. Among them, the convolutional layer has convolution kernels, which are used to capture the detailed parts in the image, including the wear and tear of the equipment, etc. The pooling layer reduces the dimension of the collected data, compresses the features, reduces the amount of data, and only retains the key features. The fully connected layer is responsible for feature integration, which is used for classification and regression. Finally, it determines whether the device will fail. If a failure occurs, it determines the type and severity of the failure. The long short-term memory network is used to process time series data to effectively handle the long-term dependencies in this data. The long short-term memory network contains memory units that filter the information to be remembered, forget the useless information, and summarize the changing trend of the device's operating status more accurately. After storing the analysis results of historical data, the long short-term memory network can predict the operation of the equipment in the future period and troubleshoot possible faults (Yu, 2025).

During the operation of elevators, Internet of Things technology can be applied for real-time monitoring and fault prediction. Liu (2025) selected multiple elevators in A certain high-rise building and compared and evaluated three elevator operation states. The Internet of Things technology was used to monitor the experimental group. Compared with Group A and Group B, traditional sensing technology and image recognition methods were adopted, respectively. Then, four types of operation abnormalities were selected and combined with experimental and historical fault data (door not closing tightly (A), operation overload (B)). Speed anomaly (C) and motor overheating (D), and then select a part of the data different from the above from the experimental recorded data to obtain the comparison of fault prediction delay time of different monitoring methods. The amounts of abnormal data in abnormal types A, B, C, and D were 210.36, 315.48, 422.15, and 512.04, respectively, while the amounts of data detected in the experimental group were 210.12,315.25,421.86, and 511.79, respectively. The data volumes monitored in Group A were 165.42,245.31,310.22, and 378.58, respectively, and those monitored in Group B were 145.26,223.18,289.45, and 365.33, respectively. The data shows that the error between Group A and Group B is relatively large. The monitoring results of Group B have a significant deviation in the fault types of motor overheating and abnormal speed. Therefore, the traditional methods have obvious deficiencies in the prediction and monitoring accuracy of equipment failures. In contrast, the monitoring accuracy of Internet of Things (IoT) technology when devices fail is much stronger than that of traditional methods, indicating the reliability of 2.3 Internet of Things technology.

2.3.3 Advantages and Challenges of IoT in Fault Prediction

Wang & Wang (2025) adopted Internet of Things technology to implement remote monitoring and fault prediction for the coal mining machines of a certain coal mining enterprise. They upload the data collected by the sensors to the cloud and, in combination with machine learning models, identify the information processed by big data to predict faults. The results show that the monthly unplanned downtime was 40 hours before deployment and 10 hours after deployment, a decrease of 30 hours, with an improvement rate of -75%. The average failure response time decreased from 60 minutes to 15 minutes, a reduction of 75%. The average failure repair time dropped from 5 hours to 2 hours, with an improvement rate of -60%. The equipment utilization rate was 70% before deployment. After deployment, it was 85%, with an overall increase of 21%. The maintenance cost was significantly reduced, from 1.2 million per month to 800,000 per month, a decrease of 33%. The accuracy rate of fault prediction increased by 85%, and the data collection coverage rate rose from 60% to 95%, with an improvement of 58%. It can be seen from the data that the unplanned downtime, fault response, and repair time of the coal mining machine have been significantly shortened after the deployment of Internet of Things technology. While the maintenance cost has decreased, the accuracy of fault prediction and the coverage rate of data collection have further improved, indicating the significant advantages of Internet of Things technology in equipment fault prediction.

Although Internet of Things technology has shown significant advantages in equipment failure prediction, there are still challenges in data privacy protection, system stability, and cross-platform compatibility. Zhang et al. (2025) pointed out that Internet of Things systems may face the risks of signal interference and data loss in high-concurrency data processing and complex environments, affecting the accuracy of prediction and the reliability of the system.

3 CONCLUSION

This paper takes equipment failure prediction as the research object and summarizes the relevant research results of equipment failure prediction from three aspects: based on time series algorithms, deep learning models, and Internet of Things technology.

As a traditional method for equipment failure prediction, the time series algorithm can better ensure the accuracy of the prediction results. However, it has many limitations and is only applicable to continuous regular data, when there is a small amount of historical data, or in medium and short-term prediction scenarios. When facing more complex or high-dimensional data, the time series algorithm cannot guarantee the accuracy of the prediction. The application of deep learning technology in equipment failure prediction can significantly improve the efficiency and effectiveness of fault detection. It has a strong ability to process high-dimensional data and can effectively analyze complex and high-dimensional data in images. Deep learning models can discover various examples existing in complex and high-dimensional data in images, which are difficult to detect by traditional detection methods.

The Internet of Things technology and predictive maintenance of equipment failures can effectively promote the transformation of industrial operation and maintenance models to intelligent models. Based on the analysis of the predictive maintenance mode for industrial equipment failures and combined with the powerful sensor network of the Internet of Things technology, this paper analyzes a large number of opportunities for real-time high-coverage equipment operation status perception existing in the transformation of the predictive maintenance mode for industrial equipment failures to the intelligent mode at present.

REFERENCES

Chen, S., Xing, C., 2024. Network device fault prediction model based on time series algorithm.

Liu, H., 2025. Real-time monitoring and fault prediction of elevator operation status based on Internet of Things technology.

Liu, L., 2025. Optimization strategy for power equipment condition monitoring and predictive maintenance technology.

Sun, Y., Guo, L., 2024. Research on fault detection and diagnosis of wind turbine equipment based on sensor technology and I-LSTM algorithm.

Wang, S., Wang, S., 2025. Research on remote monitoring and fault diagnosis system for coal mining machines based on Internet of Things.

Wang, W., 2024. Deep learning-based fault diagnosis and intelligent prediction algorithm for power equipment.

Yu, C., 2025. Deep learning-based equipment fault diagnosis and prediction technology.

Zhang, F., Fu, D., Yang, Q., Zhang, K., 2025. Review on fault diagnosis methods for large-scale wind power transmission chains under complex working conditions. *Mechanical Transmission*, 49(04), 156–168.

Zhang, L., Tang, D., Zhu, H., Liu, C., Wang, Z., 2023. Convolutional neural network-based machine tool health management system.

Zhang, Q., Guo, C., Yu, Y., Huang, J., Lian, Y., 2025. Design of medical device fault prediction system based on Internet of Things and neural network.