Proceedings of The 6th International Conference on Signal Processing and Machine Learning
DOI: 10.54254/2755-2721/2026.30832

Potential Fault Prediction of Industrial Robots Based on Machine Learning Algorithms

Ziang Wu

New York University, New York, USA
cyruswu2025@outlook.com.

Abstract. The intelligent manufacturing industry is accelerating its upgrading. Industrial robots have become the core support of the flexible production system, and their operational status directly affects the continuous operation capacity and safety of the production line. In large-scale production scenarios, industrial robots are constantly operating under high loads and multiple conditions. Potential faults such as joint wear, abnormal load transmission, and motor performance degradation tend to accumulate continuously. If not predicted in time, it can easily lead to unplanned shutdowns, resulting in reduced production efficiency and increased operation and maintenance costs. Aiming at the problems of insufficient feature focusing and inadequate dynamic correlation capture in the fault prediction of industrial robots by existing machine learning algorithms, this paper proposes the RBMO-BiLSTM-Attention classification algorithm. Firstly, violin graph analysis and correlation analysis are carried out, and then comparative experiments are conducted through multiple machine learning algorithms. The results show that among the traditional machine learning algorithms, decision trees and random forests perform relatively outstandingly, with accuracy and recall rates both reaching 94.1%, while the performance of other algorithms is slightly inferior. The proposed algorithm has significant advantages in all evaluation indicators. Its accuracy rate, recall rate, precision rate, F1 value all reach 95.9%, and the AUC value is as high as 97.8%. It has stronger feature extraction capabilities and classification performance, and can more accurately complete the risk level classification task of industrial Internet of Things devices. This research provides an effective technical solution for the precise prediction of industrial robot faults, which is of great significance for reducing unplanned downtime, improving production efficiency and the economic efficiency of operation and maintenance.

Keywords: Cloud computing, load prediction, BiLSTM, Transformer.

1. Introduction

The rapid upgrading of the intelligent manufacturing industry has driven industrial robots to become the core carriers of flexible production systems, and their operational status is directly related to the continuous operation capacity and safety level of production lines [1]. In large-scale production scenarios, industrial robots are constantly operating under high loads and multiple conditions. Potential faults such as joint wear, abnormal load transmission, and motor performance degradation

DOI: 10.54254/2755-2721/2026.30832

tend to accumulate gradually. If not predicted in time, it can easily lead to unplanned shutdowns, resulting in a decline in production efficiency and an increase in operation and maintenance costs [2]. Traditional fault prediction methods mostly rely on human experience or a single physical model, which is difficult to adapt to the dynamic change characteristics of time series sensor data during the operation of robots, and also cannot effectively mine the physical correlation and coupling rules among multi-dimensional signals, resulting in the accuracy and timeliness of fault early warning being unable to meet the high requirements of intelligent manufacturing [3].

Machine learning algorithms provide data-driven solutions for the fault prediction of industrial robots and have become the core technical path to break through the limitations of traditional methods [4]. Compared with traditional threshold judgment or mechanism modeling, machine learning algorithms can automatically extract hidden fault features from massive time series sensing data and capture the subtle changes in equipment operation status without relying on precise physical models. For instance, recurrent neural network-based algorithms can effectively handle the sequence-dependent characteristics of time series data, and convolutional neural networks can extract local feature associations. However, existing algorithms still have shortcomings - a single recurrent network is prone to ignore the weight differences in key fault feature intervals, and the specificity of feature extraction is insufficient. When facing the coupling relationship of multi-dimensional sensing data, The generalization ability and prediction accuracy of the model still have room for improvement [5].

Aiming at the problems of insufficient feature focusing and inadequate dynamic association capture in the fault prediction of industrial robots by existing machine learning algorithms, this paper proposes the RBMO-BiLSTM-Attention classification algorithm. This algorithm is based on the bidirectional long short-term memory network, fully explores the forward and backward dependencies of temporal sensing data, and ADAPTS to the temporal characteristics of robot operation data. Introduce the attention mechanism to assign differentiated weights to the key fault characteristic intervals, and enhance the ability to capture core fault precursor signals such as joint torque and vibration amplitude. At the same time, the RBMO algorithm is used to optimize the hyperparameter configuration of the model, solving the problems that the traditional BiLSTM model is prone to fall into local optimum and has a slow convergence speed. This further improves the classification and recognition accuracy of the algorithm for potential faults of industrial robots, providing more reliable technical support for predictive maintenance of equipment in intelligent manufacturing scenarios [6].

2. Data sources

This dataset contains 729 pieces of operational status data of industrial robots, covering 12 input variables and 1 binary predictive variable. The input variables include the Angle, angular velocity, torque, end effector load, load transfer efficiency, joint temperature, vibration amplitude and motor current of joint 1 and joint 2. The predictive variable is the fault label value of 0 or 1. They respectively represent normal operation and fault status. Calculate the correlations among various variables and draw a correlation heat map, as shown in Figure 1.

DOI: 10.54254/2755-2721/2026.30832

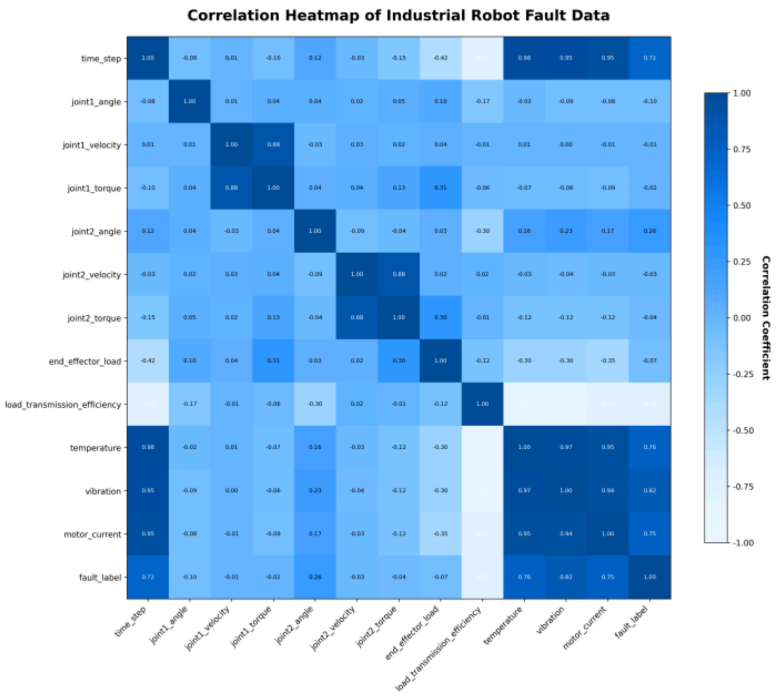

Figure 1. The correlation heat map

Violin plot analysis was conducted on each variable, and a violin plot was drawn as shown in Figure 2.

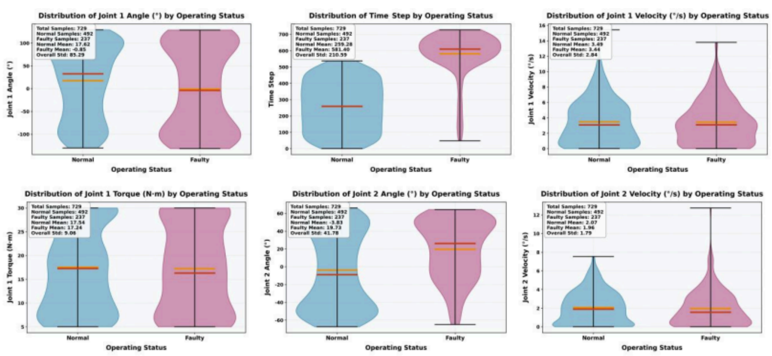

Figure 2. The violin plots of each variable and observe the distribution

3. Method

3.1. RBMO

The Red-billed blue magpie optimization algorithm simulates the foraging, alerting and group collaboration behaviors of the red-billed blue magpie to achieve optimization. The dynamic process of individual position update can be described by a first-order ordinary differential equation:

$$\frac{dx_i(t)}{d_t} = \alpha \cdot rand() \cdot (x_{best} - x_i(t)) + \beta \cdot (x_{neighbor}) - x_i(t)) \tag{1}$$

Here, $x_i(t)$ represents the position of the t candidate solution at step t of the iteration, α is the global exploration coefficient (corresponding to wide-area foraging), β is the local development

DOI: 10.54254/2755-2721/2026.30832

coefficient (corresponding to alert behavior), x_{best} is the global optimal solution, and $x_{neighbor}$ is the neighborhood optimal solution. By adjusting α and β to balance global exploration and local development, traditional algorithms can be avoided from falling into local optimum, which is suitable for model hyperparameter optimization scenarios.

3.2. BiLSTM

The bidirectional Long short-term memory network (BiLSTM) achieves long-term temporal information storage through a gating mechanism, and the dynamic evolution of its cell state can be expressed by ordinary differential equations:

$$\frac{dC_t}{dt} = f_t \cdot C_{t-1} + i_t \cdot {}_{C}t \tag{2}$$

Among them, C_t represents the cell state at time t , f_t is the output of the forget gate (controlling the retention of historical information), i_t is the output of the input gate (controlling the input of new information), and $\tilde{_C}t$ are the candidate cell states. After the forward and backward branches of BiLSTM are fused, the temporal dependence is completely characterized. In the processing of industrial robot sensor data, it can accurately describe the continuous change law of equipment operation status over time.

3.3. Attention

The attention mechanism focuses on key information by assigning differentiated weights to the input features. The model can automatically locate the key time series intervals corresponding to the weight peaks (such as the precursors of industrial robot failures) [7]. This weight distribution modeling avoids the limitations of the traditional fixed weight method. It can dynamically adjust the weights based on the correlation between the features and the target task, weaken noise interference, and enhance the pertinence of feature extraction [8].

3.4. RBMO-BiLSTM-Attention

The RBMO-BiLSTM-Attention algorithm integrates the advantages of three technologies. Firstly, the RBMO algorithm uses ODE to describe the dynamic update of candidate hyperparameters, optimizes parameters such as the number of hidden layer nodes and learning rate of BiLSTM, and ensures that the network structure ADAPTS to the data characteristics. Then, BiLSTM characterizes the temporal evolution of cell states and extracts bidirectional temporal features [9]; Finally, the attention mechanism focuses on the key features related to faults through the spatial distribution of temporal feature weights [10]. The three are dynamically related. It not only enhances the network stability through the optimization of RBMO, but also ADAPTS to time series data with BiLSTM and strengthens key features through Attention, significantly improving the fault classification accuracy of industrial robots.

4. Result

The parameter Settings of the project include optimization parameters and model training parameters. In terms of optimization, the population size is 12 and the maximum number of iterations is 6. The three parameters that need to be optimized are the L2 regularization coefficient,

DOI: 10.54254/2755-2721/2026.30832

the initial learning rate, and the number of hidden layer nodes. Their lower limits are 1e-5, 0.0001, and 10 respectively, and their upper limits are 1e-2, 0.002, and 50 respectively. During model training, the maximum number of training rounds is 500, the gradient threshold is 1, the initial learning rate is obtained through optimization, the learning rate adjustment strategy is piecewise, the adjustment period is 400, the adjustment factor is 0.2, the L2 regularization coefficient is obtained through optimization, the training environment is cpu, and the dropout layer discard rate is 0.01.

The comparison results of various machine learning algorithms are shown in Table 1. The comparison bar chart of each indicator is shown in Figure 3.

Table 1. The results of the comparative experiment

Model	Accuracy	Recall	Precison	F1	AUC
AdaBoost	0.936	0.936	0.937	0.936	0.967
Decision tree	0.941	0.941	0.942	0.939	0.94
GBDT	0.927	0.927	0.927	0.926	0.94
Random Forest	0.941	0.941	0.944	0.94	0.954
CatBoost	0.927	0.927	0.935	0.925	0.943
ExtraTrees	0.936	0.936	0.941	0.934	0.957
XGBoost	0.936	0.936	0.938	0.935	0.966
RBMO-BiLSTM-Attention	0.959	0.959	0.959	0.959	0.978

From the results of the comparative experiments, it can be seen that the performance of decision trees and random forests in traditional machine learning algorithms is relatively outstanding. The accuracy rates of both are 94.1%, the recall rates are both 94.1%, the precision rates are 94.2% and 94.4% respectively, the F1 values are 93.9% and 94.0% respectively, and the AUC values are 94.0% and 95.4% respectively. The performance indicators of AdaBoost, ExtraTrees and XGBoost are similar. Their accuracy rates are all 93.6%, recall rates are all 93.6%, precision rates are 93.7%, 94.1% and 93.8% respectively, and F1 values are 93.6%, 93.4% and 93.5% respectively. The AUC values were 96.7%, 95.7% and 96.6% respectively. The performance of GBDT and CatBoost was relatively weaker, with accuracy and recall rates both at 92.7%, precision rates at 92.7% and 93.5% respectively, F1 values at 92.6% and 92.5% respectively, and AUC values both around 94.0%. The RBMO-BiLSTM-Attention algorithm proposed in this paper demonstrates significant advantages in all evaluation indicators. The accuracy rate, recall rate, precision rate and F1 value all reach 95.9%, and the AUC value is as high as 97.8%. All indicators are significantly higher than those of traditional machine learning algorithms, demonstrating stronger feature extraction ability and classification performance. It can more accurately complete the risk level classification task of industrial Internet of Things devices. Although traditional machine learning algorithms perform similar in some indicators, their overall classification accuracy and generalization ability are both inferior to the proposed algorithm.

DOI: 10.54254/2755-2721/2026.30832

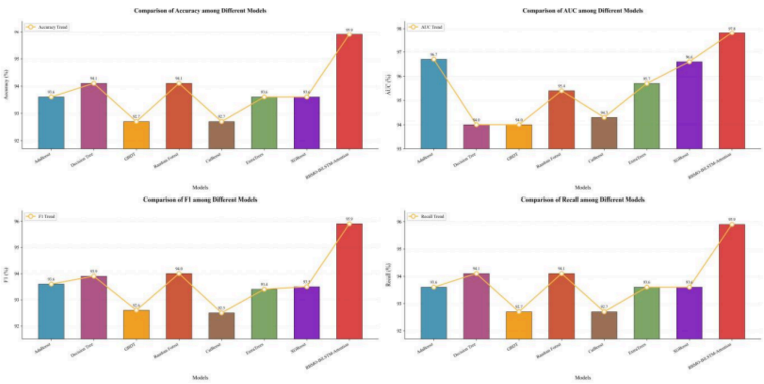

Figure 3. The comparison bar chart of each indicator

Output the confusion matrix of the RBMO-BiLSTM-Attention test set, as shown in the figure 4.

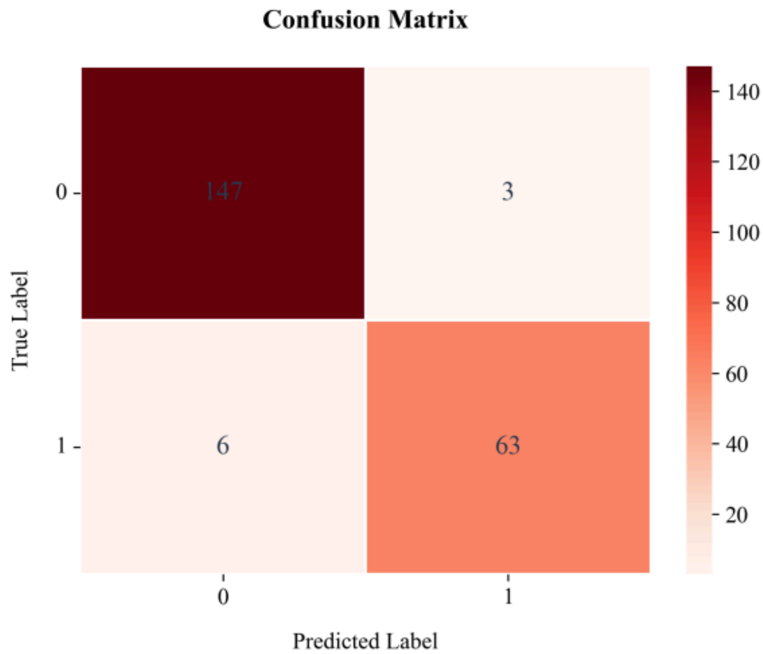

Figure 4. The confusion matrix of the RBMO-BiLSTM-Attention

5. Conclusion

The accelerated upgrading of the intelligent manufacturing industry has made industrial robots the core of the flexible production system. Their operational status directly affects the continuous operation capacity and safety of the production line. In large-scale production, industrial robots operate under high loads and multiple working conditions for a long time. Potential faults such as joint wear, abnormal load transmission, and motor performance degradation will gradually accumulate. If not predicted in time, it is easy to cause unplanned shutdowns, resulting in a decline in production efficiency and an increase in operation and maintenance costs. Aiming at the problems of insufficient feature focusing and inadequate dynamic correlation capture in the fault prediction of industrial robots by existing machine learning algorithms, this paper proposes the RBMO-BiLSTM-Attention classification algorithm. Firstly, violin graph analysis and correlation analysis are carried

DOI: 10.54254/2755-2721/2026.30832

out, and then comparative experiments are conducted through multiple machine learning algorithms. The results show that the decision tree and random forest in the traditional algorithm perform better, with an accuracy rate and recall rate of 94.1% respectively, while other indicators vary. The indicators of AdaBoost, ExtraTrees and XGBoost are similar, with accuracy and recall rates both being 93.6%. GBDT and CatBoost performed slightly weaker, with accuracy and recall rates both at 92.7%. The algorithm proposed in this paper leads significantly in all indicators. Its accuracy rate, recall rate, precision rate, F1 value reach 95.9%, and the AUC value reaches 97.8%. It has stronger feature extraction and classification performance and can more accurately complete the risk level classification of industrial Internet of Things devices. The overall accuracy and generalization ability of traditional algorithms are inferior to it. This algorithm provides a reliable technical solution for the precise prediction of industrial robot faults, which is of great practical significance for ensuring the stable operation of production lines and reducing operation and maintenance costs.

References

[1] Geng, Sujie, and Xiuli Wang. "Predictive maintenance scheduling for multiple power equipment based on data-driven fault prediction." Computers & Industrial Engineering 164 (2022): 107898.
[2] Panda, Prasanta, Debaryaan Sahoo, and Debarjun Sahoo. "Automating Fault Prediction in Software Testing Using Machine Learning Techniques: A Real-World Applications." 2024 2nd International Conference on Sustainable Computing and Smart Systems (ICSCSS). IEEE, 2024.
[3] Zhou, X. W., et al. "An Fe–Ni–Cr–H interatomic potential and predictions of hydrogen-affected stacking fault energies in austenitic stainless steels." International Journal of Hydrogen Energy 47.1 (2022): 651-665.
[4] Shang, Jie. "Simulation of Power System Fault Detection and Prediction Model Based on Deep Learning." 2024 Asia-Pacific Conference on Software Engineering, Social Network Analysis and Intelligent Computing (SSAIC). IEEE, 2024.
[5] Deng, Jian, et al. "Research on equipment fault diagnosis model based on gan and inverse PINN: Solutions for data imbalance and rare faults." PLoS One 20.5 (2025): e0324180.
[6] Ūselis, Robertas, Raimondas Pomarnacki, and Artūras Serackis. "Developing a Machinery Fault Prediction Solution in Resource-Constrained Environments." 2025 IEEE 12th Workshop on Advances in Information, Electronic and Electrical Engineering (AIEEE). IEEE, 2025.
[7] Tao, Laifa, et al. "LLM-based framework for bearing fault diagnosis." Mechanical Systems and Signal Processing 224 (2025): 112127.
[8] Kumar, Kailash, et al. "Intelligent controller design and fault prediction using machine learning model." International Transactions on Electrical Energy Systems 2023.1 (2023): 1056387.
[9] Mase, Lindung Zalbuin, et al. "Analysis of ground response and potential seismic damage to sites surrounding Cimandiri Fault, West Java, Indonesia." Natural Hazards 119.3 (2023): 1273-1313.
[10] Rahman, Md M., et al. "Accelerating cavity fault prediction using deep learning at Jefferson Laboratory." Machine Learning: Science and Technology 5.3 (2024): 035078.