Journal of Organizational and End User Computing
Volume 37 • Issue 1 • January-December 2025

Intelligent Monitoring of Industrial Equipment:

A Study on Fault Prediction Based on Deep Learning

Jingbo Song
Institute of Agriculture and Rural Development, Heilongjiang Provincial Academy of Social Sciences, China

Yingli Wu
https://orcid.org/0009-0003-3637-1639
School of Business Management, Jiaxing Nanhu University, China

Duo Zhao
Graduate School of Heilongjiang Academy of Social Sciences, China

Jingqi Li
https://orcid.org/0009-0002-3929-1696
Centre for Cultural and Media Policy Studies, The University of Warwick, UK

Liqi Ding
https://orcid.org/0009-0002-5313-0792
Department of Media and Communications,London School of Economics and Political Science, UK

ABSTRACT

Predictive maintenance is gaining increasing attention in the field of industrial equipment management as an effective strategy to enhance equipment reliability and reduce maintenance costs. Deep learning has become a focal point due to its exceptional ability to process time series data and recognize complex patterns. To address challenges related to accuracy and robustness in predicting equipment failures, this study proposes a novel model that combines deep reinforcement learning (DDPG) with gated recurrent units (GRU), alongside Bayesian Optimization for hyperparameter tuning. The DDPG component learns the dynamic interactions between actions and states, adapting to the specific characteristics of different devices. The GRU module is designed to capture temporal dependencies in sensor data.

KEYWORDS

Machine Learning, Data Analytics, Asset Reliability, Continuous Monitoring, Predictive Maintenance, Industrial Internet of Things (IIoT)

INTRODUCTION

With the continuous development of industrial equipment, especially in manufacturing, the complexity and diversity of equipment have been steadily increasing, leading to a gradual rise in the risk of equipment failures. Equipment failures not only affect production efficiency and product quality but can also result in significant economic losses (Namuduri et al., 2020). Therefore, improving equipment reliability and reducing downtime have become key challenges in industrial production

DOI: 10.4018/JOEUC.369157

(Li, Wang, & Wang, 2019). Traditional equipment fault detection methods typically rely on rules and manual experience, but as the variety of equipment and the volume of data increase these methods are gradually showing their limitations (Carvalho et al., 2019; Panchi et al., 2022). However, despite significant successes in various fields, the application of deep learning in industrial equipment monitoring still faces many challenges. Traditional models, in particular, often exhibit instability in performance and poor generalization ability when dealing with complex sequential data, which limits their applicability in practical industrial scenarios.

Among existing deep learning models, Recurrent Neural Networks (RNNs) are a classic sequence modeling method. By introducing a recurrent structure, RNNs can capture dependencies in sequential data more effectively (Q. Wang et al., 2020). RNNs have been widely used in industrial equipment monitoring, but they face the problem of vanishing gradients when processing long sequences, leading to the loss of long-term dependencies and impairing the model's performance (Zhu et al., 2022). Long Short Term Memory Networks (LSTMs) introduce a gating mechanism that successfully alleviates the vanishing gradient problem, making them more stable when handling long-term sequential dependencies (Jiang et al., 2022). However, LSTMs still face challenges, such as high computational complexity and long training times, especially when dealing with complex equipment states or large-scale sequential data, which limits their efficiency and applicability (Lai et al., 2019). Similar to LSTMs, temporal convolutional networks (TCNs) have made significant progress in processing long sequences in recent years (Li, Li, et al., 2019). TCNs extract information in the spatiotemporal domain through convolution operations, effectively overcoming the bottleneck in traditional RNNs and LSTMs when processing long sequences. With a larger receptive field, TCNs can better capture global patterns (Yueze et al., 2023). However, TCNs still face challenges in industrial equipment monitoring, such as sensitivity to sequence length and high computational complexity, which persist as efficiency issues in real-time monitoring and large-scale data processing (Yuan et al., 2021). Another common deep learning model is a graph convolutional network, which focuses on processing graph-structured data and is particularly useful in capturing complex relationships between equipment in industrial monitoring (Luo et al., 2022). By modeling the topological structure between equipment, graph convolutional networks help understand the evolution and mutual influence of equipment states (Y. Wang et al., 2020). However, graph convolutional networks still face problems in handling large-scale graph data, such as low computational efficiency and poor real-time performance, limiting their practical application in dynamic environments.

Although the aforementioned models have their advantages in industrial equipment monitoring, they still have significant limitations when faced with more complex industrial sequential data (Huang et al., 2025; N. Li et al., 2021; Zhang et al., 2025). Traditional methods often fail to effectively handle the diversity and variability of equipment states, especially in scenarios that require high real-time performance and multi-equipment collaboration (C. Li et al., 2021). The existence of these problems has prompted the exploration of a new approach that combines the advantages of different models to overcome the limitations of existing methods. Therefore, this article proposes the deep reinforcement learning–gated recurrent units (DDPG-GRU) model, which combines DDPG and GRU and introduces Bayesian optimization for hyperparameter optimization. This model aims to address the shortcomings of traditional models in handling complex sequential data and multi-equipment states. Compared with existing models, DDPG-GRU integrates the strengths of deep reinforcement learning and sequential modeling, providing higher accuracy and stability in equipment fault prediction.

The main contributions of this study are as follows:

1. By integrating DDPG and GRU, an innovative model is constructed to achieve more effective modeling of time series data.
2. Bayesian optimization for hyperparameter adjustment in introduced so that the model can better adapt to the specific requirements of different industrial equipment monitoring tasks.

3. This model achieves an accurate understanding of timing information and equipment status in industrial equipment monitoring, discovers potential faults in real time through predictive maintenance, and takes measures to improve equipment reliability and work efficiency.

Related Works

Industrial Equipment Fault Prediction

As a core concept in the research field, industrial equipment failure prediction aims to achieve predictive management of equipment maintenance by identifying potential failures in advance. In recent years, with the development of sensor technology and big data analysis, significant progress has been made in industrial equipment failure prediction (Leukel et al., 2021). Through in-depth mining of equipment operating data, researchers have proposed a variety of prediction models that are based on statistics, machine learning, and deep learning, which provide strong support for improving the reliability of industrial equipment.

Current research focuses mainly on improving the accuracy and generalization ability of predictive models. By integrating knowledge from different fields, researchers are committed to building a more comprehensive and efficient industrial equipment failure prediction system (Liu et al., 2019). In terms of application, these models are widely used in manufacturing, energy, transportation, and other fields, providing feasible solutions for fault prevention in actual production processes.

However, current research generally faces challenges, such as the fusion application of multisource heterogeneous data, the generalization ability of the model in complex scenarios, and the operability in actual industrial applications. In addition, there are other issues, including data quality, model interpretability, and modeling of complex device relationships, that also need to be resolved urgently (Scalabrini Sampaio et al., 2019).

Applications of Deep Learning

Recent research in deep learning has focused on optimizing model architectures and expanding their range of applications. Improved models, such as LSTM networks with gating mechanisms and TCNs for processing long sequences, have been developed to enhance performance. Deep learning has shown notable success in areas like industrial equipment health monitoring, fault detection, and production process optimization (Dalzochio et al., 2020). In particular, advancements in model structure have led to improvements in time series modeling and feature extraction, which are highly beneficial for industrial applications. Despite these successes, deep learning still faces challenges, in particular in the fault prediction domain. Key limitations include the need for large datasets, substantial computational resources, and limited model interpretability, which makes it challenging to explain predictions in real-world scenarios (Bittar & Garner, 2021). This lack of transparency restricts its broader application in critical fields where understanding model decisions is essential.

Model Optimization Methods

Various optimization algorithms and parameter adjustment techniques have been introduced to better adapt the model to specific tasks and data characteristics (Shafiee & Sørensen, 2019). Model optimization methods include not only the adjustment of the deep learning model structure but also the optimization of hyperparameters, the application of ensemble learning methods, and other aspects.

Recent studies have primarily concentrated on enhancing model training efficiency, minimizing overfitting risks, and boosting generalization capabilities (Qi & Tao, 2019). Research indicates that optimization strategies are crucial for improving both the performance and adaptability of models. In response, researchers are continually advancing model optimization by exploring diverse algorithms, fine-tuning hyperparameters, and implementing other strategies. These optimization techniques have become essential in refining a wide range of machine learning and deep learning models, ultimately strengthening their application in fields like industrial equipment failure prediction.

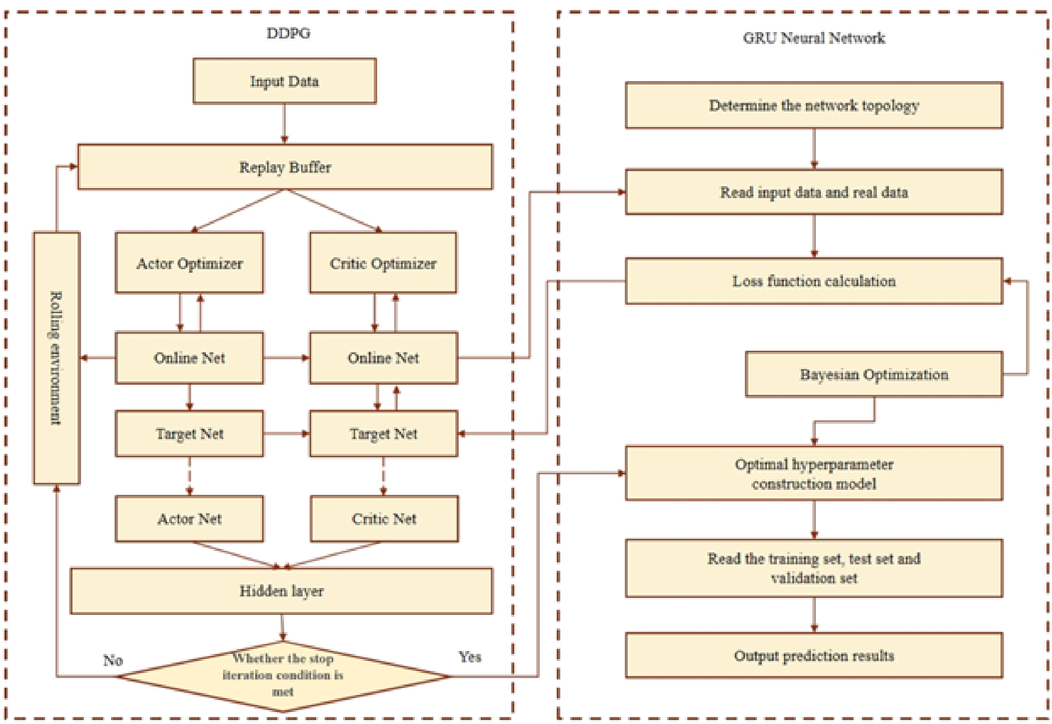
However, current model optimization methods still have some problems, such as the application effect under high-dimensional and diverse data and adaptability in special industrial scenarios (Cardeal et al., 2020). In future work, researchers are likely to work on proposing more efficient and adaptive model optimization methods to better meet the needs of actual industrial equipment monitoring tasks.

METHODOLOGY

Overview of Our Model

This study proposes a comprehensive model based on DDPG and GRUs for industrial equipment fault prediction, optimized through Bayesian optimization to improve performance. The DDPG-GRU model fully combines the advantages of reinforcement learning and time series modeling to achieve higher prediction accuracy and robustness. Its architecture, as shown in Figure 1, consists of two main components: the DDPG module and the GRU neural network module, which are closely integrated through data flow and optimization mechanisms.

Figure 1. Overall flowchart of the model



Note. DDPG = deep reinforcement learning; GRU = gated recurrent unit.

In the model operation, time series data collected by sensors is first input into the GRU module. The main task of the GRU module is to extract features from the input time series data and capture the temporal dependencies of the industrial equipment's operational state. The structure of the GRU network is adjusted through Bayesian optimization to determine the optimal network topology (e.g., the number of hidden units and sequence length) and is optimized using the training set, test set, and validation set. The optimized GRU module outputs high-dimensional time series features that represent the historical dynamics of the equipment's operational state, which are then passed as input to the DDPG module.

The DDPG module is responsible for decision optimization. After receiving the time series features from the GRU, the policy network (actor net) of the DDPG module generates the optimal control actions on the basis of these features, and the value network (critic net) evaluates the actions produced by the policy network by calculating the corresponding Q-values. Through this actor–critic architecture, the DDPG module enables continuous control and optimization of the industrial equipment's state. In addition, to improve training stability and efficiency, the DDPG module incorporates a target network (target net) and an experience replay (replay buffer) mechanism. The target network mitigates parameter fluctuations through soft updates, and the experience replay mechanism breaks the temporal correlation between data by randomly sampling stored state–action pairs, improving the model's learning efficiency.

The integration of the GRU module and the DDPG module allows the entire model to seamlessly connect time series modeling and decision optimization. To be specific, the time-context information provided by the GRU module offers sufficient input support for the DDPG decision-making process, and the optimized actions generated by the DDPG module can, in turn, influence the training objectives of the GRU, forming a collaborative feedback mechanism. This collaborative mechanism ensures that the model can simultaneously focus on both the historical trends and current state of equipment operation, thus generating more accurate predictive results.

Furthermore, the model dynamically adjusts hyperparameters through Bayesian optimization to further enhance performance. During the model training process, Bayesian optimization iteratively optimizes key hyperparameters for the GRU and DDPG modules, such as learning rate, number of hidden units, and batch size, effectively exploring the hyperparameter space to achieve the optimal performance configuration. The entire process is completed by jointly optimizing the loss function, ultimately producing predictions of the industrial equipment's operational state. This design not only improves the model's prediction accuracy but also provides strong adaptability, allowing it to be flexibly applied to various industrial scenarios and offering an efficient and scalable solution for predictive maintenance of industrial equipment.

DDPG Model

Although most reinforcement learning techniques are designed for discrete action spaces, DDPG is particularly well suited for problems where actions are continuous, such as in industrial equipment failure prediction. Unlike models that require discretizing actions, DDPG can directly output continuous values, allowing for finer control and better adaptation to the changing operational states of various devices (Liu et al., 2022). This is especially important in industrial settings, where equipment performance often varies continuously—such as temperature, vibration, or other sensor readings (Lu et al., 2022). By handling continuous action spaces, DDPG minimizes the errors introduced by discretization, making the model more accurate and relevant to real-world industrial environments.

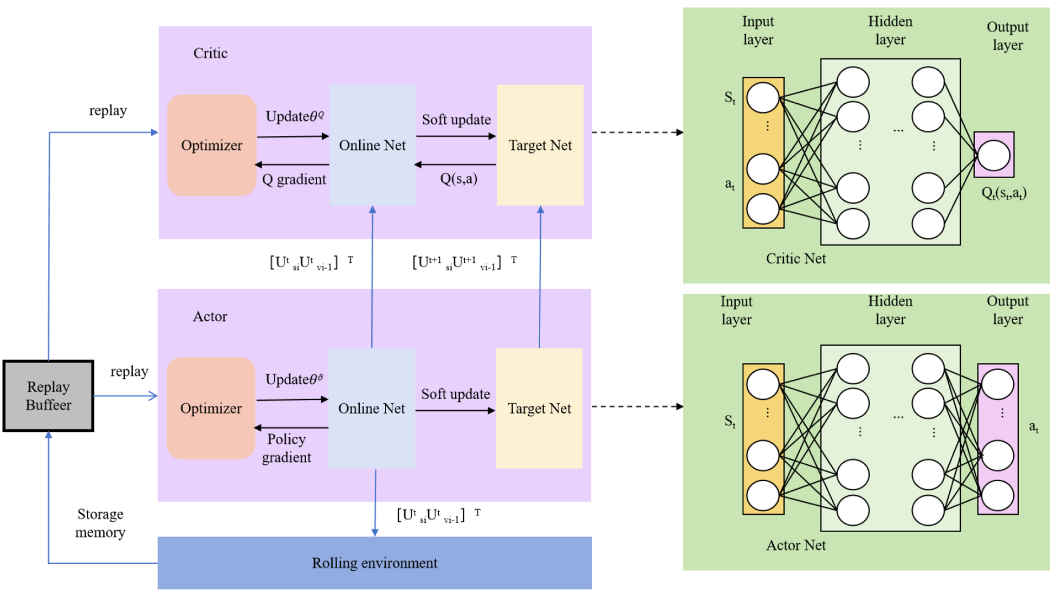
In this study, both the policy network and value network of DDPG are constructed using fully connected neural networks, which provide good expressiveness and flexibility. The policy network takes the state vector as input, with a dimensionality of 64. After processing through two hidden layers, each containing 128 neurons and using rectified linear unit as the activation function, the output layer generates the action vector, with the dimensionality equal to the number of control variables. The activation function of the output layer is Tanh, which restricts the action values to the range of [-1, 1]. The value network, on the other hand, receives the concatenated state and action vectors as input, with a dimensionality of 68. This is also processed through two hidden layers, each containing 256 neurons and using rectified linear unit as the activation function. The output layer generates a scalar

Q-value, representing the value of the current state–action pair. The target network is updated using soft updates, with a step size coefficient set to 0.005 to ensure smooth updates.

In addition, an experience replay mechanism is used to store the state–action–reward–next state quadruples generated during interactions. Random sampling is used to break the temporal correlation between samples, enhancing the diversity and utilization efficiency of the training data. To prevent the problem of gradient explosion, we apply gradient clipping, limiting the maximum gradient value to 10, ensuring the stability of parameter updates. Through these design choices, the DDPG model plays a core role in optimizing equipment operation decisions, improving reliability, and reducing fault risks in the overall architecture of this study. When combined with the GRU module's time series modeling capabilities, it further enhances the effectiveness of predictive maintenance.

The structure diagram of the DDPG model is shown in Figure 2.

Figure 2. Flowchart of the bidirectional encoder representations from transformers (BERT) model



The core equations of the DDPG model are presented as follows:

$$Q\left(s_t, a_t\right) = r_t + \gamma Q\left(s_{t+1}, \pi\left(s_{t+1}\right)\right) \tag{1}$$

$$a_t = \pi\left(s_t\right) + \mathcal{N} \tag{2}$$

$$\mathcal{L}\left(\theta^Q\right) = \mathbb{E}_{s_t,a_t,r_t,s_{t+1}}\left[\left(Q(s_t, a_t|\theta^Q) - y_t\right)^2\right] \tag{3}$$

$$y_t = r_t + \gamma Q'\left(s_{t+1}, \pi'\left(s_{t+1}\middle|\theta^{\pi'}\right)|\theta^{Q'}, \tag{4}$$

where Q is the action-value function; s_t is the state at time t; a_t is the action; r_t is the reward; γ is the discount factor; and π is the policy function; a_t is the selected action; $\pi\left(s_t\right)$ is the output of the deterministic policy at state s_t; and $\mathcal{N}$ represents noise added for exploration, typically Gaussian; and $\mathcal{L}$ is the loss function for the Q-network.

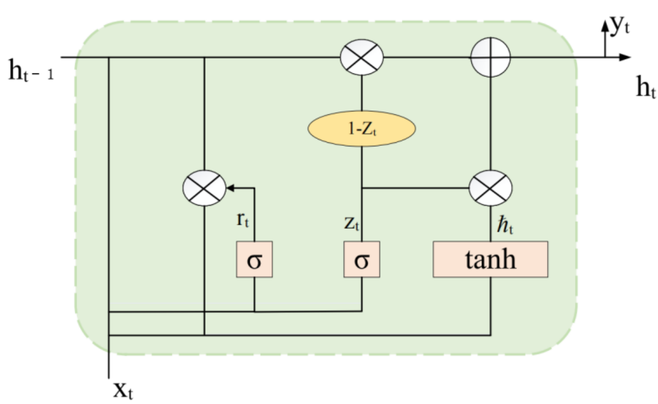
GRU Model

Unlike traditional RNNs, GRUs incorporate gates that help capture long-range dependencies more efficiently, making them ideal for applications such as industrial equipment failure prediction, where changes in equipment status are often time dependent (Wang et al., 2019). The GRU uses gates, including forget and update gates, to selectively retain or discard information as it processes sequences, improving its ability to manage long-term dependencies. In industrial equipment monitoring it is crucial to consider historical data that span extended time periods, such as the operational status of machinery over days, weeks, or months (Zhang et al., 2022). GRUs excel in this context by capturing the evolving status of equipment over time, helping identify underlying patterns that could indicate potential failures or the need for maintenance. Compared with other models, such as LSTM, GRUs have a simpler structure with fewer parameters, which makes them more computationally efficient—especially when working with large datasets common in industrial failure prediction scenarios. This efficiency is beneficial for real time applications because it reduces the computational load during training and inference.

In the model proposed in this article, the GRU component plays a vital role in modeling the temporal dynamics of equipment status. By learning key patterns from time series data, the GRU helps track the evolution of equipment conditions and predict potential failures. This ability to capture temporal relationships in operational data is crucial for effective predictive maintenance. When combined with the DDPG module, the GRU enhances the model's capacity to understand both the time-dependent nature and the dynamic changes in equipment performance. Together, these components improve the model's overall prediction accuracy, making the GRU a critical element in enhancing the reliability of industrial equipment and supporting predictive maintenance strategies.

Figure 3 illustrates the architecture of the GRU module.

Figure 3. Flowchart of the prophet module



The core equations of the GRU module are presented as follows:

$$r_t = \sigma\left(W_r x_t + U_r h_{t-1} + b_r\right) \tag{5}$$

$$z_t = \sigma\left(W_z x_t + U_z h_{t-1} + b_z\right) \tag{6}$$

$$h_t = (1 - z_t) \odot h_{t-1} + z_t \odot \tilde{h}_t), \tag{7}$$

$$\tilde{h}_t = tanh(W_h x_t + U_h(r_t \odot h_{t-1}) + b_h) \tag{8}$$

where W_h is the weight matrix for input-to-hidden, U_h is the weight matrix for hidden-to-hidden after reset applying, r_t is the reset gate vector, ⊙ denotes the element-wise multiplication, b_h is the bias vector for hidden state, h_t is the hidden state vector at time t, z_t is the update gate vector, h_{t-1} is the previous hidden state vector, and $\tilde{h}_t$ is the candidate hidden state vector.

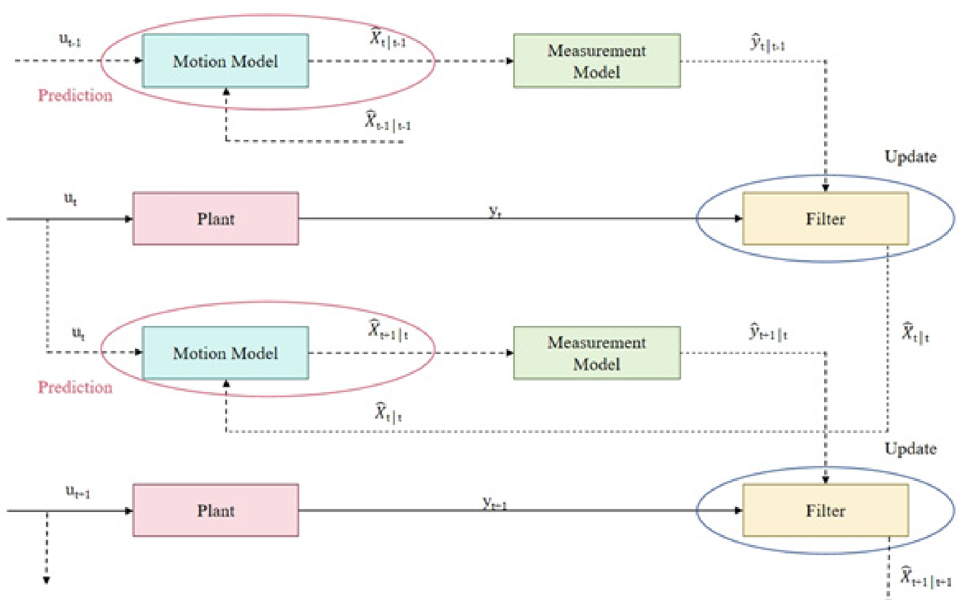
Bayesian Optimization Model

Bayesian optimization is a powerful approach for global optimization, using a probabilistic model to estimate the objective function and efficiently identify optimal solutions. It is especially suitable for complex and nonconvex optimization problems that are computationally intensive and is often used in applications such as hyperparameter tuning and enhancing machine learning models (Compare et al., 2020). In industrial contexts, this method is instrumental in adjusting control parameters of intricate systems, leading to improvements in both system performance and energy efficiency while minimizing the number of required iterations (Soltanali et al., 2021).

In our study, Bayesian optimization plays a pivotal role in tuning the hyperparameters of the DDPG-GRU model, contributing significantly to improving both accuracy and resilience. This approach ensures that our model effectively adapts to diverse types of industrial machinery and varying operational environments. By continuously optimizing hyperparameters, Bayesian optimization enhances the predictive maintenance framework, allowing the model to respond efficiently to changing operational conditions and produce consistent, reliable predictions.

The structure diagram of the Bayesian optimization model is shown in Figure 4.

Figure 4. Flowchart of the softmax model



The main formula of the Bayesian optimization model is as follows:

$$P(data|\theta) = L(\theta; data), \tag{9}$$

where Pdenotes probability, *data* represents the observed data, θ signifies the parameters of the model, and L is the likelihood function.

$$P(data) = \int P(data|\theta)P(\theta)d\theta, \tag{10}$$

where $P(data)$ is the marginal likelihood (or evidence) of the observed data, integrating the likelihood and the prior over all possible values of θ. $P(\theta)$ represents the prior probability distribution of the parameters θ, expressing our beliefs about θ before observing the data.

$$\mathbb{E}[\theta|data] = \int \theta P(\theta|data)d\theta, \tag{11}$$

where $\mathbb{E}[\theta|data]$ is the expected value (or mean) of the posterior distribution of θ, providing a point estimate of the parameters after observing the data.

$$P(new\ data|data) = \int P(new\ data|\theta)P(\theta|data)d\theta, \tag{12}$$

where $P(new\ data|data)$ is the posterior predictive distribution for new unseen data given the observed data, integrating over the posterior distribution of the parameters $P(\theta|data)$.

$$\int_{\theta_{low}}^{\theta_{high}} P(\theta|data)d\theta = 1 - \alpha, \tag{13}$$

where θ_{low} and θ_{high} are the lower and upper bounds of the $1 - \alpha$ Bayesian credible interval for the parameter θ, within which the parameter lies with probability $1 - \alpha$.

EXPERIMENT

Experimental Dataset

To deeply explore the fault prediction model of industrial equipment, we designed and conducted a series of experiments in an environment that simulated real working conditions. The experiments rely on data sets from four different sources: the Industrial Equipment Health dataset, the Power Equipment Monitoring dataset, the Aircraft Engine Failure dataset, and the Smart Manufacturing Sensor dataset. Our aim was to analyze and train advanced prediction models, thereby improving performance in various aspects.

The Industrial Equipment Health dataset covers rich information from manufacturing equipment; specifically, it contains sensor data, maintenance records, and equipment operating status history for more than 100,000 device instances. These devices cover different types of manufacturing equipment, ensuring data diversity. Data features include various sensor measurements, such as temperature and vibration, and are labeled with the operating status of the equipment, providing rich monitoring information for the deep learning model. These data were collected over many years, so the dataset contains rich temporal information, which helps the model better understand the evolution of the equipment's working status (Zhang et al., 2019). In the experiment, we used more than 5 years of data, which included the evolution of the equipment's status, allowing the model to better understand the long-term operation of the equipment. Sensor measurements in the data set include temperature, vibration, and more, providing rich monitoring information for the deep learning model. The Industrial Equipment Health dataset not only provides not only diversified manufacturing data for the model but also challenging actual industrial scenarios for model training, making the model more robust and generalizable.

The Power Equipment Monitoring dataset covers sensor data, maintenance records, and operating status of multiple power equipment; specifically, the data come from the power industry and contains monitoring data from more than 50,000 equipment instances in the power industry. In our experiments, we used data covering a wide range of collection times across multiple seasons and load conditions, ensuring data diversity. Various monitoring data, such as current, voltage, and power, are marked with the operating status of the equipment, providing the model with the opportunity to adapt to the complexity of power equipment (Ayvaz & Alpay, 2021). The application of the Power Equipment Monitoring dataset helps the model better understand the operation of power equipment and improves the adaptability and robustness of the model in the power field.

The Aircraft Engine Failure dataset is a key data source in the study. This data set contains operational data from multiple engine models in the aviation industry (Arias Chao et al., 2021). The dataset covers sensor data, maintenance records, and performance parameters for more than 10,000 engine instances. Sensor measurement values include engine speed, temperature, pressure, and more, and have a high degree of timing complexity. Long-term data collection ensures the model's accurate prediction of engine failure and improves the model's practical application value in the aviation field.

The Smart Manufacturing Sensor dataset is a treasure trove of multivariate data, including sensor data from various equipment in smart manufacturing factories (Shahbazi & Byun, 2021). the data in this dataset come from smart manufacturing enterprise partners and include information on more than 80,000 device instances from smart manufacturing enterprise partners. In the experiments, we used sensor measurements with high-resolution time series data, such as processing speed, temperature, humidity, and so on. The extensive collection of data covering different stages of the production cycle helps the model more fully understand changes in equipment status. This data set provides the model with the opportunity to adapt to the special working conditions of smart manufacturing factories and improves the generalization of the model.

Experimental Setup and Details

The experiment used time series sensor data from industrial equipment, collected over a 2-month period at a frequency of 1 Hz. The dataset included data from multiple devices, with sensor readings such as temperature, pressure, vibration, and rotational speed. These sensors operated under varying working conditions, making the dataset diverse and complex.

To ensure the quality and integrity of the data, several preprocessing steps were performed. Outliers were identified using domain knowledge, such as temperature or pressure values exceeding the operational range of the devices. These extreme values were removed from the dataset. Missing values were addressed through linear interpolation to maintain continuity and completeness, and duplicates were eliminated to ensure that each data point had a unique time stamp.

Normalization was applied in two stages. First, the Z-score method adjusted the values of each sensor to have a mean of 0 and a standard deviation of 1, preventing certain features from disproportionately influencing the model. In addition, all features were scaled to a [0, 1] range, improving model stability and convergence speed. To avoid overfitting, the dataset was divided into training (80%) and testing (20%) sets. Furthermore, five-fold cross-validation was used to provide a robust evaluation by training on four subsets and validating on the remaining subset, repeated five times.

The model was trained using the DDPG-GRU architecture, with hyperparameters fine tuned through Bayesian optimization to optimize the F1 score. The optimization focused on parameters such as learning rate, batch size, and the number of hidden units in the GRU. The learning rate was adjusted between 0.001 and 0.01, and the optimal value was found to be 0.003. The GRU's hidden units were set to 64 on the basis of optimization results. In addition, the learning rates for the actor and critic networks in the DDPG algorithm were optimized, with the final value selected as 0.0005. The batch size for experience replay was optimized to 32, which improved the model's performance.

In each training iteration, a maximum of 500 iterations were allowed. The first 100 iterations focused on exploring different hyperparameter configurations, and the remaining 400 iterations refined the model's performance. A sequence length of 30 time steps was set for the GRU to ensure that the model could adequately capture temporal dependencies in the data. Bayesian optimization used a Gaussian process model to estimate the objective function, iteratively updating the sampling strategy to identify the best hyperparameter combination. The Adam optimizer, with an initial learning rate of 0.001, was used to accelerate convergence and ensure training stability.

Experimental Results and Analysis

As detailed in Tables 1 and 2, our model demonstrates considerable advantages across a variety of datasets, particularly excelling in industrial equipment health monitoring tasks. On the Industrial Equipment Health dataset, the model achieved accuracy, recall, F1 score, and area under the curve (AUC) values of 96.57%, 95.76%, 91.92%, and 92.68%, respectively, significantly outperforming the competing models. In comparison, the best accuracy achieved by other approaches was only 92.91%, with the highest recall reaching 92.89%, and F1 score and AUC values of 85.92% and 88.64%, respectively. This clear margin underscores our model's effectiveness in capturing relevant patterns for predicting equipment health.

Moreover, similar performance gains are observed across the other datasets. Our model's generalization capabilities are highlighted by its consistent high performance, demonstrating adaptability to diverse industrial scenarios. A notable example is the Smart Manufacturing Sensor dataset, where our model achieved an impressive accuracy of 98.43%. In contrast, the next best model achieved only 95.48%, a significant gap that underlines the strengths of our model in precision-sensitive environments like smart manufacturing. Such superior accuracy is indicative of our model's ability to handle complex sensor data and make highly reliable predictions, which is crucial for real-time decision making in smart industrial settings.

Beyond just accuracy, the model also excels in other key metrics, such as recall, F1 score, and AUC. The recall rates observed indicate that our model is particularly effective at identifying true positives, meaning it has a higher ability to correctly predict instances of equipment failure. This feature is vital for minimizing unplanned downtimes and ensuring equipment reliability. The F1 score and AUC further reflect a balanced and comprehensive capability, combining precision and recall to give a strong measure of overall effectiveness. The robust performance in these metrics confirms that the model is well suited for scenarios where both timely intervention and accurate failure prediction are essential.

Furthermore, our model's architecture has proven to be more efficient in dealing with the noise and variability inherent in complex industrial environments. By capturing nuanced relationships in the time series data, it ensures predictive robustness even under challenging operational conditions. Figure 5 provides a visual comparison that effectively illustrates the advantages of our model over alternative approaches, giving a clearer, intuitive understanding of its performance strengths.

Table 1. Comparisons in accuracy, recall, F1 scores, and area under the curve indicators on the industrial equipment health dataset and the power equipment monitoring dataset

Model	Industrial Equipment Health dataset				Power Equipment Monitoring dataset			
	Accuracy	Recall	F1 score	AUC	Accuracy	Recall	F1 score	AUC
Zhao (Zhao & Zhou, 2022)	86.36	86.76	90.11	89.3	88.07	86.04	90.64	83.86
Ma (Ma et al., 2022)	90.98	90.87	90.4	87.8	87.46	84.32	85.19	92.67

continued on following page

Table 1. Continued

Model	Industrial Equipment Health dataset				Power Equipment Monitoring dataset			
	Accuracy	Recall	F1 score	AUC	Accuracy	Recall	F1 score	AUC
Bin (Bin & Sun, 2022)	87.37	84.28	89.5	88.97	96.32	92.27	86.15	85.26
Souza (Souza et al., 2021)	92.91	92.89	85.92	88.64	95.7	92.69	86.37	92.29
Khalil (Khalil et al., 2021)	86.13	92.6	91.08	86.77	94.95	88.19	89.46	84.67
Bampoula (Bampoula et al., 2021)	89.11	92.89	91.08	90.9	88.18	85.01	87.07	89.66
Ours	96.57	95.76	91.92	92.68	96.66	94.3	91.36	93.9

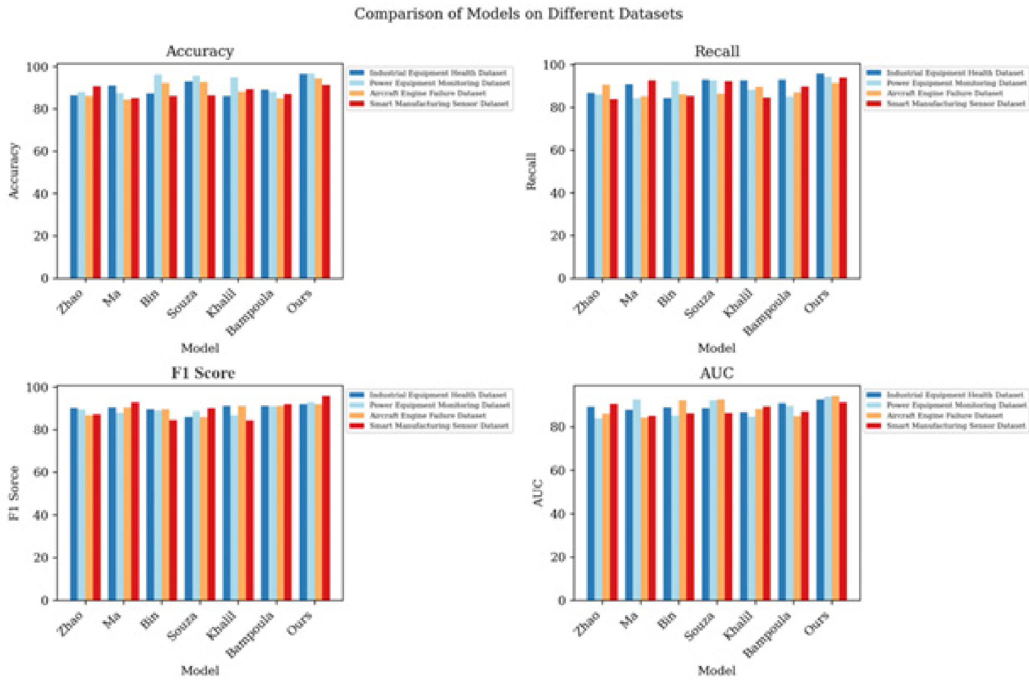
Note. AUC = area under the curve.

Table 2. Comparisons of accuracy, recall, F1 scores, and area under the curve indicators on the aircraft engine failure dataset and the smart manufacturing sensor dataset

Model	Aircraft Engine Failure dataset				Smart Manufacturing Sensor dataset			
	Accuracy	Recall	F1 score	AUC	Accuracy	Recall	F1 score	AUC
Zhao	95.98	85.29	84.45	87.16	89.87	87.32	88.38	86.31
Ma	94.58	87.41	84.09	92.71	94.06	91.55	86.62	86.45
Bin	85.91	85.6	90.98	84.42	86.21	89.09	86.25	91.13
Souza	90.27	85.9	87.76	90.02	87.17	85.74	88.62	84.19
Khalil	94.88	85.68	85.56	84.3	89.59	93.44	86.26	91.47
Bampoula	94.3	89.4	90.04	91.91	95.48	93.43	84.88	88.71
Ours	92.33	96.16	92.03	95.78	98.43	94.99	91.78	91.39

Note. AUC = area under the curve.

Figure 5. Comparisons of models on different datasets



Note. AUC = area under the curve.

As summarized in Tables 3 and 4, our model demonstrates considerable advantages across several performance metrics; specifically, evaluated on four different datasets, the model shows strong efficiency in terms of parameter count, computational cost (flops), inference speed, and training duration. For example, on the Industrial Equipment Health dataset, our model uses only 336.85 M parameters, significantly fewer compared with the highest count of 481.62 M in other models, resulting in improved efficiency in computation and memory usage. Furthermore, our model's flops is 3.52 G, markedly lower than the maximum of 8.04 G found in other models, highlighting its reduced computational complexity. With respect to inference and training times, our model also performs notably better. On the Smart Manufacturing Sensor dataset, our model's inference time is just 5.33 ms, much faster than the highest reported time of 13.46 ms of other models, thus offering better real time performance. In addition, the training time is shorter compared with other approaches. Overall, these results indicate that our model excels in aspects such as parameter efficiency, computational demands, and both inference and training speeds, underlining its effectiveness for industrial equipment fault prediction tasks. Figure 6 provides a visualization of the table, offering a clearer comparison of our model's performance against others.

Table 3. Efficiency verification and comparison in parameters, flops, inference times, and training time indicators on the industrial equipment health dataset and the power equipment monitoring dataset

Model	Industrial Equipment Health dataset				Power Equipment Monitoring dataset			
	Parameters (M)	Flops (G)	Inference time (ms)	Training time (s)	Parameters (M)	Flops (G)	Inference time (ms)	Training time (s)
Zhao	481.62	6.26	8.91	553.88	561.37	5.85	9.94	484.54
Ma	819.29	7.96	13.00	778.36	689.35	9.02	10.70	726.64
Bin	769.17	5.12	6.09	721.35	652.96	8.42	7.77	685.67
Souza	699.06	8.04	10.38	707.34	706.90	8.10	13.46	729.01
Khalil	496.52	5.07	7.67	450.60	466.27	4.47	7.22	489.28
Bampoula	338.69	3.55	5.33	325.81	319.14	3.64	5.61	338.60
Ours	336.85	3.52	5.33	327.06	318.05	3.66	5.60	335.55

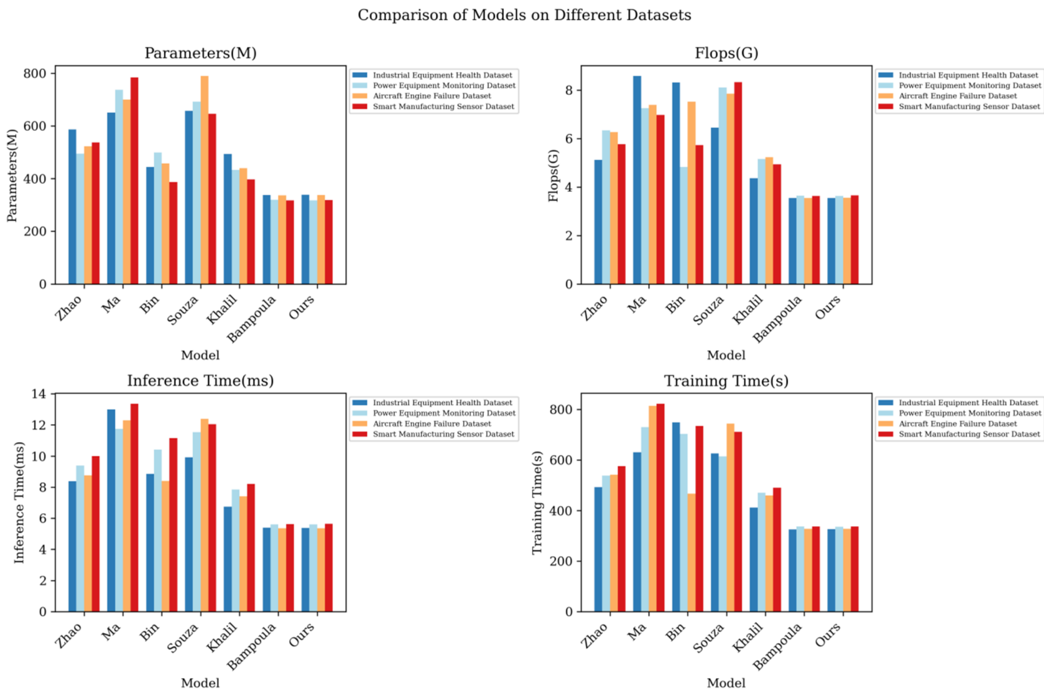
Note. M = xxxx; G = xxxx.

Table 4. Efficiency verification and comparison in parameters, flops, inference times, and training time indicators on the aircraft engine failure dataset and the smart manufacturing sensor dataset

Model	Aircraft Engine Failure dataset				Smart Manufacturing Sensor dataset			
	Parameters (M)	Flops (G)	Inference time (ms)	Training time (s)	Parameters (M)	Flops (G)	Inference time (ms)	Training time (s)
Zhao	577.06	5.72	9.00	572.05	487.14	5.71	8.42	586.75
Ma	832.25	8.54	11.83	733.86	754.11	8.09	12.89	743.22
Bin	773.70	6.15	5.98	613.40	729.11	8.37	10.60	581.69
Souza	649.49	6.83	10.09	638.14	707.32	7.91	11.78	730.74
Khalil	492.79	4.39	8.02	452.23	396.77	5.22	7.18	423.20
Bampoula	338.45	3.54	5.36	325.58	320.49	3.64	5.63	337.21
Ours	337.13	3.56	5.32	326.37	320.10	3.65	5.62	337.01

Note. M = xxxx; G = xxxx.

Figure 6. Comparison of model performance on different datasets



We conducted ablation studies to evaluate the impact of each component of the DDPG-GRU model across four distinct datasets, in addition to analyzing the performance of the entire proposed model. The results are given in Tables 5 and 6. These ablation experiments aimed to identify the contributions of different elements within the architecture, ultimately assessing their influence on prediction accuracy and overall performance. On the Industrial Equipment Health dataset, the complete model delivered the highest results for all evaluated metrics, achieving an accuracy of 96.54%, a recall of 95.85%, an F1 score of 92.79%, and an AUC of 95.36%. This performance underscores the effectiveness of integrating all model components compared with partial configurations.

Similarly, on the Power Equipment Monitoring dataset the comprehensive model outperformed all other versions, reaching an accuracy of 96.03%, a recall of 95.19%, an F1 score of 93.9%, and an AUC of 95.03%. This consistency suggests that the individual components, including the reinforcement learning module and the temporal sequence model, each play critical roles in improving prediction performance. The Aircraft Engine Failure dataset further validated the robustness of our complete model given that it achieved superior accuracy and AUC scores, specifically, 96.98% and 95.54%, respectively, demonstrating the versatility of the model across diverse types of data.

For the Smart Manufacturing Sensor dataset, the full model again led in all metrics, with 97.88% accuracy, 94.12% recall, an F1 score of 92.21%, and an AUC of 95.65%. These results are particularly notable because they highlight the model's capability in handling noisy sensor data, emphasizing its adaptability in different industrial environments. Each ablation experiment's results point to the importance of including all designed components—such as the GRU for capturing temporal dependencies and the DDPG for optimal decision making—to enhance the model's overall predictive accuracy and reliability.

The visualization provided in Figure 7 further illustrates the clear performance advantages of our comprehensive model over other configurations, offering insights into how different components contribute to enhancing key metrics. The findings not only affirm the overall superior performance of our proposed architecture in predictive maintenance tasks but also highlight the significance of a holistic model design in which each component adds meaningful value to time series data analysis, ultimately boosting the accuracy, robustness, and reliability of the system.

Table 5. Ablation experiments on the industrial equipment health dataset and the power equipment monitoring dataset

Model	Industrial Equipment Health dataset				Power Equipment Monitoring dataset			
	Accuracy	Recall	F1 score	AUC	Accuracy	Recall	F1 score	AUC
GRU-Bayesian	91.35	89.95	86.8	93.59	91.18	92.53	87.15	93.32
DDPG-Bayesian	95.49	87.9	85.04	89.59	95.3	86.88	87.83	93.45
DDPG-GRU	94.24	84.77	90.76	87.97	95.78	92.84	86.81	84.26
Ours	96.54	95.85	92.79	95.36	96.03	95.19	93.9	95.03

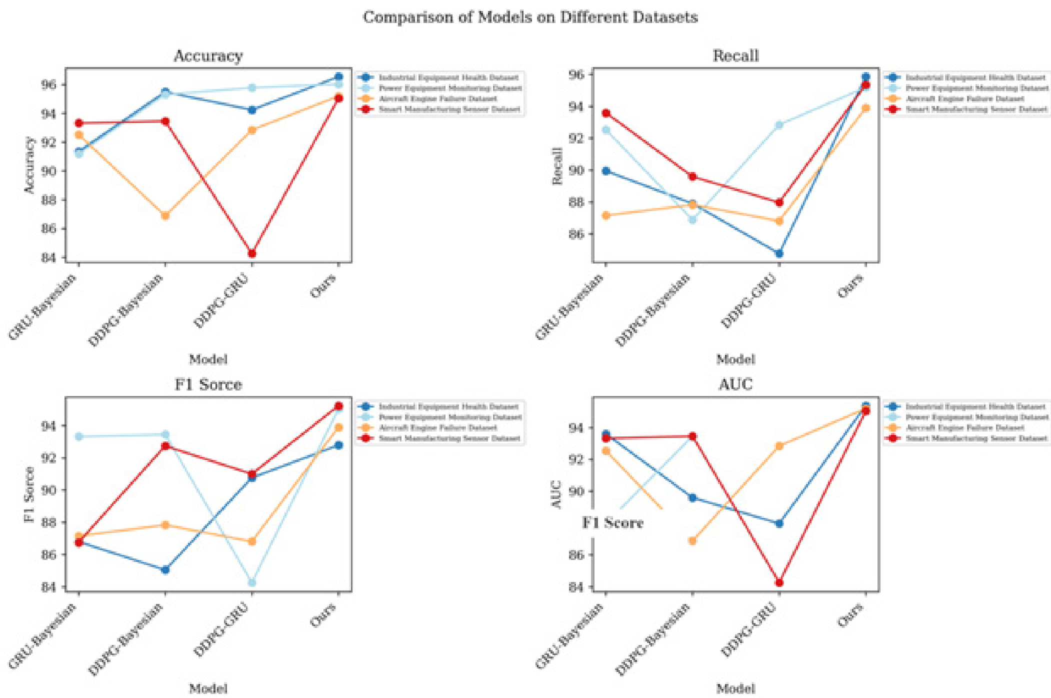
Note. PSO = Particle Swarm Optimization; LSTNet = Long Short-Term Neural Networks; AUC = area under the curve; GRU = gated recurrent unit; DDPG = deep reinforcement learning.

Table 6. Ablation experiments on the aircraft engine failure dataset and the smart manufacturing sensor dataset

Model	Aircraft Engine Failure dataset				Smart Manufacturing Sensor dataset			
	Accuracy	Recall	F1 score	AUC	Accuracy	Recall	F1 score	AUC
GRU-Bayesian	88.84	90.77	86.76	86.75	90.12	90.55	90.15	89.28
DDPG-Bayesian	87.06	87.02	87.41	92.74	87.2	92.97	84.67	89.53
DDPG-GRU	90.84	85.41	84.27	91	95.76	85.39	89.56	85.69
Ours	96.98	95.54	93.41	95.23	97.88	94.12	92.21	95.65

Note. PSO = xxxx; LSTNet = xxxx; AUC = area under the curve; GRU = gated recurrent unit; DDPG = deep reinforcement learning.

Figure 7. Comparison of model performances on different datasets



Note. AUC = area under the curve; GRU = gated recurrent unit; DDPG = deep reinforcement learning.

In Tables 7 and 8, we present the results of a series of comparative experiments designed to assess the impact of different optimization algorithms on the DDPG-GRU model's performance. The comparisons were conducted across several datasets, focusing on metrics such as the number of parameters, computational complexity (flops), inference time, and training time. Our model, optimized with Bayesian optimization, demonstrates a clear advantage over other algorithms, including Adam, Root Mean Square Propagation (RMSprop), and Proximal Policy Optimization (PPO), in particular in reducing the number of parameters. This reduction makes the model more lightweight, suitable for resource-limited environments. For example, on the Industrial Equipment Health dataset, our model has 214.9 M parameters—significantly lower than Adam and RMSprop, with reductions of 42.7% and 45.6%, respectively.

In terms of computational complexity (flops), our model consistently shows lower values, indicating higher computational efficiency during inference. On the Smart Manufacturing Sensor dataset, our model's flops value is 206.1 G, compared with 335.98 G and 380.49 G for Adam and RMSprop, representing a decrease of 39% and 45.8%, respectively. Our model also achieves shorter inference times, highlighting its practical efficiency. For instance, inference time on the Aircraft Engine Failure dataset is 211.47 ms, notably faster than RMSprop's 269.28 ms, improving real time capabilities. In addition, training time is considerably shorter, with a significant reduction of 70.8% on the Power Equipment Monitoring dataset compared with PPO. These results underscore the strengths of our proposed model, in particular in efficiency metrics, such as parameter count, computational load, inference, and training time, while also validating the effectiveness of Bayesian optimization in enhancing performance. Figure 8 visually illustrates these performance differences, providing a clear comparison across optimization strategies.

Table 7. Comparation experiments on the PSO module based on the industrial equipment health dataset and the power equipment monitoring dataset

Model	Industrial Equipment Health dataset				Power Equipment Monitoring dataset			
	Parameters (M)	Flops (G)	Inference time (ms)	Training time (s)	Parameters (M)	Flops (G)	Inference time (ms)	Training time (s)
Adam	373.83	260.96	258.87	309.35	359.76	382.46	222.88	399.17
RMSprop	394.83	307.33	269.28	280.56	270.77	350.09	396.7	349.06
PPO	346.89	369.68	256.39	320.64	354.85	327.43	269.24	371.43
Ours	214.9	186.2	211.47	226.94	170.51	172.63	200.33	108.4

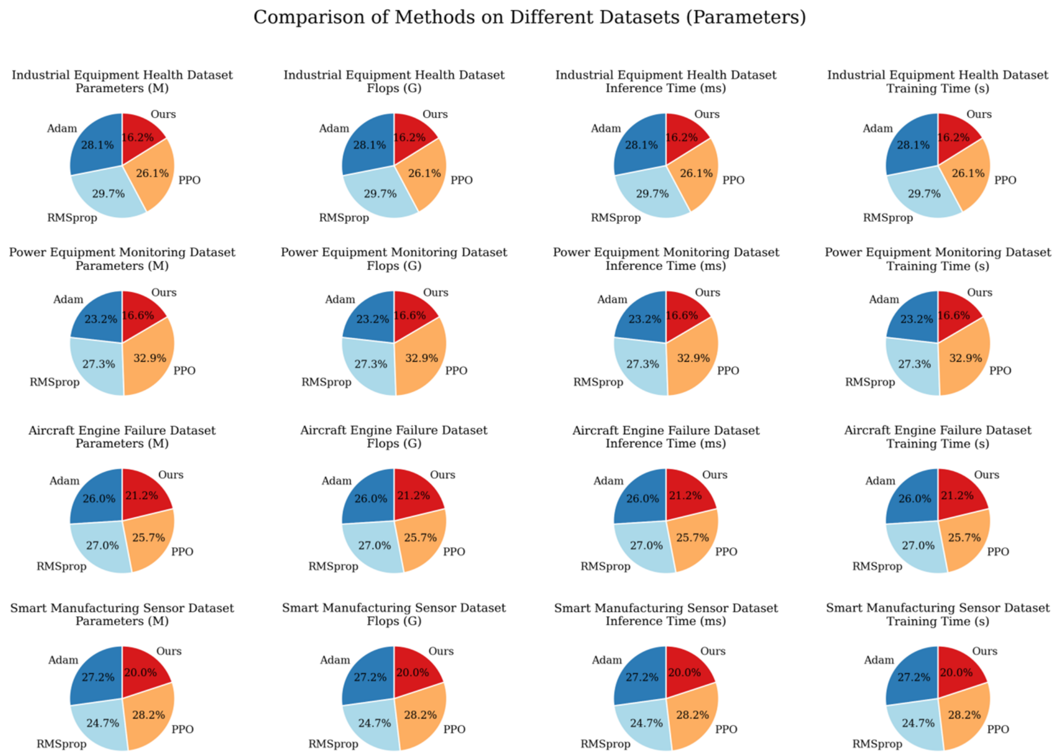
Note. PSO = Particle Swarm Optimization; RMSprop = Root Mean Square Propagation; M = Million; G = Giga.

Table 8. Comparation experiments on the PSO module based on the aircraft engine failure dataset and the smart manufacturing sensor dataset

Model	Aircraft Engine Failure dataset				Smart Manufacturing Sensor dataset			
	Parameters (M)	Flops (G)	Inference time (ms)	Training time (s)	Parameters (M)	Flops (G)	Inference time (ms)	Training time (s)
Adam	386.84	304.24	308.22	389.13	281.93	243.27	335.98	384.68
RMSprop	376.21	268.9	249.06	290.48	380.49	296.91	212.55	399.33
PPO	308.68	310.57	238.72	289.99	360.03	285.37	393.96	406.2
Ours	184.73	128.17	227.44	192.26	206.1	216.29	208.58	183.31

Note. PSO = Particle Swarm Optimization; RMSprop = Root Mean Square Propagation; M = Million; G = Giga.

Figure 8. Comparison of model performance on different datasets



Note. PPO = xxxx; RMSprop = xxxx.

CONCLUSION AND DISCUSSION

The main objective of this study was to enhance the reliability of industrial equipment and provide an effective solution for predictive maintenance. To achieve this goal, we proposed and implemented a novel DDPG-GRU model that combines deep reinforcement learning and time series data analysis. In terms of model design, we fully consider the characteristics of equipment fault prediction, using a DDPG to optimize equipment operation decisions and using the GRU network to capture key dependencies in time series data. Through detailed data preprocessing, Bayesian optimization for hyperparameter tuning, ablation studies, and comparative experiments, we comprehensively evaluated the model's performance and the contributions of its individual components. The experimental results showed that the DDPG-GRU model offers significant advantages in industrial equipment fault prediction. Our model outperformed traditional prediction methods in several performance metrics, including accuracy, recall, and F1 score, and demonstrated stronger robustness and generalization ability when handling complex time series data and large-scale industrial datasets. By using Bayesian optimization, we further improved the model's hyperparameter adjustment capability, allowing the model to better adapt to and perform well in different industrial scenarios. The results of ablation and comparative experiments helped us gain a deeper understanding of the role of each component in the model's performance, providing important insights for future model optimization and deployment.

Despite these achievements, this study has some limitations and areas that need improvement. First, because of the potential uncertainty in the quality and labeling accuracy of real-world industrial data, our model may experience some performance degradation when faced with noisy data or incorrect labels. Second, the model's robustness in dealing with sudden events and extreme operating conditions still needs improvement; specifically, under extreme conditions, such as equipment malfunctions, sensor failures, or environmental changes, the model's performance may be significantly affected, highlighting the need for further enhancement of its adaptability and interference resistance.

In future research, we plan to address these issues. First, we will explore more advanced robust learning algorithms to improve the model's stability under noisy data and abnormal events. In addition, given the diversity and complexity of industrial equipment, we aim to expand the experimental dataset to include more types of equipment and complex operating conditions to enhance the model's generalization ability. We also plan to incorporate more domain expertise to improve the model's interpretability and deepen our understanding of equipment operating mechanisms. Furthermore, to meet the practical needs of large-scale industrial environments, we intend to optimize the model's inference speed and computational efficiency on distributed computing platforms, further enhancing its practical application value in industrial production.

The DDPG-GRU model proposed in this study provides an innovative solution for predictive maintenance of industrial equipment, with strong academic value and potential for application. We believe that by further optimizing the model's robustness and generalization ability, it will play an important role in smart manufacturing, the industrial internet of things, and related fields and will provide valuable experience and insights for the future intelligent development of equipment fault prediction and maintenance management.

COMPETING INTERESTS

The authors of this publication declare there are no competing interests.

FUNDING STATEMENT

This work was supported by Research on the strategy of transformation and upgrading of equipment manufacturing industry in Heilongjiang Province under the background of dual circulation [grant number QD21JYE402].

CORRESPONDING AUTHOR

Correspondence concerning this article should be addressed to Yingli Wu, (China, 562484197@qq.com)

PROCESS DATES

February 3, 2025
Received: November 27, 2024, Revision: January 1, 2025, Accepted: January 20, 2025

REFERENCES

Arias Chao, M., Kulkarni, C., Goebel, K., & Fink, O. (2021). Aircraft engine run-to-failure dataset under real flight conditions for prognostics and diagnostics. *Data*, *6*(1), 5. DOI: 10.3390/data6010005

Ayvaz, S., & Alpay, K. (2021). Predictive maintenance system for production lines in manufacturing: A machine learning approach using IoT data in real-time. *Expert Systems with Applications*, *173*(C), 114598. DOI: 10.1016/j.eswa.2021.114598

Bampoula, X., Siaterlis, G., Nikolakis, N., & Alexopoulos, K. (2021). A deep learning model for predictive maintenance in cyber-physical production systems using LSTM autoencoders. *Sensors (Basel)*, *21*(3), 972. DOI: 10.3390/s21030972 PMID: 33535642

Bin, S., & Sun, G. (2022). Research on the influence maximization problem in social networks based on the multi-functional complex networks model. *Journal of Organizational and End User Computing*, *34*(3), 1–17. DOI: 10.4018/JOEUC.302662

Bittar, A., & Garner, P. N. (2021). A Bayesian interpretation of the light gated recurrent unit. In *ICASSP 2021–2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Toronto, ON, Canada, 2021 (pp. 2965–2969). IEEE. DOI: 10.1109/ICASSP39728.2021.9414259

Cardeal, G., Höse, K., Ribeiro, I., & Götze, U. (2020). Sustainable business models—Canvas for sustainability, evaluation method, and their application to additive manufacturing in aircraft maintenance. *Sustainability (Basel)*, *12*(21), 9130. DOI: 10.3390/su12219130

Carvalho, T. P., Soares, F. A., Vita, R., Francisco, R. P., Basto, J. P., & Alcalá, S. G. (2019). A systematic literature review of machine learning methods applied to predictive maintenance. *Computers & Industrial Engineering*, *137*, 106024. DOI: 10.1016/j.cie.2019.106024

Compare, M., Baraldi, P., Bani, I., Zio, E., & McDonnell, D. (2020). Industrial equipment reliability estimation: A Bayesian Weibull regression model with covariate selection. *Reliability Engineering & System Safety*, *200*, 106891. DOI: 10.1016/j.ress.2020.106891

Dalzochio, J., Kunst, R., Pignaton, E., Binotto, A., Sanyal, S., Favilla, J., & Barbosa, J. (2020). Machine learning and reasoning for predictive maintenance in Industry 4.0: Current status and challenges. *Computers in Industry*, *123*, 103298. DOI: 10.1016/j.compind.2020.103298

Huang, J., Yu, X., An, D., Ning, X., Liu, J., & Tiwari, P. (2025). Uniformity and deformation: A benchmark for multi-fish real-time tracking in the farming. *Expert Systems with Applications*, *264*, 125653. DOI: 10.1016/j.eswa.2024.125653

Jiang, Y., Dai, P., Fang, P., Zhong, R. Y., Zhao, X., & Cao, X. (2022). A2-LSTM for predictive maintenance of industrial equipment based on machine learning. *Computers & Industrial Engineering*, *172*, 108560. DOI: 10.1016/j.cie.2022.108560

Khalil, R. A., Saeed, N., Masood, M., Fard, Y. M., Alouini, M.-S., & Al-Naffouri, T. Y. (2021). Deep learning in the industrial internet of things: Potentials, challenges, and emerging applications. *IEEE Internet of Things Journal*, *8*(14), 11016–11040. DOI: 10.1109/JIOT.2021.3051414

Lai, C.-F., Chien, W.-C., Yang, L. T., & Qiang, W. (2019). LSTM and edge computing for big data feature recognition of industrial electrical equipment. *IEEE Transactions on Industrial Informatics*, *15*(4), 2469–2477. DOI: 10.1109/TII.2019.2892818

Leukel, J., González, J., & Riekert, M. (2021). Adoption of machine learning technology for failure prediction in industrial maintenance: A systematic review. *Journal of Manufacturing Systems*, *61*, 87–96. DOI: 10.1016/j.jmsy.2021.08.012

Li, C., Mo, L., & Yan, R. (2021). Rotating machinery fault diagnosis based on spatial–temporal GCN. In *2021 International Conference on Sensing, Measurement & Data Analytics in the era of Artificial Intelligence (ICSMD)*, Nanjing, China, 2021, (pp. 1-6). IEEE. https://doi.prg/10.1109/ICSMD53520.2021.9670851

Li, N., Zhao, Y., Li, D., & Guan, W. (2021). Real-time prediction of highway equipment faults based on GCN and GRU algorithms. In *2021 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, Melbourne, Australia, 2021 (pp. 3252–3257). IEEE. DOI: 10.1109/SMC52423.2021.9658637

Li, Z., Li, J., Wang, Y., & Wang, K. (2019). A deep learning approach for anomaly detection based on SAE and LSTM in mechanical equipment. *International Journal of Advanced Manufacturing Technology*, *103*(1-4), 499–510. DOI: 10.1007/s00170-019-03557-w

Li, Z., Wang, Y., & Wang, K. (2019). A deep learning driven method for fault classification and degradation assessment in mechanical equipment. *Computers in Industry*, *104*, 1–10. DOI: 10.1016/j.compind.2018.07.002

Liu, Y., Liang, H., Xiao, Y., Zhang, H., Zhang, J., Zhang, L., & Wang, L. (2022). Logistics-involved service composition in a dynamic cloud manufacturing environment: A DDPG-based approach. *Robotics and Computer-integrated Manufacturing*, *76*, 102323. DOI: 10.1016/j.rcim.2022.102323

Liu, Y., Qin, H., Zhang, Z., Pei, S., Wang, C., Yu, X., Jiang, Z., & Zhou, J. (2019). Ensemble spatiotemporal forecasting of solar irradiation using variational Bayesian convolutional gate recurrent unit network. *Applied Energy*, *253*, 113596. DOI: 10.1016/j.apenergy.2019.113596

Lu, S., Liu, S., Zhu, Y., Liang, W., Li, K., & Lu, Y. (2022). A DRL-based decentralized computation offloading method: An example of an intelligent manufacturing scenario. *IEEE Transactions on Industrial Informatics*, *19*(9), 9631–9641. DOI: 10.1109/TII.2022.3227652

Luo, D., Fang, J., He, H., Lee, W.-J., Zhang, Z., Zai, H., Chen, W., & Zhang, K. (2022). Prediction for dissolved gas in power transformer oil based on TCN and GCN. *IEEE Transactions on Industry Applications*, *58*(6), 7818–7826. DOI: 10.1109/TIA.2022.3197565

Ma, S., Wang, Y., & Zhang, S. (2022). Improved artificial bee colony algorithm for multimodal optimization based on crowding method. *Journal of Organizational and End User Computing*, *34*(3), 1–18. DOI: 10.4018/JOEUC.302661

Namuduri, S., Narayanan, B. N., Davuluru, V. S. P., Burton, L., & Bhansali, S. (2020). Deep learning methods for sensor based predictive maintenance and future perspectives for electrochemical sensors. *Journal of the Electrochemical Society*, *167*(3), 037552. DOI: 10.1149/1945-7111/ab67a8

Panchi, L., Tong, S. H. I., & Xuegui, L. (2022). Denoising method for microseismic data based on recurrent neural networks. *Journal of Jilin University Science Edition*, *60*(3), 685–696.

Qi, Q., & Tao, F. (2019). A smart manufacturing service system based on edge computing, fog computing, and cloud computing. *IEEE Access : Practical Innovations, Open Solutions*, *7*, 86769–86777. DOI: 10.1109/ACCESS.2019.2923610

Scalabrini Sampaio, G., Vallim Filho, A. R. A., Santos da Silva, L., & Augusto da Silva, L. (2019). Prediction of motor failure time using an artificial neural network. *Sensors (Basel)*, *19*(19), 4342. DOI: 10.3390/s19194342 PMID: 31597304

Shafiee, M., & Sørensen, J. D. (2019). Maintenance optimization and inspection planning of wind energy assets: Models, methods and strategies. *Reliability Engineering & System Safety*, *192*, 105993. DOI: 10.1016/j.ress.2017.10.025

Shahbazi, Z., & Byun, Y.-C. (2021). Smart manufacturing real-time analysis based on blockchain and machine learning approaches. *Applied Sciences (Basel, Switzerland)*, *11*(8), 3535. DOI: 10.3390/app11083535

Soltanali, H., Khojastehpour, M., Farinha, J. T., & Pais, J. E. D. A. E. (2021). An integrated fuzzy fault tree model with Bayesian network-based maintenance optimization of complex equipment in automotive manufacturing. *Energies*, *14*(22), 7758. DOI: 10.3390/en14227758

Souza, R. M., Nascimento, E. G., Miranda, U. A., Silva, W. J., & Lepikson, H. A. (2021). Deep learning for diagnosis and classification of faults in industrial rotating machinery. *Computers & Industrial Engineering*, *153*, 107060. DOI: 10.1016/j.cie.2020.107060

Wang, J., Yan, J., Li, C., Gao, R. X., & Zhao, R. (2019). Deep heterogeneous GRU model for predictive analytics in smart manufacturing: Application to tool wear prediction. *Computers in Industry*, *111*, 1–14. DOI: 10.1016/j.compind.2019.06.001

Wang, Q., Bu, S., & He, Z. (2020). Achieving predictive and proactive maintenance for high-speed railway power equipment with LSTM-RNN. *IEEE Transactions on Industrial Informatics*, *16*(10), 6509–6517. DOI: 10.1109/TII.2020.2966033

Wang, Y., Chen, J., Chen, X., Zeng, X., Kong, Y., Sun, S., Guo, Y., & Liu, Y. (2020). Short-term load forecasting for industrial customers based on TCN-LightGBM. *IEEE Transactions on Power Systems*, *36*(3), 1984–1997. DOI: 10.1109/TPWRS.2020.3028133

Yuan, X., Qi, S., Wang, Y., Wang, K., Yang, C., & Ye, L. (2021). Quality variable prediction for nonlinear dynamic industrial processes based on temporal convolutional networks. *IEEE Sensors Journal*, *21*(18), 20493–20503. DOI: 10.1109/JSEN.2021.3096215

Yueze, L., Xianglin, Z., Wanli, Z., Shining, L., Yijia, Z., & Yuan, Z. (2023). Causality extraction based on BERT–GCN. *Journal of Jilin University Science Edition*, *61*(2), 325–330.

Zhang, J., Zhang, F., Guo, X., Yang, X., & Mu, P. (2022). A predictive maintenance model of equipment system based on GRU model. In *2022 International Conference on Virtual Reality, Human–Computer Interaction and Artificial Intelligence (VRHCIAI)*, Changsha, China, 2022 (pp. 100–105). IEEE. DOI: 10.1109/VRHCIAI57205.2022.00024

Zhang, L., Liu, J., Wei, Y., An, D., & Ning, X. (2025). Self-supervised learning-based multi-source spectral fusion for fruit quality evaluation: A case study in mango fruit ripeness prediction. *Information Fusion*, *117*, 102814. DOI: 10.1016/j.inffus.2024.102814

Zhang, W., Yang, D., & Wang, H. (2019). Data-driven methods for predictive maintenance of industrial equipment: A survey. *IEEE Systems Journal*, *13*(3), 2213–2227. DOI: 10.1109/JSYST.2019.2905565

Zhao, Y., & Zhou, Y. (2022). Measurement method and application of a deep learning digital economy scale based on a big data cloud platform. *Journal of Organizational and End User Computing*, *34*(3), 1–17. DOI: 10.4018/JOEUC.295092

Zhu, J., Jiang, Q., Shen, Y., Qian, C., Xu, F., & Zhu, Q. (2022). Application of recurrent neural network to mechanical fault diagnosis: A review. *Journal of Mechanical Science and Technology*, *36*(2), 527–542. DOI: 10.1007/s12206-022-0102-1