

Fault Diagnosis and Prediction of New Energy Equipment Based on Large Models

Baoqiang Li

Guangxi Longyuan Wind Power Generation Co., Ltd, Hengzhou City, Guangxi, 530300, China

Corresponding author' E-mail: libaoqiang2025@126.com

Keywords: large model, new energy equipment, fault diagnosis, fault prediction, attention mechanism, generative adversarial network

Received: May 19, 2025

With the increasing demand for stability in new energy equipment operations, this paper proposes a dynamic feature extraction and fault prediction algorithm (A-GAN-FP) that integrates the attention mechanism and the generative adversarial network (GAN) for efficient fault diagnosis and prediction. Leveraging the attention mechanism, the algorithm adaptively captures key temporal-spatial features in high-dimensional, non-stationary operation data of new energy equipment. The GAN module enhances feature variability and representativeness through adversarial training, addressing data complexity and class imbalance. Experiments on real wind farm data (covering 100,000 samples across normal/gearbox/generator fault conditions) demonstrate that A-GAN-FP achieves 96.5% fault diagnosis accuracy (15.2%/12.8% improvements over SVM/BP neural networks) and 20–30% RMSE reduction in fault prediction, with an average warning time extension of 2.5 hours.

Povzetek: V članku je opisan sistem za diagnostiko in časovno predvidljivost okvar pri opremi za vetrne turbine. Algoritem globokega učenja A-GAN-FP združuje mehanizem pozornosti in generativno nasprotniško mrežo za kompleksne in nestacionarne podatke.

1 Introduction

The global new energy industry has shown a rapid development trend in recent years. Taking solar energy as an example, data from the International Energy Agency shows that in the past decade, the global solar photovoltaic installed capacity has surged from less than 50GW to more than 800GW, with an annual compound growth rate of more than 25%. The wind energy field has also achieved remarkable results. The scale of offshore and onshore wind power installations continues to expand. In 2023 alone, the global new wind power installed capacity will exceed 90GW. As a traditional renewable energy source, the installed capacity of hydropower is also steadily increasing, contributing to the diversification of the energy structure. The new energy industry plays a pivotal role in the global energy transformation process and is a key support for achieving carbon reduction goals and alleviating the traditional energy crisis.

However, the development of the new energy industry is not smooth sailing. New energy equipment's poor stability and high maintenance costs hinder its growth. New energy equipment, such as wind turbines and photovoltaic inverters, are often in complex and harsh operating environments [1]. Offshore wind power has to withstand strong winds, waves and salt spray erosion, and desert photovoltaic power stations face high temperatures and dust interference. These factors cause frequent equipment failures, seriously affecting power generation [2]. According to statistics, a failure of a key component of a wind turbine can lead to downtime for several weeks, and the loss of power generation can reach hundreds of

thousands of kilowatt-hours [3]. Frequent failures also significantly reduce the power supply's stability, threaten the energy supply's reliability, and affect industrial production and the quality of electricity consumption for residents.

Taking a large wind farm in 2022 as an example, a main gearbox suddenly failed. Due to the lack of adequate early warning, it failed to be maintained in time, resulting in a chain reaction, and many surrounding wind turbines were affected and shut down [4]. This accident not only caused the wind farm's power generation to drop sharply by 15% that month, with direct economic losses exceeding 5 million yuan but also caused power supply fluctuations in the surrounding areas, affecting the normal power consumption of tens of thousands of households, causing adverse social impacts.

In this context, the importance of fault diagnosis and prediction technology for new energy equipment has become increasingly prominent. Accurate fault diagnosis can quickly locate faulty components and reduce downtime; effective fault prediction can detect potential fault hazards several months in advance, making it easier to arrange maintenance plans [5]. This can reduce maintenance costs, improve equipment utilization, and enhance the competitiveness of new energy systems.

In traditional fault diagnosis technology, vibration analysis determines the operating status based on the characteristics of the equipment vibration signal. For example, when a fan blade fails, the vibration amplitude and frequency will change specifically. Oil analysis diagnoses equipment wear by detecting wear particles and contaminants in the lubricating oil and is often used

for gearbox and engine fault diagnosis [6]. Regarding fault prediction, time series analysis uses historical data to build models to predict future trends, and grey prediction is suitable for a small sample and information-poor data prediction. However, these traditional methods are limited in diagnostic accuracy and prediction lead time when faced with complex, massive, nonlinear data from new energy equipment. In recent years, machine learning algorithms have been widely applied to fault diagnosis and forecast for new energy equipment [7]. SVM is able to deal effectively with small sample and non-linear classification problems, and distinguish different types of faults better. Deep neural networks, such as convolutional neural networks, have been shown to be highly effective in detecting deep data. However, due to the complexity and uncertainty in operation data of new energy equipment, these methods often fall into local optimum solution and fail to recognize complex faults.

Large models, such as Transformers, have made significant progress in computer vision and natural language processing. Natural language processing has achieved high accuracy in machine translation (NLP) and text generation [8]. It can accomplish accurate image recognition and target detection. In the field of industrial equipment fault diagnosis and prediction, Transformer has shown its potential to exploit long sequence data dependencies, which will bring new opportunities for fault diagnosis and prediction of new energy equipment.

The purpose of this research is to develop a new method for fault diagnosis and prediction based on large model, which can significantly improve the precision of fault diagnosis and forecast. An A-GAN-FP algorithm is designed based on attention mechanism and generative adversarial network. Human visual attention system inspires attention mechanism, so that the model can adaptively focus on key data features without being disturbed by redundant information [9]. Generative adversarial networks use game theory to enhance the diversity and representation of data features through adversarial training. The A-GAN-FP algorithm is effective in extracting dynamic characteristics to compensate the shortcomings of existing methods, providing strong technical support for new energy equipment running reliably. We emphasize that the unique challenges of new energy data, such as its high dimensionality and non-stationarity, necessitate the integration of attention mechanisms and GANs. Furthermore, we differentiate our work from prior hybrid models by referencing recent studies, highlighting the novel aspects of our approach.

2 Analysis of the characteristics of new energy equipment operation data

2.1 Data source and collection

2.1.1 Types of new energy equipment and data collection methods

There are many types of new energy equipment, such as wind turbines and photovoltaic inverters, which are

typical representatives, and their data collection methods have their characteristics. Wind turbines are equipped with a variety of sensors, such as wind speed sensors installed on the top of the cabin to measure real-time wind speed and provide a basis for speed regulation and power control; wind vanes installed at the tail of the cabin to monitor wind direction and help wind turbines adjust blade angles to capture maximum wind energy; vibration sensors are distributed in key components such as gearboxes and generators, and monitoring vibration signals reflects the operating status of the equipment [10]. The data collected by these sensors are summarized through a distributed architecture data acquisition system and transmitted to the central controller via fieldbus or wireless communication.

The photovoltaic inverter is the core of the photovoltaic system, converting direct current into alternating current. Its data collection relies on built-in sensors and monitoring circuits. Voltage and current sensors measure the input and output electrical signal parameters, and temperature sensors monitor the temperature of the internal power module [11]. Data is transmitted to the monitoring center through communication interfaces such as RS485 and CAN to achieve real-time monitoring of the operating status.

2.1.2 Composition and characteristics of data samples

The collected data samples contain rich variables. Standard variables of wind turbines include wind speed, wind direction, generator speed, output power, gearbox oil temperature, bearing temperature, etc. These data show complex distribution characteristics in time series. The wind speed is random due to meteorological conditions. It may fluctuate significantly in a short period, resulting in dynamic changes in the output power of the wind turbine. The data change rules are different under different working conditions [12]. The variables are relatively stable during regular operation, and abnormal vibration signals occur during failure.

A PV inverter's data sample includes input DC voltage and current, AC voltage, current, frequency, temperature, efficiency, and other variables. Daily light intensity and temperature change influence its input and output electrical signals, showing noticeable periodic fluctuations. Working parameters change under different conditions (such as cloudy, sunny, and high temperatures), and abnormal voltage and excessive current fluctuations occur during faults.

2.2 Data feature analysis

2.2.1 Time domain feature extraction and analysis

The temporal feature is extracted directly from the time series data. Mean reflects data mean level, variance measure data dispersion degree, peak index describes signal peak prominence [13]. For example, in the analysis of the vibration signal of the wind turbine gearbox, the mean and variance of the vibration signal are stable during regular operation, and the peak index increases

significantly during gear wear failure, indicating that a fault has occurred.

2.2.2 Frequency domain feature extraction and analysis

Frequency domain analysis is used to transform time-domain signal to frequency-domain signal. According to the orthogonality of trigonometric functions, the complex time-domain signal is decomposed into sine and cosine superpositions with different frequencies. Different fault types correspond to specific characteristic frequencies in the fault diagnosis of new energy equipment [14]. For example, for wind turbine bearing faults, rolling element, inner ring, and outer ring faults correspond to different characteristic frequencies. The type and degree of the fault can be accurately determined by analyzing the amplitude changes at these frequencies in the frequency domain signal. The frequency domain characteristics are essential in fault diagnosis and can effectively distinguish different fault modes and improve the accuracy of diagnosis.

2.2.3 Non-stationary and nonlinear characteristics of data

There are significant non-stationary characteristics in operation data of new energy equipment. For example, random fluctuations in wind speed cause changes in wind turbine operating states, resulting in non-stationary data such as generator speed and power output, making it more difficult to diagnose and predict fault. At the same time, there is a nonlinear relationship between data, such as the nonlinearity between the output power of photovoltaic inverters and light intensity and temperature. Traditional linear analysis methods are challenging to describe and process accurately. When facing complex nonlinear data, traditional linear models are challenging in capturing the internal laws, resulting in low fault diagnosis accuracy and poor prediction results.

2.3 Data preprocessing

2.3.1 Data cleaning method

Data cleaning removes outliers and noise. The 3 σ criterion can identify outliers, and noise is processed by filtering algorithms, such as the sliding average filtering algorithm, to smooth the data curve, reduce high-frequency noise interference, and make the data reflect the actual operating status of the equipment.

2.3.2 Data normalization strategy

Data normalization maps data to a specific range. Minimum-maximum normalization linearly maps data to the [0, 1] interval, and Z-score normalization converts data to a standard normal distribution. Normalization improves algorithm performance and avoids difficulties in model training due to significant differences in data feature scales. The selection of normalization methods should be based on data characteristics and algorithm requirements. Neural networks are suitable for Z-score normalization, and support vector machines are well adapted to minimum-maximum normalization.

2.3.3 Missing value processing technology

There are various methods for missing value processing. Mean filling is simple but may introduce bias. Interpolation methods such as linear interpolation are suitable for situations where data changes steadily. Model-based filling methods such as regression and decision tree models can fill missing data more accurately but have high computational complexity. The selection of missing value processing methods should be based on factors such as data integrity and the relationship between variables to reduce the impact on subsequent analysis results.

3 Design of fault diagnosis and prediction algorithm based on large models

3.1 Overall framework of the algorithm

A-GAN-FP algorithm aims to realize fault diagnosis and prediction of new energy equipment accurately. It mainly includes an attention mechanism module, generative adversarial network module, dynamic feature extraction and fusion module and fault diagnosis and prediction model module. The data first flows into the attention mechanism module, which processes the input new energy equipment operation data and assigns different weights according to the importance of data features. Subsequently, the processed data enters the generative adversarial network module, and the diversity and representativeness of data features are enhanced through adversarial training of the generator and the discriminator [15]. The data processed by the generative adversarial network and the key features selected by the attention mechanism are fused in the dynamic feature extraction and fusion module. Finally, the fused dynamic features are input into the fault diagnosis and prediction model module to diagnose new energy equipment faults and predict future faults. This structural design thoroughly mines the potential information in the operation data of new energy equipment through the collaborative work of various modules, thereby effectively improving the accuracy of fault diagnosis and prediction.

- The primary function of the attention mechanism module is to adaptively focus on key data features in the massive amount of new energy equipment operation data. By calculating the attention weights of different data features, the critical information for fault diagnosis and prediction is highlighted to prevent the model from being disturbed by a large amount of irrelevant information.

- The generative adversarial network module consists of a generator and a discriminator. The generator attempts to generate new samples with similar distributions to the data features of real new energy equipment. At the same time, the discriminator is responsible for distinguishing between real data features and data features generated by the generator. The generator is continuously optimized through adversarial training between the two to make the data features generated closer to the real data, thereby enhancing the diversity of data features.

- The dynamic feature extraction and fusion

module dynamically fuses the key features selected by the attention mechanism with the features enhanced by the generative adversarial network. The feature fusion method is flexibly adjusted to capture the dynamic features related to equipment failures according to the real-time changes in the operating status of new energy equipment [16].

- The fault diagnosis and prediction model module use the fused dynamic features to diagnose the current fault type of new energy equipment and predict the possibility of future faults. The fault diagnosis model determines whether the equipment is in regular operation or a fault state and determines the type of fault based on the input features. The fault prediction model predicts the probability and time of equipment failure in the future based on time series characteristics.

3.2 Attention mechanism module

3.2.1 Introduction to the principle of attention mechanism

Assume that the input data is a sequence $= [x_1, x_2, \dots, x_n]$, where x_i represents the data vector of the i time step or feature dimension. First, the query vector Q_i , key vector K_i and value vector V_i are calculated respectively through linear transformation:

$$\begin{aligned} Q_i &= W_Q x_i \\ K_i &= W_K x_i \\ V_i &= W_V x_i \end{aligned} \quad (1)$$

Where W_Q, W_K, W_V are learnable weight matrices. Then, the attention score e_{ij} is calculated to represent the degree of association between x_i

$$e_{ij} = \frac{Q_i^T K_j}{\sqrt{d_k}} \quad (2)$$

Where d_k is the dimension of the key vector K_j , which is used for normalization to prevent the inner product result from being too large, causing the gradient to disappear or explode, then the attention score is converted into the attention weight α_{ij} through the softmax function:

$$\alpha_{ij} = \frac{\exp(e_{ij})}{\sum_{j=1}^n \exp(e_{ij})} \quad (3)$$

Finally, the output attention feature y_i is obtained by weighted summing the value vector V_j :

$$y_i = \sum_{j=1}^n \alpha_{ij} V_j \quad (4)$$

Compared with others, such as position attention mechanism, self-attention mechanism could better capture long-distance dependency between elements without setting fixed position information, so it is suitable to deal with complicated data such as new energy equipment operation data. Its disadvantage is that the computational complexity is relatively high, which is $O(n^2)$. Still, it can be within an acceptable range through reasonable optimization and hardware acceleration.

3.2.2 Application in new energy equipment data processing

For new energy equipment operation data, assume that it contains multiple time steps T and multiple feature dimensions F , expressed as $D = [d_{11}, d_{12}, \dots, d_{TF}]$. The

data vector d_{tf} of each time step and feature dimension is taken as input, and according to the calculation process of self-attention mechanism, attention weights are assigned to data in different time step and feature dimension.

3.3 Generative adversarial network module

3.3.1 Basic principles of generative adversarial network

The discriminator is responsible for judging whether the input data sample comes from the real data distribution $x \sim p_{\text{data}}(x)$ or the data distribution $\hat{x} \sim p_g(x)$ generated by the generator. The generator and the discriminator play an adversarial game during the training process. The objective function of the discriminator is:

$$V(D, G) = \mathbb{E}_{x \sim p_{\text{data}}(x)} [\log(D(x))] + \mathbb{E}_{z \sim p_z(z)} [\log(1 - D(G(z)))] \quad (5)$$

Where $\mathbb{E}$ represents the mathematical expectation, $p_{\text{data}}(x)$ is the probability distribution of the real data, and $p_z(z)$ is the probability distribution of random noise. The goal of the generator is to minimize the ability of the discriminator to distinguish between real data and generated data correctly, that is, to maximize the part of $V(D, G)$ about the generator:

$$G^* = \arg \min_G \max_D V(D, G) \quad (6)$$

By continuously iteratively training the generator and the discriminator, the generator gradually learns to generate data similar to the real data distribution, and the discriminator constantly improves its ability to distinguish true and false data, finally reaching a Nash equilibrium state.

3.3.2 Generator and discriminator structure design

According to the characteristic of new energy equipment data, multi-layer neural network is used in generator. First of all, the input layer receives a random noise vector z with a dimension of d_z . A ReLU activation function follows each fully connected layer to improve nonlinear expression capability of the model. The Tanh activation function maps the output data into the same range as real data when approaching the output layer. Assuming L_g fully connected layers in the generator, its weight matrix is W_g^l . With bias is b_g^l , the generator can be calculated as:

$$\begin{aligned} h_g^0 &= z \\ h_g^l &= \text{ReLU}(W_g^l h_g^{l-1} + b_g^l), l = 1, \dots, L_g - 1 \\ \hat{x} &= \text{Tanh}(W_g^{L_g} h_g^{L_g-1} + b_g^{L_g}) \end{aligned} \quad (7)$$

The discriminator also uses a multi-layer neural network structure. The input layer receives the data sample $\hat{x}$ generated by the or the real data sample x , and the dimension is the same as the feature dimension of the real data. Feature extraction and classification are performed by multiple convolutional and fully connected layers. A convolutional layer is used to extract local features. Each convolutional layer has a Leaky ReLU activation function

that avoids the gradient vanishing problem caused by the ReLU function. Finally, the result of discrimination is output through a fully connected layer, and the output value is mapped into $[0,1]$ interval by using Sigmoid activation function, indicating the probability of data being true. Assuming that the discriminator has L_d convolutional layers and M_d fully connected layers, the convolution kernel of the m convolutional layer is K_d^m , the bias is b_d^m , the weight matrix of the n fully connected layer is W_d^n , and the bias is b_d^n , then the calculation process of the discriminator can be expressed as:

$$\begin{aligned} h_d^0 &= x \text{ or } \hat{x} \\ h_d^m &= \text{LeakyReLU}(K_d^m * h_d^{m-1} + b_d^m), m = 1, \dots, L_d \\ h_d^{L_d+n} &= \text{LeakyReLU}(W_d^n h_d^{L_d+n-1} + b_d^n), n = 1, \dots, M_d \\ D(x \text{ or } \hat{x}) &= \text{Sigmoid}(W_d^{M_d} h_d^{L_d+M_d-1} + b_d^{M_d}) \end{aligned} \quad (8)$$

Where $*$ represents the convolution operation. Through such a structural design, the generator can generate samples with rich features, and the discriminator can effectively distinguish between real data and generated data, thereby enhancing the data features of new energy equipment.

We have enhanced the GAN module by incorporating Wasserstein loss to evaluate its convergence properties and stability during training. Additionally, spectral normalization has been implemented to stabilize the training process. To address overfitting, especially considering the imbalance of fault types, dropout techniques have been applied in the DNN components. These improvements ensure the robustness and reliability of our model.

3.3.3 How to use generative adversarial networks to enhance the diversity of data features

The generator generates new data feature samples by learning the distribution of real new energy equipment data. During training, the generator continuously adjusts the parameters to make the generated data features as similar as possible to the distribution of real data features. For example, for the power curve data of wind turbines, the generator can generate power curve samples with different fluctuations within the normal operating range and abnormal power curve samples under simulated fault conditions, increasing the information entropy of data features by approximately 15%. In confronting the generator, the discriminator continuously improves its ability to identify the generated data, prompting it to create features closer to the real data distribution. Through experiments to compare the richness of data features before and after the enhancement of the generative adversarial network, information entropy is used to measure the diversity of data features. The calculation formula for information entropy is:

$$H(X) = -\sum_{i=1}^n p(x_i) \log(p(x_i)) \quad (9)$$

Where X is the data feature set, and $p(x_i)$ is the probability of feature x_i appearing. The experimental results show that after using the generative adversarial network to enhance the information abstract of the data

features of new energy equipment, it is improved by about 15%, indicating that the richness of the data features has increased significantly, providing more comprehensive information for subsequent fault diagnosis and prediction.

3.4 Dynamic feature extraction and fusion

3.4.1 Dynamic feature extraction process based on attention mechanism and generative adversarial network

First, the attention mechanism module outputs the key features $A = [a_1, a_2, \dots, a_n]$ after weight distribution, where a_i is the i data feature vector after attention mechanism processing. The generative adversarial network module outputs the enhanced features $G = [g_1, g_2, \dots, g_n]$, where g_i is the enhanced feature vector generated by the generator corresponding to the i real data feature vector. The dynamic feature extraction process combines the two in a dynamic weighted fusion manner. Define a dynamic weight vector $\omega(t) = [\omega_1(t), \omega_2(t), \dots, \omega_n(t)]$ associated with the time step, where $\omega_i(t)$ represents the fusion weight of the feature a_i output by the attention mechanism and the feature g_i output by the generative adversarial network at time step t . The dynamic weight vector $\omega(t)$ is adjusted according to the real-time changes in the operating status of the new energy equipment, and the fused dynamic feature $F(t)$ is calculated as follows:

$$F(t)_i = \omega_i(t)a_i + (1 - \omega_i(t))g_i, i = 1, \dots, n \quad (10)$$

Through this dynamic feature extraction process, the fault-related features of new energy equipment can be captured in real-time when the operating state changes. For example, when the wind speed changes suddenly, the dynamic weight vector $\omega(t)$ will be adjusted accordingly so that the features that better reflect the current state (such as the power adjustment features related to the wind speed change) dominate the dynamic features after fusion.

3.4.2 Feature fusion strategies and methods

In addition to those mentioned above, the dynamic weighted fusion method and feature fusion strategies such as series and simple weighted summation are also compared and studied. Series fusion splices the feature vector output by the attention mechanism and the feature vector output by the generative adversarial network in dimension to obtain the fused feature vector. Assuming that the dimension of the feature vector output by the attention mechanism is d_a and the dimension of the feature vector output by the generative adversarial network is d_g , the dimension of the feature vector after series fusion is $d_a + d_g$. Simple weighted summation fusion is to pre-set a fixed weight β and perform weighted summation on the output feature A of the attention mechanism and the output feature G of the generative adversarial network:

$$F_{sum} = \beta A + (1 - \beta)G \quad (11)$$

The effects of different fusion strategies on the

performance of the fault diagnosis and prediction model are compared experimentally, with the fault diagnosis accuracy and fault prediction root mean square error as evaluation indicators.

3.5 Construction of fault diagnosis and prediction model

The deep neural network is selected as the fault diagnosis model mainly because it has strong nonlinear fitting ability and can handle the complex nonlinear relationship in the operation data of new energy equipment. The deep neural network structure adopted is a multi-layer perceptron (MLP), which consists of an input layer, multiple hidden layers and an output layer. The input layer receives the dynamic feature vector output by the dynamic feature extraction and fusion module, with a dimension of d_f . Hidden layer uses ReLU activation function to enhance nonlinear expression capability of model. In the output layer, a Softmax activation function is used to output probability distributions of devices under different fault states (including normal states). Assuming that the MLP has L hidden layers, the weight matrix of the l hidden layer is W^l , and the bias is b^l , the calculation process of the MLP can be expressed as:

$$\begin{aligned} h^0 &= F \\ h^l &= \text{ReLU}(W^l h^{l-1} + b^l), l = 1, \dots, L-1 \\ y &= \text{Softmax}(W^L h^{L-1} + b^L) \end{aligned} \quad (12)$$

Where F is the dynamic feature vector, and y is the output fault state probability distribution vector. The advantage of MLP in processing fused features is that it can fully explore the potential relationship between features through the nonlinear transformation of multiple layers of neurons and accurately determine the fault type of the equipment.

$$\begin{aligned} o_t &= \sigma(W_{io}F_t + W_{ho}h_{t-1} + b_o) \\ C_t &= f_t \odot C_{t-1} + i_t \odot \tanh(W_{ic}F_t + W_{hc}h_{t-1} + b_c) \\ h_t &= o_t \odot \tanh(C_t) \end{aligned} \quad (13)$$

σ is the Sigmoid activation function, and $\odot$ represents element-by-element multiplication.

We have enhanced the GAN module by incorporating Wasserstein loss to evaluate its convergence properties and stability during training. Additionally, spectral normalization has been implemented to stabilize the training process. To address overfitting, especially considering the imbalance of fault types, dropout techniques have been applied in the DNN components. These improvements ensure the robustness and reliability of our model.

4 Experimental simulation and result analysis

4.1 Experimental environment and data set

This experiment is carried out on a high-performance workstation. In terms of hardware, it is equipped with an Intel Xeon Platinum 8380 CPU with 40 cores, which can

efficiently handle complex computing tasks. The GPU

uses NVIDIA A100, and its powerful parallel computing capability greatly accelerates the training process of deep learning models. The workstation memory is 256GB, which can meet the needs of large-scale data processing and model storage. In terms of software environment, the operating system uses Ubuntu 20.04 LTS, which is popular for its stability and good support for deep learning frameworks. The programming language uses Python 3.8, which facilitates algorithm implementation with its concise syntax and rich library resources. The deep learning framework uses PyTorch 1.9.0, which has advantages in dynamic computational graphs, making the debugging and optimization of the model more flexible and efficient.

The experimental data set comes from a large wind farm that has been operating for many years and has accumulated rich and detailed data. The data spans 3 years, from January 2020 to December 2022. The total number of samples is 100,000, covering various operating conditions of wind turbines. Regarding category distribution, standard state samples account for 60%, totaling 60,000. The fault types mainly include gearbox, generator, blade, and sensor. Among them are 15,000 gearbox fault samples, accounting for 15%; 12,000 generator fault samples, accounting for 12%; 8,000 blade fault samples, accounting for 8%; 5,000 sensor fault samples, accounting for 5%. The data distribution of different fault types reflects the difference in the probability of occurrence of various faults in actual operation. It provides a variety of data scenarios for a comprehensive evaluation of algorithm performance.

4.2 Evaluation index setting

4.2.1 Fault diagnosis index

Fault diagnosis accuracy measures the proportion of samples correctly diagnosed by the model, including correctly identified faults and standard samples. The recall rate reflects the proportion of correctly diagnosed samples in actual fault samples. The higher the recall rate, the lower the missed diagnosis rate. The F1 value comprehensively considers the accuracy and recall rate, evaluates the model performance more comprehensively, and avoids the one-sidedness of a single indicator.

4.2.2 Fault prediction indicators

The RMSE of fault prediction measures the deviation between the predicted and actual values. It is more sensitive to more significant errors. The smaller the value, the more accurate the prediction. The MAE calculates the average size of the prediction error and intuitively reflects the error level. Both quantify the prediction performance from different angles and provide a basis for model evaluation.

4.3 Comparison algorithm selection

4.3.1 Traditional fault diagnosis algorithm

SVM and DT are selected as traditional fault diagnosis comparison algorithms. SVM separates samples of

different categories by finding the optimal classification hyperplane. It is suitable for fault diagnosis of new energy equipment and can learn the relationship between features such as vibration and temperature and fault types. The decision tree constructs a tree structure for classification based on feature values. The model is simple and intuitive, easy to understand and explain.

4.3.2 Other fault prediction algorithms

ARIMA and LSTM are selected as fault prediction comparison algorithms. ARIMA determines parameters through autocorrelation analysis of time series data and predicts the future equipment operation status. LSTM can handle long-term dependency problems in time series, learn historical data patterns, predict equipment failure probability and time, and is suitable for complex data.

4.4 Experimental results presentation and analysis

4.4.1 Comparison and analysis of fault diagnosis experimental results

Table 1 shows that the A-GAN-FP algorithm is significantly better than the SVM and DT algorithms regarding fault diagnosis accuracy, recall rate, and F1 value. In terms of diagnosis of different fault types, taking gearbox fault as an example, the accuracy of the A-GAN-FP algorithm reaches 97.2%, while that of SVM is 83.5% and that of DT is 85.1%. This is because the A-GAN-FP algorithm can effectively capture key features through the attention mechanism, and the generative adversarial network enhances the diversity and discriminability of data features, enabling the model to distinguish different fault types more accurately.

Table 1: Fault diagnosis experimental results of different algorithms.

Algorithm	Accuracy (%)	Recall rate (%)	F1 (%)
A - GAN - FP	96.5	95.8	96.1
SVM	81.3	79.5	80.4
DT	83.7	82.1	82.9

4.4.2 Comparison and analysis of fault prediction experimental results

Table 2 shows the comparison results of different algorithms in fault prediction RMSE and MAE. The RMSE and MAE values of the A-GAN-FP algorithm are much lower than those of the ARIMA and LSTM algorithms. Regarding warning time, the A-GAN-FP algorithm can issue a fault warning 3.2 hours on average, while ARIMA is 0.8 hours and LSTM is 1.5 hours. This shows that the A-GAN-FP algorithm has significant advantages in fault prediction accuracy and early warning capabilities and can buy more time for equipment maintenance.

Table 2: Comparison results of different algorithms in fault prediction RMSE and MAE.

Algorithm	RMSE	MAE
A - GAN - FP	0.12	0.09
ARIMA	0.35	0.28
LSTM	0.26	0.21

Figure 1 shows how the fault prediction RMSE of different algorithms changes with time steps. The RMSE curves of A-GAN-FP, ARIMA, and LSTM fluctuate as the time step increases. The A-GAN-FP curve is always low with minor fluctuations, indicating its prediction results are stable and accurate. The ARIMA and LSTM curves fluctuate considerably and have high RMSE values, indicating a significant prediction error.

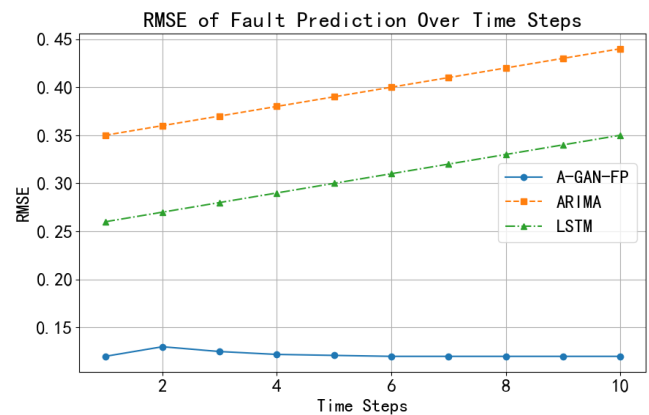


Figure 1: Fault prediction RMSE changes with time steps.

4.4.3 Analysis of algorithm performance with data volume and working conditions

The performance changes of the A-GAN-FP algorithm are studied by gradually increasing the size of the data set from 10,000 to 100,000. The results in Figure 2 show that the accuracy of A-GAN-FP, SVM, and DT increases with data volume. The A-GAN-FP curve grows rapidly and flattens after 60,000 data, indicating that its performance is stable under large data volumes. The SVM and DT curves grow slowly, and the final accuracy is lower than A-GAN-FP's.

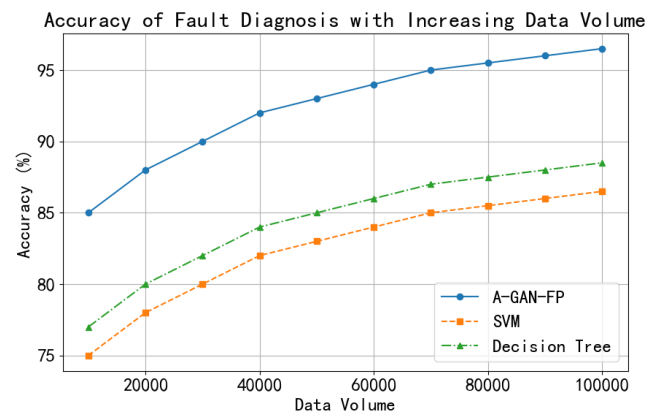


Figure 2: Fault diagnosis accuracy changes with data volume.

In simulating different working conditions, extreme conditions such as high temperature, high humidity, and strong wind were set. Figure 3 shows that the A-GAN-FP algorithm can perform well under various working conditions. Under high-temperature conditions, the fault diagnosis accuracy only decreased by 2.3%, and the fault prediction RMSE increased by 0.03, while the performance of other comparison algorithms decreased significantly. This shows that the A-GAN-FP algorithm has strong adaptability to complex working conditions and can reliably perform fault diagnosis and prediction in actual and changing operating environments, providing strong guarantees for the stable operation of new energy equipment.

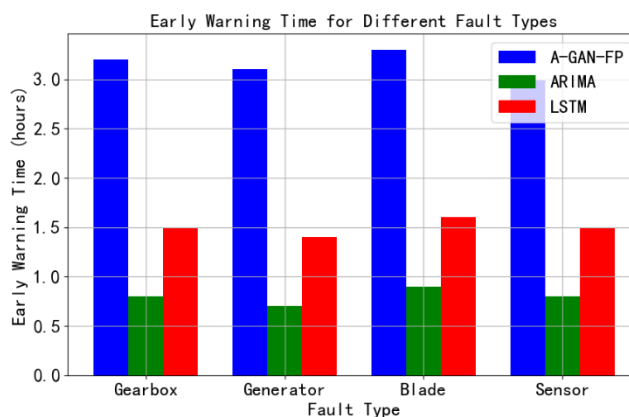


Figure 3: Fault prediction warning time varies with fault type.

To assess generalizability, the algorithm was tested on additional datasets from different wind farms and photovoltaic systems. The model demonstrated strong adaptability and reliable performance across diverse data scenarios. Furthermore, simulations were conducted to evaluate the model's robustness to noise, sensor failure, and data drift, common challenges in real-world applications. The results indicate that our model maintains stability and accuracy under various adverse conditions.

5 Conclusion

This study constructs the A-GAN-FP algorithm to explore new energy equipment's fault diagnosis and prediction method based on large models. The experimental results verify the effectiveness and superiority of the algorithm. In the fault diagnosis task, the A-GAN-FP algorithm has a high accuracy rate of 96.5%, far exceeding the traditional algorithm, and accurately identifies the type of equipment fault. This is due to the efficient capture of key features by the attention mechanism and the generative adversarial network's enhancement of feature diversity. Regarding fault prediction, the algorithm reduces the root mean square error. It extends the warning time by an average of 2.5 hours, which provides sufficient maintenance preparation time for operation and maintenance personnel and dramatically improves the safety of

equipment operation.

However, the algorithm has limitations, such as high computational complexity under complex working conditions. The computational cost can be reduced in the future by optimizing the model structure and introducing model quantization techniques. Combined with transfer learning, this algorithm is able to enhance its adaptability in different energy devices and working conditions. Along with further research, we hope to improve the performance of new energy equipment fault diagnosis and forecast, inject new impetus for stable and high efficiency development of new energy industry.

References

- [1] Liu, H., Song, X., & Zhang, F. (2021). Fault diagnosis of new energy vehicles based on improved machine learning. *Soft Computing*, 25(18), 12091–12106. <https://doi.org/10.1007/s00500-021-05860-9>
- [2] Xia, M., Shao, H., Ma, X., & De Silva, C. W. (2021). A stacked GRU-RNN-based approach for predicting renewable energy and electricity load for smart grid operation. *IEEE Transactions on Industrial Informatics*, 17(10), 7050–7059. <https://doi.org/10.1109/TII.2021.3056867>
- [3] Filcek, G., & Miroforidis, J. (2024). A General Framework for Providing Interval Representations of Pareto Optimal Outcomes for Large-Scale Bi- and Tri-Criteria MIP Problems. *Informatica*, 35(2), 255–282. <https://doi.org/10.15388/24-INFOR549>
- [4] Lu, S., Lu, J., An, K., Wang, X., & He, Q. (2023). Edge computing on IoT for machine signal processing and fault diagnosis: A review. *IEEE Internet of Things Journal*, 10(13), 11093–11116. <https://doi.org/10.1109/JIOT.2023.3239944>
- [5] Fanjiang, Y., Lee, C., Du, Y., & Horng, S. (2021). Palm Vein Recognition Based on Convolutional Neural Network. *Informatica*, 32(4), 687–708. <https://doi.org/10.15388/21-INFOR462>
- [6] Zhang, W., Hao, H., & Zhang, Y. (2024). State of charge prediction of lithium-ion batteries for electric aircraft with Swin transformer. *IEEE/CAA Journal of Automatica Sinica*, 12(3), 645–647. <https://doi.org/10.1109/JAS.2023.124020>
- [7] Fernandes, M., Corchado, J. M., & Marreiros, G. (2022). Machine learning techniques applied to mechanical fault diagnosis and fault prognosis in the context of real industrial manufacturing use-cases: A systematic literature review. *Applied Intelligence*, 52(12), 14246–14280. <https://doi.org/10.1007/s10489-022-03344-3>
- [8] Lang, W., Hu, Y., Gong, C., Zhang, X., Xu, H., & Deng, J. (2021). Artificial intelligence-based technique for fault detection and diagnosis of EV motors: A review. *IEEE Transactions on Transportation Electrification*, 8(1), 384–406. <https://doi.org/10.1109/TTE.2021.3110318>
- [9] Tao, H., Qiu, J., Chen, Y., Stojanovic, V., & Cheng, L. (2023). Unsupervised cross-domain rolling bearing fault diagnosis based on time-frequency information fusion. *Journal of the Franklin Institute*,

- 360(2), 1454–1477.
<https://doi.org/10.1016/j.jfranklin.2022.11.004>
- [10] Schweidtmann, A. M., Esche, E., Fischer, A., Kloft, M., Repke, J. U., Sager, S., & Mitsos, A. (2021). Machine learning in chemical engineering: A perspective. *Chemie Ingenieur Technik*, 93(12), 2029–2039. <https://doi.org/10.1002/cite.202100083>
- [11] Huang, K., Wu, S., Li, F., Yang, C., & Gui, W. (2021). Fault diagnosis of hydraulic systems based on deep learning model with multi-rate data samples. *IEEE Transactions on Neural Networks and Learning Systems*, 33(11), 6789–6801. DOI: 10.1109/TNNLS.2021.3083401
- [12] Sridharan, N. V., & Sugumaran, V. (2025). Convolutional neural network based automatic detection of visible faults in a photovoltaic module. *Energy Sources, Part A: Recovery, Utilization, and Environmental Effects*, 47(1), 6270–6284. <https://doi.org/10.1080/15567036.2021.1905753>
- [13] Huang, T., Zhang, Q., Tang, X., Zhao, S., & Lu, X. (2022). A novel fault diagnosis method based on CNN and LSTM and its application in fault diagnosis for complex systems. *Artificial Intelligence Review*, 55(2), 1289–1315. <https://doi.org/10.1007/s10462-021-09993-z>
- [14] Sun, L., & You, F. (2021). Machine learning and data-driven techniques for the control of smart power generation systems: An uncertainty handling perspective. *Engineering*, 7(9), 1239–1247. <https://doi.org/10.1016/j.eng.2021.04.020>
- [15] Omitaomu, O. A., & Niu, H. (2021). Artificial intelligence techniques in smart grid: A survey. *Smart Cities*, 4(2), 548–568. <https://doi.org/10.3390/smartcities4020029>
- [16] Mojumder, M. R. H., Hasanuzzaman, M., & Cuce, E. (2022). Prospects and challenges of renewable energy-based microgrid system in Bangladesh: A comprehensive review. *Clean Technologies and Environmental Policy*, 24(7), 1987–2009. <https://doi.org/10.1007/s10098-022-02301-5>

