

## Retraction

# Retracted: Research on Motor Bearing Fault Diagnosis Based on the AdaBoost Algorithm and the Ensemble Learning with Bayesian Optimization in the Industrial Internet of Things

### Security and Communication Networks

Received 10 October 2023; Accepted 10 October 2023; Published 11 October 2023

Copyright © 2023 Security and Communication Networks. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

### References

- [1] J. Jiang, Y. Liu, C. Xu, and H. Shen, "Research on Motor Bearing Fault Diagnosis Based on the AdaBoost Algorithm and the Ensemble Learning with Bayesian Optimization in the Industrial Internet of Things," *Security and Communication Networks*, vol. 2022, Article ID 4569954, 12 pages, 2022.

## Research Article

# Research on Motor Bearing Fault Diagnosis Based on the AdaBoost Algorithm and the Ensemble Learning with Bayesian Optimization in the Industrial Internet of Things

Jiahui Jiang <sup>1</sup>, Yuqing Liu <sup>1</sup>, Chaozheng Xu <sup>2</sup>, and Hanting Shen <sup>1</sup>

<sup>1</sup>College of Engineering Science and Technology, Shanghai Ocean University, Shanghai 201306, China

<sup>2</sup>College of Electrical and Electronic Engineering, Shijiazhuang Tiedao University, Hebei, Shijiazhuang 050043, China

Correspondence should be addressed to Yuqing Liu; [yqliu@shou.edu.cn](mailto:yqliu@shou.edu.cn)

Received 12 May 2022; Revised 13 June 2022; Accepted 29 June 2022; Published 18 July 2022

Academic Editor: Mukesh Soni

Copyright © 2022 Jiahui Jiang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In order to adapt to the development of the industrial Internet of Things, the relationship between the internal components of electromechanical equipment is getting closer and closer, such as motor bearings. Nowadays, timely diagnosis of motor bearing faults is urgently needed. Most traditional methods for motor bearing fault diagnosis use a single learner and emphasize the role of feature extraction, which usually requires a large amount of sample support and computer runtime to obtain satisfactory performance. In this article, the Bayesian optimized decision tree with ensemble classifiers after feature extraction of the original data is finally proposed which has good performance. We use multiple feature extraction to establish the feature matrix and construct a decision tree model with the ensemble method for AdaBoost and a Bayesian optimized decision tree model with ensemble classifiers to conduct experiments on the accuracy, prediction speed, etc., of the model. We derived four sets of experimental data. The results show that the optimal method is the Bayesian optimized decision tree with ensemble classifiers after feature extraction. The accuracy of this method is as high as 99.9%. At the same time, unlike previous studies, we found in our study that feature extraction does not improve the accuracy of diagnosis for the decision tree with ensemble method for AdaBoost and there is a precipitous decline. In the industrial Internet of Things, the conclusion can improve certain reference value for the future fault diagnosis of motor bearings.

## 1. Introduction

With the continuous development of the Internet of Things (IoT), information technology, and sensor technology, various terminals with environmental awareness, computing models based on ubiquitous technologies, and mobile communication technologies are being continuously integrated into all aspects of industrial production. As an application of Internet of Things technology in industry, the core concept of the industrial Internet of Things is an interdisciplinary combination involving several disciplines such as network communication, information security, and automation [1]. The use of various emerging technologies in the industrial Internet of Things can significantly improve manufacturing efficiency, improve product quality, and

reduce product costs and resource consumption [2]. Simultaneously, in order to adapt to the development of the industrial Internet of Things, electromechanical equipment develops in the direction of large-scale, high-speed, precision, systematization, and automation, and the relationship between components within the equipment becomes closer.

Motor bearings play an important role in industrial production, and bearings are one of the most important components of rotating machines; bearing fault can lead to mechanical failure, economic loss, and even personal injury; and their monitoring and fault diagnosis can provide a reliable guarantee for the normal operation of motors [3–5]. Therefore, effective fault monitoring and diagnosis of bearings in rotating machinery and equipment are of great relevance to promote the development of industrial

networking. Motor bearings are prone to faults on the inner race, rolling element, and outer race, and if the faulty bearings are not detected and run with load, serious safety accidents can occur. Accurate fault diagnosis is the key to ensure the safe and reliable operation of rotating machinery. In the era of the industrial Internet of Things, the motor bearing fault diagnosis method with high recognition accuracy and low leakage rate has become a hot spot for domestic and international research [6–8].

For the problem of motor bearing fault detection, a lot of research has been conducted by domestic and foreign scholars. Lucena-Junior et al. proposed a technique for detecting three-phase induction motor bearing faults using acoustic signals collected by a single sensor [9]. Kim et al. proposed empirical mode decomposition (EMD) and probabilistic filtering techniques to eliminate interference peaks in the acoustic emission data [10]. Hiruta et al. proposed a Gaussian mixture model (GMM) to show the increase of abnormal bearing condition corresponding to insufficient grease [11]. Wang et al. proposed the improved cyclostationary analysis method based on TKEO [12]. Zhang et al. proposed a DCGAN-RCCNN permanent magnet motor fault diagnosis model, which relies on stator current data to detect permanent magnet motor faults [13]. Nikfar et al. suggested that integrating machine learning components as part of a predictive maintenance system could improve confidence in the condition of the motor, reduce maintenance costs, and enhance operator and machine safety [14]. Wang et al. proposed a new attention-guided joint learning convolutional neural network (JL-CNN) for condition monitoring of mechanical equipment [15]. Zhi et al. performed feature extraction for motor faults based on decision trees, and found that there was a great improvement in accuracy and diagnostic speed compared to the traditional CART algorithm [16]. Wang et al. proposed an integrated fault diagnosis and prediction method based on wavelet transform and particle filtering, which can infer the hidden defect status of bearings from noise measurements by Bayesian inference [17]. Kong et al. proposed a wind turbine condition monitoring method based on the fusion of spatiotemporal features of GRU SCADA data with good performance in a certain data range [18]. Chang et al. proposed a novel neural network structure for wind turbine fault diagnosis, which can adaptively extract generic features from the original vibration signal [19]. Yang et al. investigated a multipoint data fusion-assisted noise suppression method for feature frequency extraction, which can effectively suppress white noise and short-term disturbance noise [20]. In addition, Bhatnagar study found that in the field of IoT technology, machine learning techniques compared to traditional classification and prediction methods, these algorithms not only increase the processing speed but also produce better results [21]. They can recover lost data, eliminate noise, promote messaging, form network classifiers, and predict the state and location of IoT devices in order to process data faster and more accurately, resulting in faster and better results. Artificial intelligence methods such as machine learning are heavily used in the field of fault diagnosis and are effective. Further improvement in

traditional machine learning methods through optimization algorithms will help to further improve the accuracy of fault diagnosis.

Typically, motor vibration signals are collected from data acquisition systems. Signal processing methods such as time domain, frequency domain, and time-frequency domain are used to analyze the signals to extract sensitive and robust features for fault type identification. Signal acquisition and transmission can be implemented using cable or wireless technology. Cable transmission can provide high throughput rates and power supply capabilities. However, the capability of cable transmission is limited by the transmission distance and operating environment. In contrast, IoT technologies collect signals from distributed motors and transmit them through wireless communication [22, 23]. Thus, the Internet of Things offers great flexibility and convenience for remote motor troubleshooting. IOT nodes can be installed on industrial motors for condition monitoring and fault diagnosis. Two MEMS accelerometers are installed at both ends of the motor, and the vibration signals are collected by the Internet of Things node. The signals are transmitted through a GPRS network and received by a remote server for further analysis and fault diagnosis [24].

In summary, the current research on motor bearing fault diagnosis consists of two main aspects: feature extraction and classification algorithms. A key step that affects the accuracy of fault diagnosis is feature extraction and feature selection. The fault features of motor bearings in the early damage stage are usually very weak, so the study of feature extraction and selection methods that can effectively extract fault features becomes a breakthrough to solve the fault diagnosis problem [25]. The Internet of Things technology transmits signals acquired from decentralized nodes to routers and then transmits them to data centers or clouds for further processing. IoT technology provides a convenient and flexible network model as it does not require complex cable wiring. IoT nodes can also be easily added, removed, and replaced as per the requirement of conditional monitoring [26].

Most traditional methods for motor bearing fault diagnosis use a single learner and emphasize the role of feature extraction, which usually requires a large amount of sample support and computer runtime to obtain satisfactory performance. This study is based on multiple feature extraction, decision tree with ensemble method for AdaBoost and Bayesian optimized decision tree with ensemble classifiers for fault diagnosis study of motor bearing equipment, and the final conclusion has some scientific significance and reference value.

## 2. Materials and Methods

*2.1. Data Acquisition.* The data were obtained from experiments conducted at the Bearing Data Center at Case Western Reserve University. The test equipment is shown in Figure 1. Experiments were conducted using 2 hp Reliance Electric motors, and acceleration data were measured at locations near and away from the motor bearings. The motor bearings were implanted with failures using electric

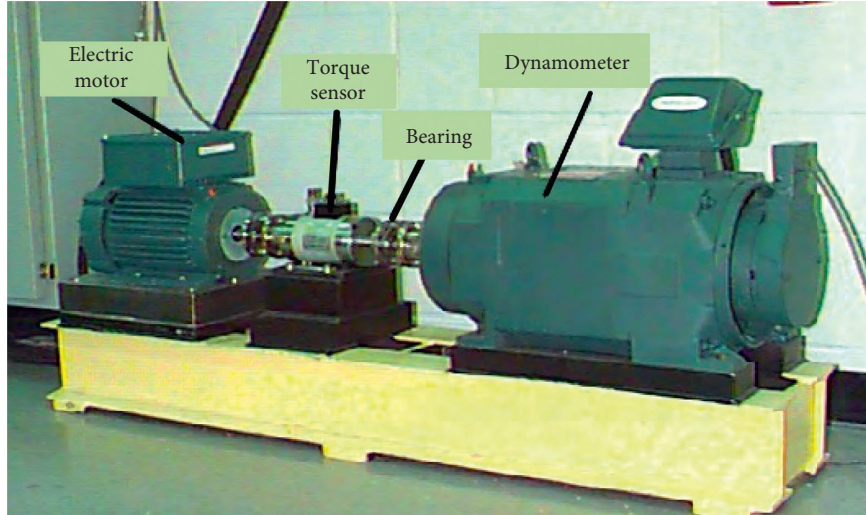


FIGURE 1: Bearing test system equipment.

TABLE 1: Four bearing condition types and labels.

Label	Fault type	Sampling frequency (Hz)	Damage to the inner diameter (mm)
0	Normal bearing	1200	0
1	Rolling element fault	1200	0.1778
2	Inner race fault	1200	0.3556
3	Outer race fault	1200	0.5334

discharge machining (EDM). Faults ranging from 0.007" to 0.040" in diameter were introduced in the inner race, rolling element, and outer race, respectively. The faulty bearings were reinstalled into the test motor, and vibration data were recorded for motor loads of 0 to 3 hp (motor speeds of 1797 to 1720 rpm).

The test equipment was a motor with a power of 1500 watts and the bearing under test supported the motor. The experimental data were measured at a sampling frequency of 1200 Hz and a speed of 1772 r/min. The four types included normal bearings, rolling element failed bearings, inner ring failed bearings, and outer ring failed bearings. The latter three bearings were damaged inside diameters of 0.1778 mm, 0.3556 mm, and 0.5334 mm, respectively. The measured data had only one vibration signal per fault, but the vibration signal was long, so it was intercepted once every 200 sampling points, and then the time domain, frequency domain, and distance characteristics were calculated separately, as shown in Table 1.

## 2.2. Feature Extraction

**2.2.1. Time Domain Analysis.** Time domain analysis enables to analyze the stability and transient and steady-state performance of the system based on the time domain expressions of the output quantities. Extraction includes dimensional features, such as maximum, peak, and mean values, and dimensionless features, such as waveform factor, pulse factor, margin factor, and margin factor, for analysis [27]. Time domain analysis offers the advantages of intuition

and accuracy, but it is susceptible to noise signal interference and error. In motor bearing fault diagnosis, time domain features can usually be extracted from the bearing fault data set.

**2.2.2. Frequency Domain Analysis.** The frequency domain analysis is based on the principle of splitting the original waveform into a number of harmonic components of different frequencies after the Fourier transform, and through the analysis, processing and filtering of specific components, more discriminative data features can be obtained [27]. In motor bearing fault diagnosis, the procedure of mapping time domain data to frequency domain by Fourier transform is called time domain to frequency domain conversion. The analysis of frequency domain indicators including center of gravity frequency and mean square frequency in the frequency domain leads to the diagnostic analysis of the bearing signal.

**2.2.3. Mahalanobis Distance.** The Mahalanobis distance is a method to calculate the similarity of unknown samples, and the use of Marxist distance can consider the connection between the features and allows the comparative analysis of different measures [28].

For the sample space  $X = \{x_z/1 \leq z \leq m\} \subset R^n$  containing  $m$  samples, the marginal distance from one of the sample points  $x_z$  to the sample mean  $u_x$  is

$$d_M(x_z, u_x) = \sqrt{(x_z - u_x)^T C_X^{-1} (x_z - u_x)}, \quad (1)$$

where  $C_X$  is the covariance matrix of the sample space  $X$ . The equations for  $C_X$  and  $u_X$  are as follows:

$$\begin{aligned} u_X &= \sum_{z=1}^m \frac{x_z}{m}, \\ C_X &= \sum_{z=1}^m \left[ \frac{(x_z - u_X)^T (x_z - u_X)}{m} \right]. \end{aligned} \quad (2)$$

**2.3. AdaBoost Algorithm.** The AdaBoost algorithm was proposed by Yoav Freund and Robert Schapire in 1995. It is adaptive in the sense that samples that are incorrectly classified by the previous basic classifier are strengthened, the weights of the accurately classified samples are reduced, and the weighted whole sample is used again to train the next basic classifier. At the same time, a new weak classifier is added in each round until some predefined sufficiently small error rate is reached or a prespecified maximum number of iterations is reached. After the training process of each weak classifier is completed, the weights of the weak classifiers with small classification error rates are increased so that the weak classifiers with low error rates take up a larger weight in the final classifier and the weak classifiers with high error rates take up a smaller weight, and finally, the weak classifiers obtained from each training are combined into strong classifiers.

Assume that the training sample set is

$$T = \{(x_1, y_1), (x_2, y_2), \dots, (x_m, y_m)\}, \quad (3)$$

where  $x$  is the instance space,  $y$  is the mark set, and  $T$  is the number of training set samples.

The weight distribution of the trainer is initialized, and the training set at the  $k$ th weak learner is trained with the following weights.

$$\begin{aligned} D(k) &= (w_{k1}, w_{k2}, \dots, w_{km}); \\ w_{ki} &= \frac{1}{m}; i = 1, 2, \dots, m; k = 1, 2, \dots, K, \end{aligned} \quad (4)$$

where  $m$  is the number of samples,  $k$  is the number of weak learners,  $D$  is the model weight,  $w$  is the sample weight, and  $w_{ki}$  is the weight of the  $i$ th sample in the  $k$ th training.

The first iteration is performed first, that is the value of  $k$  is 1. According to the optional threshold, a threshold with the smallest classification error rate in the current trainer is selected, and the classification error rate at this time is calculated as

$$e_k = P(G_k(x_i) \neq y_i) = \sum_{i=1}^m w_{ki} I(G_k(x_i) \neq y_i), \quad (5)$$

where  $e_k$  is the sum of weights of sample points with the wrong classification, and  $G_k(x)$  is a base learner. If the actual value is different from the predicted value, the output is 1; if it is the same, the output is 0. All sample points are multiplied with the wrong classification by the sample weight and then accumulated.

The weight coefficient of the  $k$ th weak classifier  $G_k(x)$  is:  $\alpha_k = (1/2) \log_1 - e_k/e_k$ , substituting  $f(x)$  additive model,

$$f(x) = \sum_{k=1}^K \alpha_k G_k(x). \quad (6)$$

If the accuracy requirement is reached in the first iteration, the operation is stopped. If the requirement is not met, the weights are updated.

$$w_{k+1,i} = \frac{w_{ki}}{Z_k} \exp(-\alpha_k y_i G_k(x_i)), \quad (7)$$

where  $Z_k$  is the normalization factor, and  $\exp$  refers to the exponential function with base  $e$ .

$$Z_k = \sum_{i=1}^m w_{ki} \exp(-\alpha_k y_i G_k(x_i)). \quad (8)$$

The above steps are repeated to calculate the value of  $f(x)$  by the  $f(x)$  addition model, and finally the final learner  $G(x)$  is obtained as

$$G(x) = \text{sign}[f(x)] = \text{sign} \left[ \sum_{k=1}^K \alpha_k G_k(x) \right], \quad (9)$$

where the sign function means to output 1 when  $f(x)$  is greater than or equal to 0 and  $-1$  when  $f(x)$  is less than 0.

**2.4. Bayesian Optimized Decision Tree.** Decision tree is an inductive statistical model training method based on the original data set. The tree is used to frame decision planning by training a tree classification model based on data attributes to estimate the relationship between independent and dependent variables, which has the advantages of extensive sample data handling, comprehensibility, coupling multiple features, and avoiding the influence of correlation between features. However, in some cases, traditional decision trees can lead to decreased fault diagnosis accuracy and increased leakage, for example traditional decision tree algorithms like CART and C4.5 are prone to overfitting [29]. To address these problems, Bayesian optimization is proposed. Bayesian optimization has been proved to be able to quickly and efficiently determine the best-fitting algorithm and its optimal hyperparameters to achieve the global optimum for many multimodal functions [30]. The main problem situations for which Bayesian optimization is oriented are

$$X^* = \text{argmax}_x f(x) (x \in S), \quad (10)$$

where  $S$  is the candidate set of the variable  $x$ , that is the set of possible values of the parameter  $x$ . Assuming that the function is sampled from a Gaussian process,  $x$  is first selected randomly from the set  $S$ . By using different hyperparameters for the learning algorithm experiments, the posterior distribution of values taken at any sample point is obtained for any sample point under the condition that the values taken at the previous  $n$  sampling points are known.

$$P(f|f_{1:n}) = \frac{P(f, f_{1:n})}{P(f_{1:n})}. \quad (11)$$

The posterior distribution is used to infer a currently optimal  $X^*$  as the configuration parameter for the next training verification attempt, but the optimization can only be performed for a deterministic function, so the posterior distribution needs to be transformed into a deterministic function  $\alpha_{EI}(x)$  before optimization, which is the acquisition function.

$$\alpha_{EI}(x) = E[[f(x) - f_{\max}]^+]. \quad (12)$$

The meaning of the acquisition function is the expectation of the excess of the value of point  $x$  under the posterior distribution with respect to the previously observed maximum value of the objective function. After defining this function, the automatic iterative calculation finds the point that makes the maximum value of the acquisition function  $\alpha_{EI}(x)$  in the above equation, which is the optimal point, that is,

$$X^* = \operatorname{argmax}_x \alpha_{EI}(x), \quad (13)$$

where  $X^*$  is the eventual optimal machine learning model configuration. For the Bayesian optimized decision tree with ensemble classifiers, the ensemble method, the number of learners, the learning rate, the maximum number of splits, and the number of predictor variables to be collected are the main parameters of Bayesian optimization, which will affect the classification accuracy of the decision tree and require a reasonable setting of the search range.

**2.5. Fault Diagnosis System Based on Decision Tree with Ensemble Method for AdaBoost and Bayesian Optimized Decision Tree with Ensemble Classifiers.** After preprocessing the collected bearing data, four models were established for experiments, in order to further compare and verify the effectiveness of the method. The specific flow block diagram is shown in Figure 2.

The first group of experiments (later referred to as model-1): The preprocessed motor bearing data were directly inputted into the model about decision tree with ensemble method for AdaBoost, and the model parameters were trained to obtain the results, and the model was used for fault classification.

The second group of experiments (later referred to as model-2): The motor bearing data preprocessed with time domain, frequency domain, and distance features are inputted into the model about decision tree with ensemble method for AdaBoost, and the model parameters are trained to obtain the results, and the model is used for fault classification.

The third group of experiments (later referred to as model-3): The motor bearing data preprocessed by the data are inputted to the model about Bayesian optimized decision tree with ensemble classifiers for training and fault classification with this model.

The fourth group of experiments (later referred to as model-4): The motor bearing data preprocessed with the data are inputted into the model about Bayesian optimized decision tree with ensemble classifiers for training in time domain, frequency domain, and distance features, and the model is used for fault classification.

### 3. Results

**3.1. Feature Extraction Results.** After the extraction of time domain, frequency domain, and distance features, the feature parameter matrix is obtained. At this time, the feature parameter matrix contains four kinds of data of motor bearings, which are the feature parameters of normal bearings, rolling element fault bearings, inner race fault bearings, and outer race fault bearings. Among them, the data labeled as 0 are the data of the normal bearing, the data labeled as 1 are the rolling element fault bearing, the data labeled as 2 are the inner race fault bearing, and the data labeled as 3 are the outer race fault bearing.

Figure 3 shows the partial signal data distribution of the original data, and Figure 4 shows the partial signal data distribution after feature extraction. It can be observed that most of the signals are not very different before feature extraction, while some of the signals have obvious differences for classification after feature extraction, and there are still some signals with little difference after feature extraction, which are not easy to distinguish.

**3.2. Decision Tree Diagnosis Results with Ensemble Method for AdaBoost in the Case That the Original Data Are Not Feature Extracted.** The experimental data in this section are the original bearing data without feature extraction. The decision tree with ensemble method for AdaBoost is trained using these data, setting the maximum number of splits to 20, the number of learners to 30, and the learning rate to 0.1. The final diagnosis accuracy of the decision tree with ensemble method for AdaBoost is 86.1%, the total misclassification cost is 333, the prediction speed is 7900 obs/sec, and the training time is 103.61 seconds.

The ROC curve is plotted as shown in Figure 5. The confusion matrix of the final result is shown in Figure 6. The figure shows that the model has a certain degree of misjudgment for normal bearings, rolling element faults, inner race faults, and outer race faults, and the degree of confusion for rolling element faults is greater, with a large proportion of rolling element faults being misjudged as normal bearings. There are 16 normal bearings judged as rolling element fault. There were 154 rolling element faults judged as normal bearings, 3 rolling element faults judged as inner race faults, and 32 rolling element faults judged as outer race faults. There were 1 inner race fault judged as normal bearing, 3 inner race faults judged as rolling element faults, and 47 inner race faults judged as outer race faults. There were 9 outer race faults judged as normal bearings, 28 outer race faults judged as rolling element faults, and 40 outer race faults judged as inner race faults.

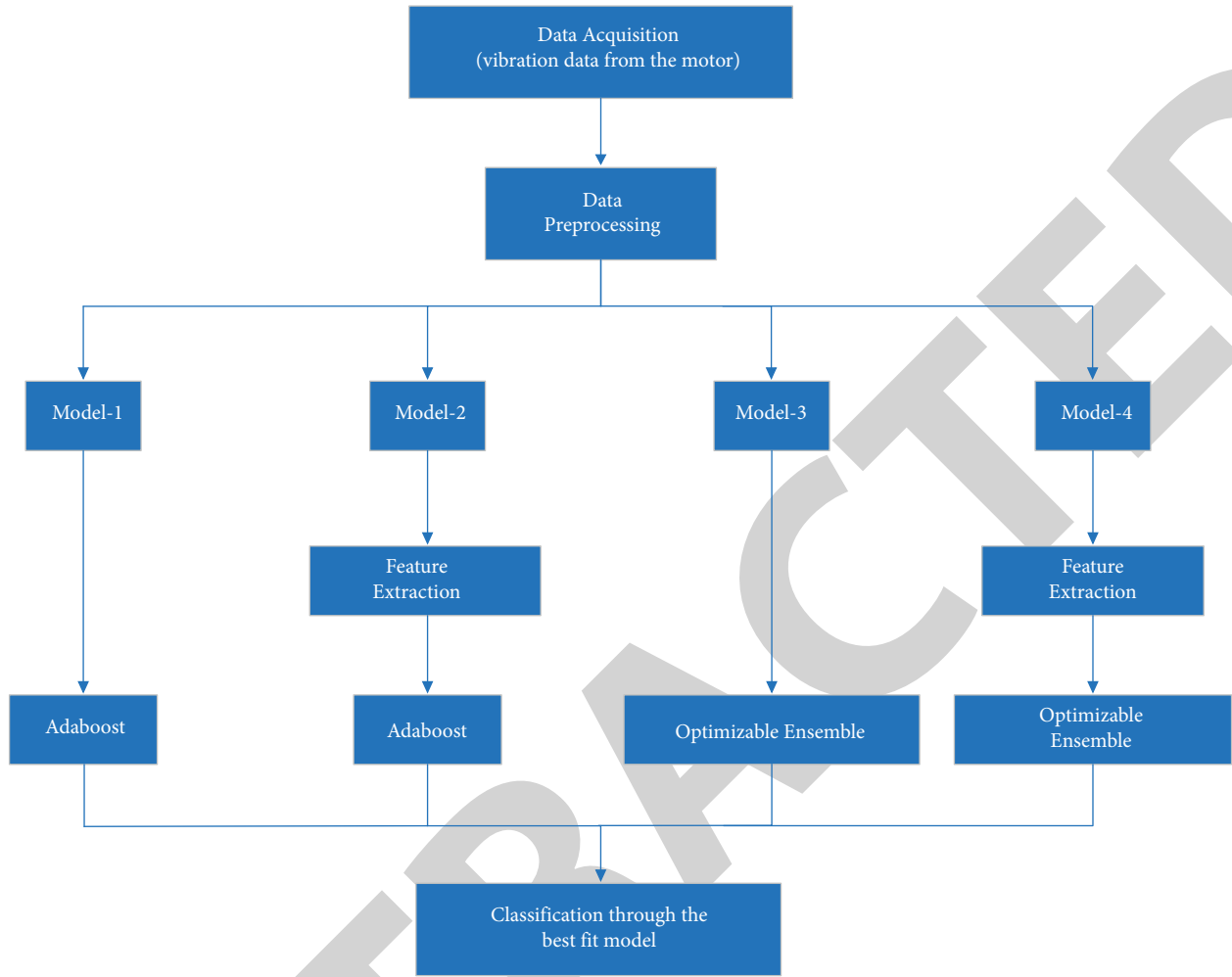


FIGURE 2: The flowchart of this article.

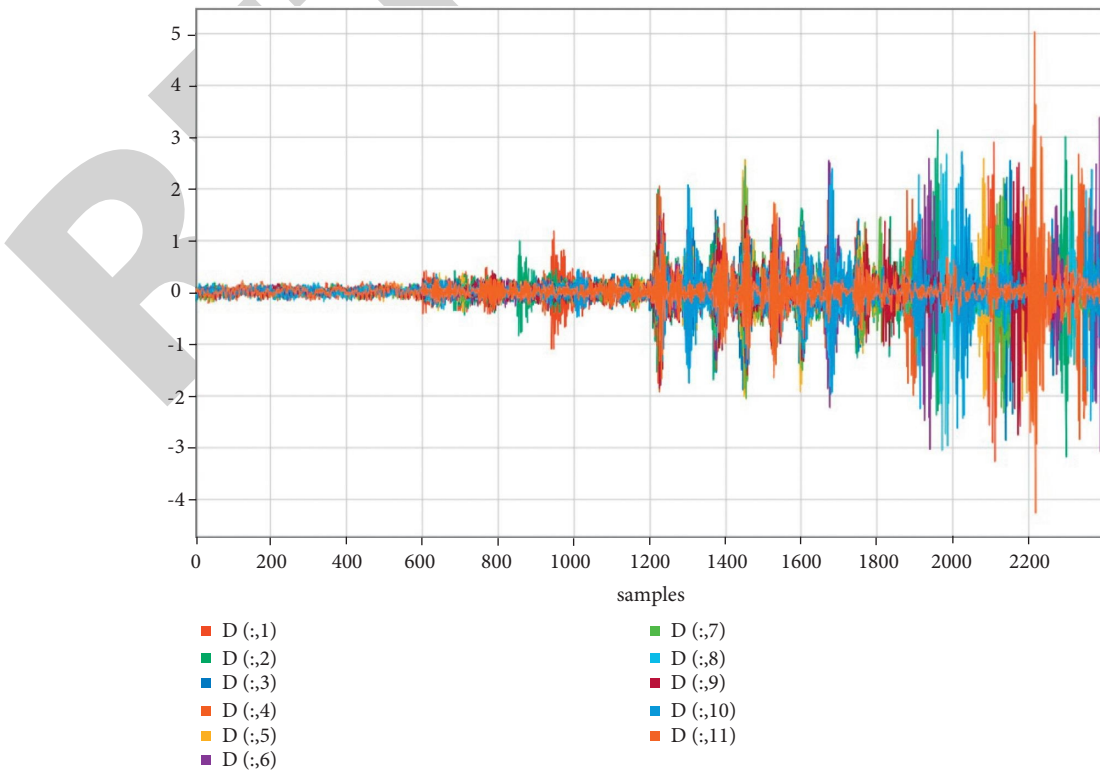


FIGURE 3: Distribution of partial signal data of the original data.

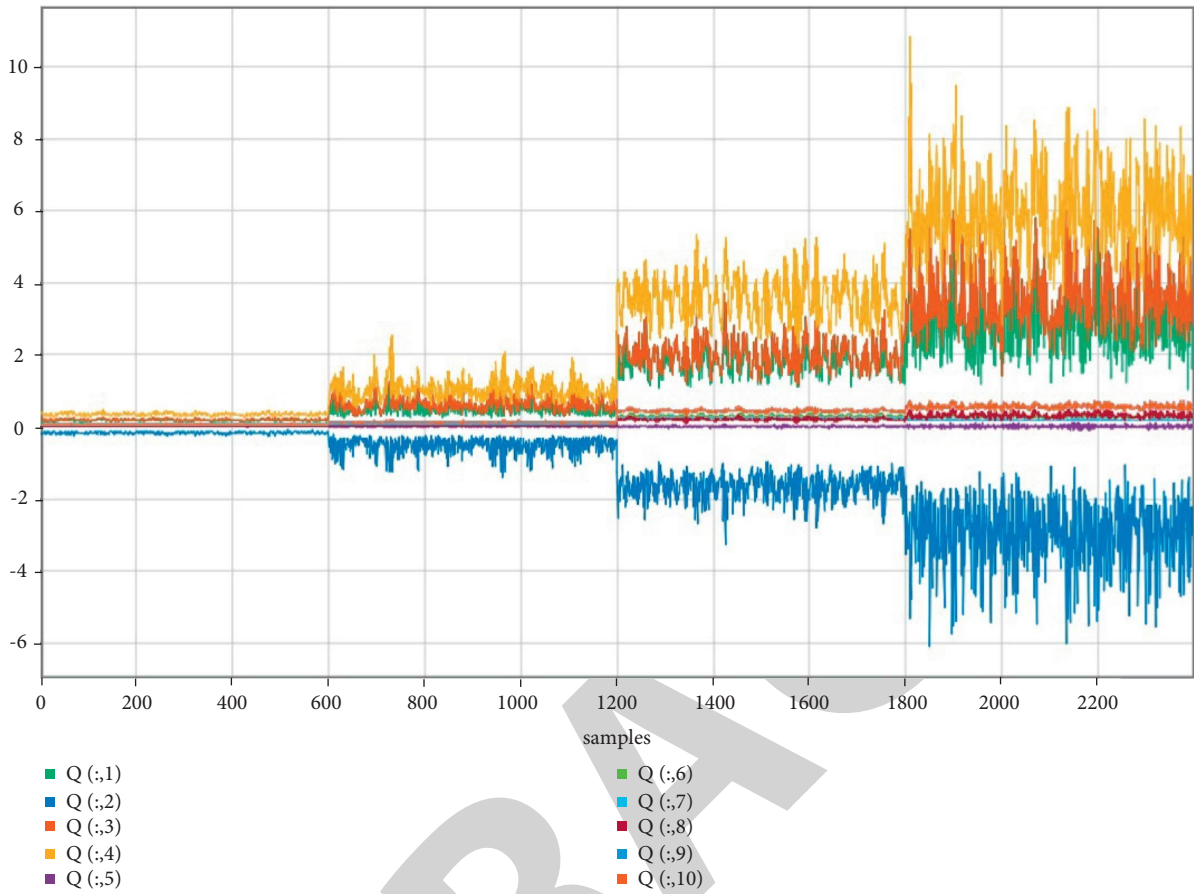


FIGURE 4: Distribution of partial signal data after feature extraction.

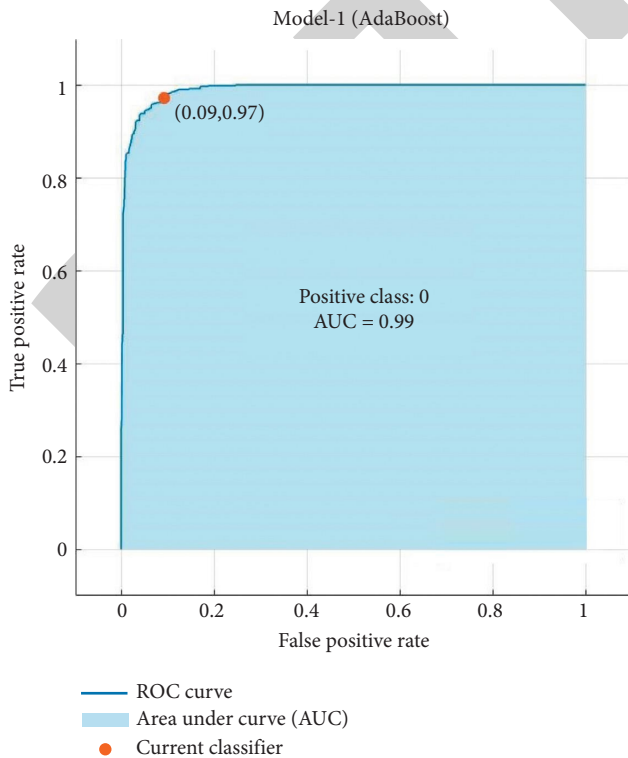


FIGURE 5: ROC curve of model-1.

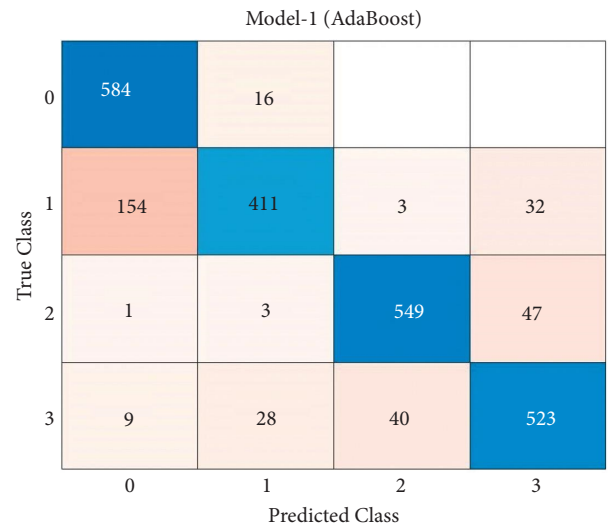


FIGURE 6: Confusion matrix of model-1.

3.3. Decision Tree Diagnosis Results with Ensemble Method for AdaBoost after the Feature Extraction of the Original Data. The experimental data in this section are the bearing data after feature extraction. The decision tree with ensemble method for AdaBoost is trained using these data, setting the maximum number of splits to 20, the number of learners to



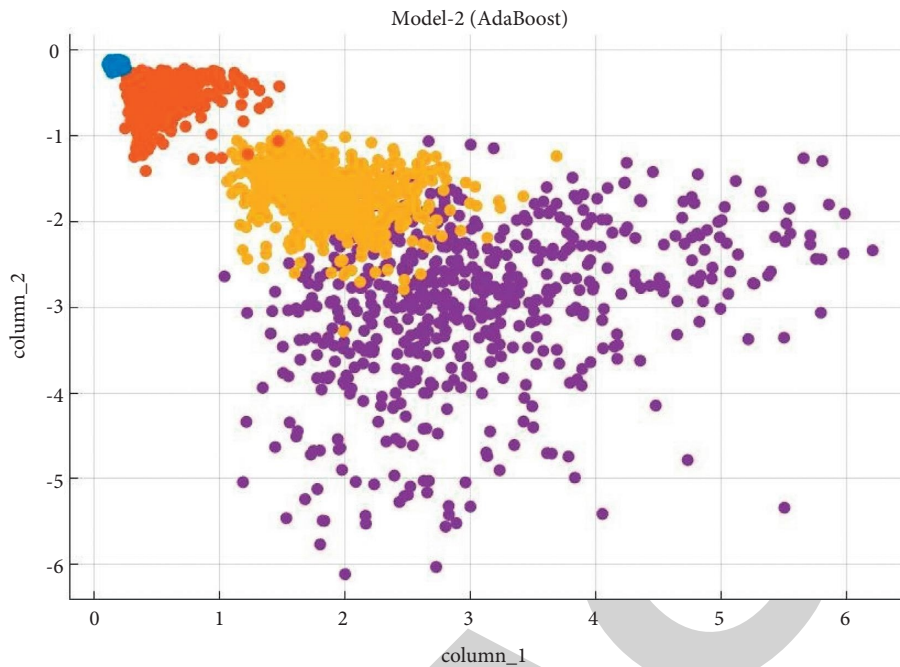


FIGURE 7: Model-2 scatterplot of classification before the experiment.

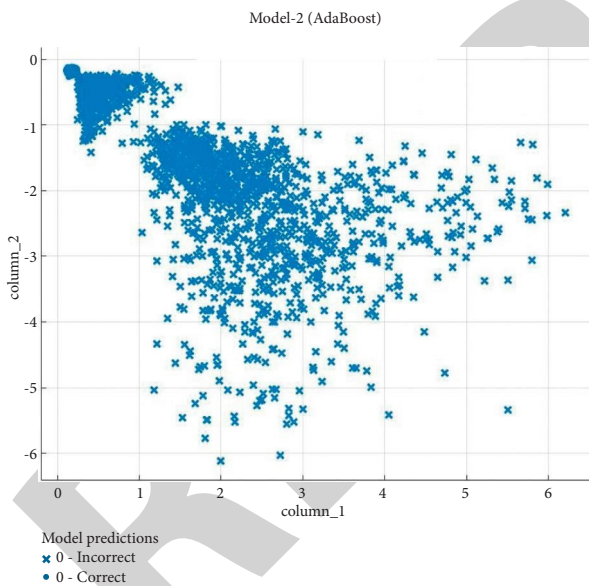


FIGURE 8: Model-2 scatterplot of classification after the experiment.

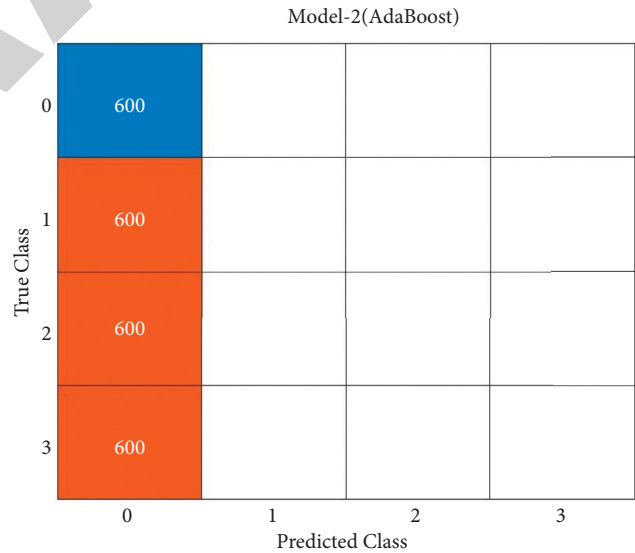


FIGURE 9: Confusion matrix of model-2.

30, and the learning rate to 0.1. The final diagnosis accuracy of the decision tree with integrated AdaBoost is 25.0%, the total misclassification cost is 1800, the prediction speed is 35000 obs/sec, and the training time is 7.4335 seconds.

Figure 7 shows the scatterplot of model-2 before the experiment, and Figure 8 shows the scatterplot of model-2 after the experiment. The confusion matrix of the final results is shown in Figure 9. The figure shows that the accuracy of the model is very low, and the misjudgment rate for rolling element fault, inner ring fault, and outer ring fault is

100%. Rolling element fault, inner race fault, and outer race fault were all judged as normal bearings.

3.4. Bayesian Optimized Decision Tree Diagnosis Results with Ensemble Classifiers in the Case That the Original Data Are Not Feature Extracted. The experimental data in this section are the original bearing data without feature extraction. The Bayesian optimized decision tree with ensemble classifiers is

TABLE 2: Model-3 hyperparametric search range.

Ensemble method	Number of learners	Learning rate	Maximum number of splits	Number of predictor variables to be collected
Bag, AdaBoost, RUSBoost	10–500	0.001–1	0–2399	1–200

TABLE 3: Model-3 optimized hyperparameters.

Ensemble method	Number of learners	Maximum number of splits	Number of predictor variables to be captured
Bag	80	583	193

trained using these data, and the hyperparameter search range is set as shown in Table 2.

The final Bayesian optimized decision tree with ensemble classifiers had a diagnostic accuracy of 99.3%, a total misclassification cost of 18, a prediction speed of ~7100 obs/sec, and a training time of 898.04 seconds.

The minimum classification error iteration diagram is shown in Figure 10. As can be seen in Figure 10, the Bayesian optimized decision tree with ensemble classifiers converges quickly, stabilizes in the late iterations, and finds the optimal hyperparameters in the 28th iteration. The results of the optimized hyperparameters are shown in Table 3.

The confusion matrix of the final results is shown in Figure 11. The figure shows that the performance of the model is good and the misclassification rate is not high. There are 6 normal bearings judged as rolling element faults, 1 rolling element fault judged as normal bearing, 4 inner race faults judged as outer race faults, 4 outer race faults judged as rolling element faults, and 3 outer race faults judged as inner race faults.

*3.5. Bayesian Optimized Decision Tree Diagnosis Results with Ensemble Classifiers after the Feature Extraction of the Original Data.* The experimental data in this section are the bearing data after feature extraction. The Bayesian optimized decision tree with ensemble classifiers is trained using these data, and the hyperparameter search range is set as shown in Table 4.

The final Bayesian optimized decision tree with ensemble classifiers had a diagnostic accuracy of 99.9%, a total misclassification cost of 2, a prediction speed of 15000 obs/sec, and a training time of 138.4 seconds.

The minimum classification error iteration diagram is shown in Figure 12. Similar to model-3, the optimized hyperparameter results are shown in Table 5.

The confusion matrix of the final results is shown in Figure 13. The figure shows that the performance of the model is good and the misclassification rate is not high. There is 1 inner race fault judged as outer race fault and 1 outer race fault judged as inner race fault.

#### 4. Discussion

In order to have a more intuitive understanding of the obtained experimental results, four experimental data are compared.

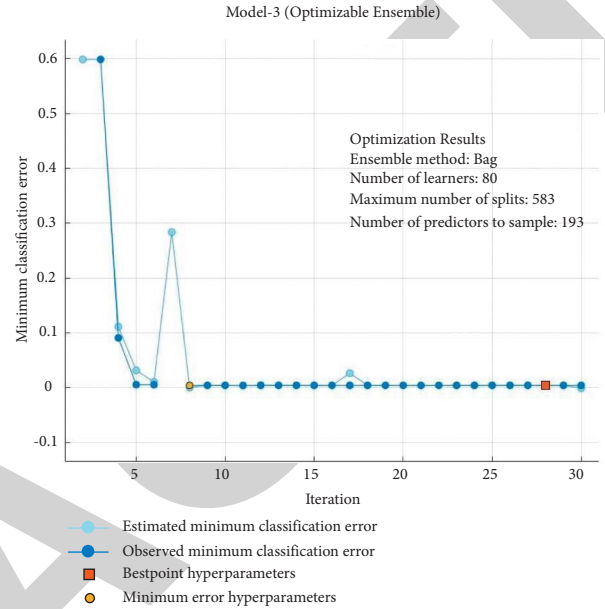


FIGURE 10: Model-3 minimum classification error iteration diagram.

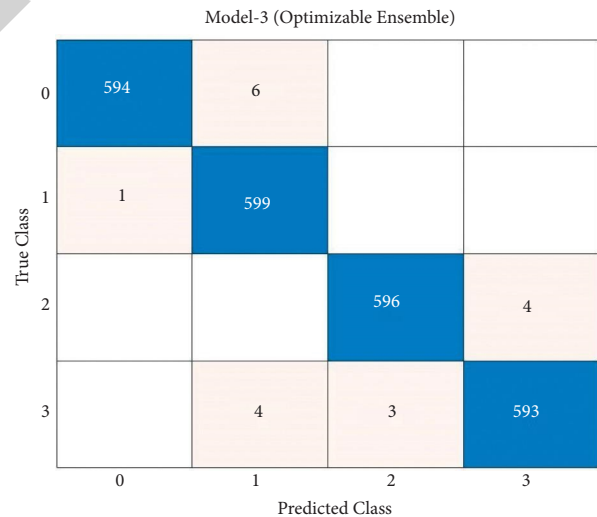


FIGURE 11: Confusion matrix of model-3.

The comparison between the first set of experimental data and the second set of experimental data revealed that the model’s prediction speed became faster after feature extraction, but the accuracy of the model showed a precipitous decrease, from 86.1% in model-1 to 25.0%. The experimental results obtained were all for normal bearings, and no diagnosis was made for rolling element fault, inner race fault, or outer race fault. For motor bearing fault diagnosis, feature extraction has been shown to be effective in

TABLE 4: Model-4 hyperparametric search range.

Ensemble method	Number of learners	Learning rate	Maximum number of splits	Number of predictor variables to be collected
Bag, AdaBoost, RUSBoost	10–500	0.001–1	0–2399	1–25

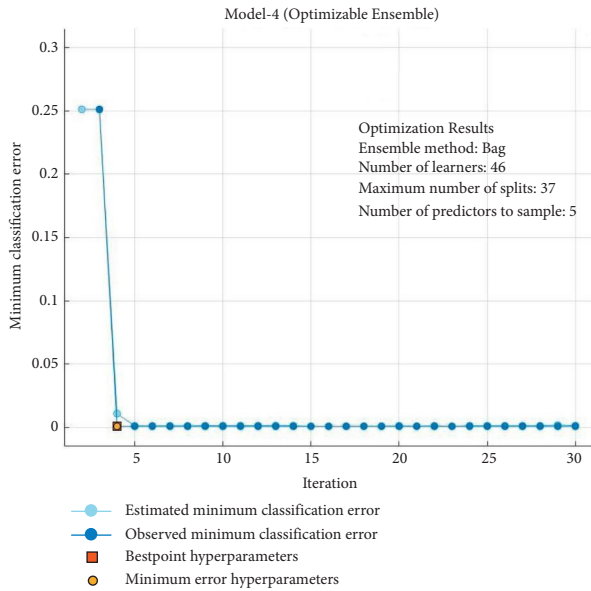


FIGURE 12: Model-4 minimum classification error iteration diagram.

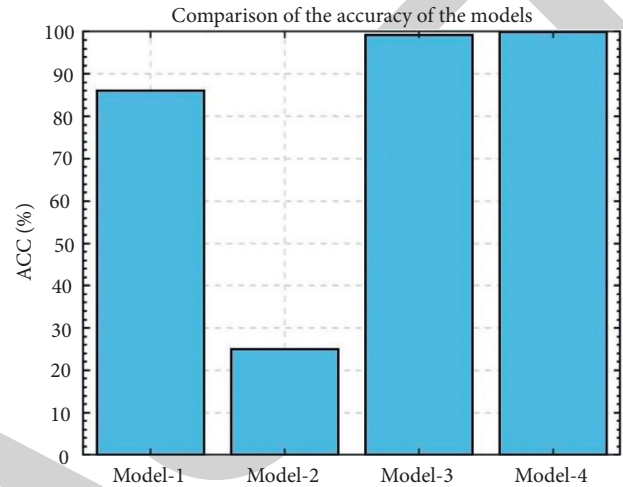


FIGURE 14: Comparison of the accuracy of the models.

TABLE 5: Model-4 optimized hyperparameters.

Ensemble method	Number of learners	Maximum number of splits	Number of predictor variables to be captured
Bag	46	37	5

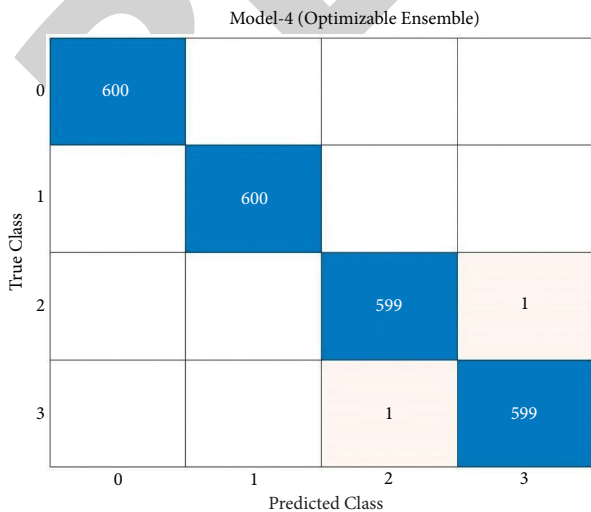


FIGURE 13: Confusion matrix of model-4.

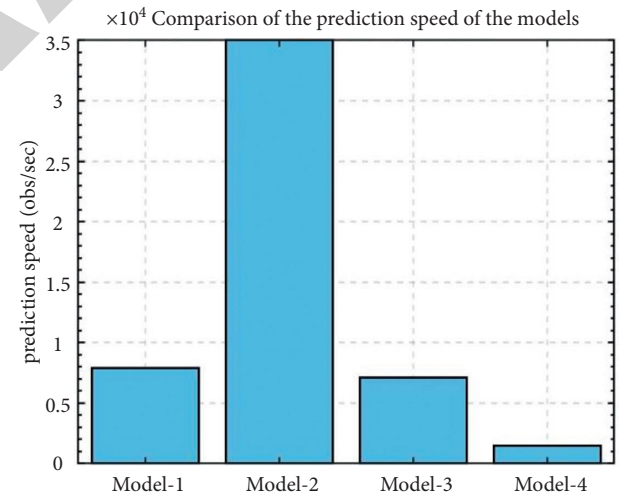


FIGURE 15: Comparison of the prediction speed of the models.

most models to improve the accuracy [31]. However, the decision tree with ensemble method for AdaBoost is not suitable for feature extraction first.

A comparison between the first set of experimental data and the third set of experimental data revealed that the accuracy was effectively improved from 86.1% to 99.3% with a smaller reduction in prediction speed after using the Bayesian optimized decision tree with ensemble classifiers. In the experimental results obtained from model-3, the ensemble method for the optimized hyperparameter

selection is the Bag algorithm. This shows that the Bayesian optimized decision tree with ensemble classifiers has a better performance compared to the decision tree with ensemble method for AdaBoost.

The comparison between the third set of experimental data and the fourth set of experimental data reveals that after feature extraction of the original data, there is a certain degree of reduction in the prediction speed and the accuracy rate increases from 99.3% to 99.9%. In the experimental results obtained from model-4, there was no misclassification of normal bearing and rolling element faults, one inner race fault was misclassified as an outer race fault, and one outer race fault was misclassified as an inner race fault. Therefore, for the Bayesian optimized decision tree with ensemble classifiers, feature extraction before the experiment can improve the accuracy of the model to some extent.

The comparison between the second set of experimental data and the fourth set of experimental data reveals that after using the Bayesian optimized decision tree with ensemble classifiers, the prediction speed is reduced to a greater extent, but the accuracy rate is effectively increased from 25.0% to 99.9%. In the experimental results obtained in model 4, the ensemble method for the optimized hyperparameter selection is the Bag algorithm. Therefore, for the bearing data for which feature extraction has been performed, the accuracy of the Bayesian optimized decision tree with ensemble classifiers is significantly better than that of the decision tree with ensemble method for AdaBoost.

The comparison of the accuracy and prediction speed of the four models can be visually expressed in Figures 14 and 15. In Figure 14, the horizontal axis indicates the classification of the model and the vertical axis indicates the accuracy of the model. The accuracy of both model-3 and model-4 reached over 99%, with model-4 having the highest accuracy and model-2 having the lowest accuracy. In Figure 15, the horizontal axis indicates the classification of the model and the vertical axis indicates the prediction speed of the model. Model-2 has the fastest prediction speed and model-4 has the slowest prediction speed.

## 5. Conclusions

Based on the experimental data from Case Western Reserve University, the performance of the AdaBoost algorithm and the Bayesian optimized decision tree with ensemble classifiers in motor bearing fault diagnosis is studied in depth in the time domain, frequency domain, and distance feature calculation methods. The ideal experimental results were obtained after the experiments. The experimental conclusions are as follows:

- (1) It is stated in many literature works that feature extraction combined with many fault diagnosis methods can get better accuracy. For the decision tree with ensemble method for AdaBoost, the extraction of time domain, frequency domain, and distance features from the original data can have an obvious negative impact on the diagnosis results.
- (2) The Bayesian optimized decision tree with ensemble classifiers can learn the correlation in the data more

accurately than the decision tree with ensemble method for AdaBoost and construct the fitting conditions for accurate diagnosis.

- (3) Regardless of whether feature extraction is performed, the experimental results of the Bayesian optimized decision tree with ensemble classifiers, the integration method for the optimal hyperparameter selection is Bag algorithm instead of AdaBoost algorithm and RUSBoost algorithm.
- (4) The decision tree with ensemble method for AdaBoost on the original data predicts faster than the machine learning method and, at the same time, has a certain accuracy (86.1%).
- (5) The Bayesian optimized decision tree with ensemble classifiers after feature extraction on the original data has the best performance and better accuracy than other combined methods.

In this article, we focus on the principles of the decision tree with ensemble method for AdaBoost and the Bayesian optimized decision tree with ensemble classifiers. Based on this theory, we derived the experimental results. The optimal method was the Bayesian optimized decision tree with ensemble classifiers, after feature extraction of the original data. The accuracy of this method is up to 99.9%. At the same time, unlike previous studies, we found that feature extraction does not improve the accuracy of diagnosis for the decision tree with ensemble method for AdaBoost and there is a precipitous decline.

Although we have made some achievements with this study, there are still some limitations. The data we used were too homogeneous and the characteristics of the collected data were limited. The quality of the data may affect the performance of the model in our experiments. In future work, more detailed data processing can be used to extract features that work better for the experiments. In addition, future research is not only applicable to bearing data, but also can be extended to other rotating machinery fault diagnosis, such as gearboxes, pumps, etc.

In the industrial Internet of Things, it is believed that the findings of this experiment can provide a certain degree of theoretical support for future research on fault diagnosis of motor equipment and rolling bearings and provide a reference value for the development of future research on motor bearing fault diagnosis.

## Data Availability

The data used to support the findings of this study are included within the article.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## References

- [1] D. J. Ning, J. Yu, and J. Huang, "An intelligent device fault diagnosis method in industrial internet of things," in *Proceedings of the 2018 international symposium in sensing and*

- instrumentation in IoT era (ISSI), pp. 1–6, IEEE, Shanghai, China, September 2018.
- [2] L. Cheng, T. Wang, H. Xiao, Z. Wang, J. Wang, and J. Liu, “A study on the architecture of manufacturing internet of things,” *International Journal of Modelling, Identification and Control*, vol. 23, no. 1, pp. 8–23, 2015.
  - [3] Q. Gao, C. Duan, H. Fan, and Q. Meng, “Rotating machine fault diagnosis using empirical mode decomposition,” *Mechanical Systems and Signal Processing*, vol. 22, no. 5, pp. 1072–1081, 2008.
  - [4] B. Li, M. Y. Chow, Y. Tipsuwan, and J. Hung, “Neural-network-based motor rolling bearing fault diagnosis,” *IEEE Transactions on Industrial Electronics*, vol. 47, no. 5, pp. 1060–1069, 2000.
  - [5] Z. K. Peng, P. W. Tse, and F. L. Chu, “A comparison study of improved Hilbert-Huang transform and wavelet transform: application to fault diagnosis for rolling bearing,” *Mechanical Systems and Signal Processing*, vol. 19, no. 5, pp. 974–988, 2005.
  - [6] C. Chen, F. Shen, and R. Q. Yan, “Enhanced least squares support vector machine-based transfer learning strategy for bearing fault diagnosis,” *Chinese Journal of Scientific Instrument*, vol. 38, no. 1, pp. 33–40, 2017.
  - [7] J. Yang and M. Zhao, “Fault diagnosis of traction motor bearings using modified bispectrum and empirical mode decomposition,” *Proceedings of the CSEE*, vol. 32, no. 18, pp. 116–122, 2012.
  - [8] S. Xiangjin, W. Zhuo, and H. Jingtao, “Diagnosis of bearing fault in induction motors using Hilbert demodulation approach,” *Transactions of China Electrotechnical Society*, vol. 33, no. 21, pp. 4941–4948, 2018.
  - [9] J. A. Lucena-Junior, T. L. de Vasconcelos Lima, G. P. Bruno et al., “Chaos theory using density of maxima applied to the diagnosis of three-phase induction motor bearings failure by sound analysis,” *Computers in Industry*, vol. 123, Article ID 103304, 2020.
  - [10] S. J. Kim, K. Kim, T. Hwang et al., “Motor-current-based electromagnetic interference de-noising method for rolling element bearing diagnosis using acoustic emission sensors,” vol. 193, Measurement, Article ID 110912, 2022.
  - [11] T. Hiruta, K. Maki, T. Kato, and Y. Umeda, “Unsupervised learning based diagnosis model for anomaly detection of motor bearing with current data,” *Procedia CIRP*, vol. 98, pp. 336–341, 2021.
  - [12] Z. Wang, J. Yang, H. Li, D. Zhen, F. Gu, and A. Ball, “Improved cyclostationary analysis method based on TKEO and its application on the faults diagnosis of induction motors,” *ISA transactions*, 2021.
  - [13] D. Zhang, Z. Q. Ning, B. Yang, T. Wang, and Y. Ma, “Fault diagnosis of permanent magnet motor based on DCGAN-RCCNN,” *Energy Reports*, vol. 8, pp. 616–626, 2022.
  - [14] M. Nikfar, J. Bitencourt, and K. Mykoniatis, “A two-phase machine learning approach for predictive maintenance of low voltage industrial motors,” *Procedia Computer Science*, vol. 200, pp. 111–120, 2022.
  - [15] H. Wang, Z. Liu, D. Peng, and Z. Cheng, “Attention-guided joint learning CNN with noise robustness for bearing fault diagnosis and vibration signal denoising,” *ISA transactions*, 2021.
  - [16] L. Zhi, J. Yuanliang, and X. Zhongheng, “Anti-noise motor fault diagnosis method based on decision tree and the feature extraction methods in the time domain and frequency domain,” in *Proceedings of the 2021 International Conference on Communications, Information System and Computer Engineering (CISCE)*, pp. 71–75, IEEE, Beijing, China, May 2021.
  - [17] J. Wang, Y. Liang, Y. Zheng, R. X. Gao, and F. Zhang, “An integrated fault diagnosis and prognosis approach for predictive maintenance of wind turbine bearing with limited samples,” *Renewable Energy*, vol. 145, pp. 642–650, 2020.
  - [18] Z. Kong, B. Tang, L. Deng, W. Liu, and Y. Han, “Condition monitoring of wind turbines based on spatio-temporal fusion of SCADA data by convolutional neural networks and gated recurrent units,” *Renewable Energy*, vol. 146, pp. 760–768, 2020.
  - [19] Y. Chang, J. Chen, C. Qu, and T. Pan, “Intelligent fault diagnosis of wind turbines via a deep learning network using parallel convolution layers with multi-scale kernels,” *Renewable Energy*, vol. 153, pp. 205–213, 2020.
  - [20] D. Yang, H. Li, Y. Hu, J. Zhao, H. Xiao, and Y. Lan, “Vibration condition monitoring system for wind turbine bearings based on noise suppression with multi-point data fusion,” *Renewable Energy*, vol. 92, pp. 104–116, 2016.
  - [21] A. Bhatnagar, S. Shukla, and N. Majumdar, “Machine learning techniques to reduce error in the internet of things,” in *Proceedings of the 2019 9th International Conference on Cloud Computing, Data Science & Engineering (Confluence)*, pp. 403–408, IEEE, Noida, India, January 2019.
  - [22] A. Mavromatis, C. Colman-Meixner, A. P. Silva, X. Vasilakos, R. Nejabati, and D. Simeonidou, “A software-defined IoT device management framework for edge and cloud computing,” *IEEE Internet of Things Journal*, vol. 7, no. 3, pp. 1718–1735, 2020.
  - [23] R. Yan, H. Sun, and Y. Qian, “Energy-aware sensor node design with its application in wireless sensor networks,” *IEEE Transactions on Instrumentation and Measurement*, vol. 62, no. 5, pp. 1183–1191, 2013.
  - [24] H. Tang, S. Lu, G. Qian, J. Ding, Y. Liu, and Q. Wang, “IoT-based signal enhancement and compression method for efficient motor bearing fault diagnosis,” *IEEE Sensors Journal*, vol. 21, no. 2, pp. 1820–1828, 2021.
  - [25] X. Wang, X. Yan, and Y. He, “Weak fault feature extraction and enhancement of wind turbine bearing based on OCYCBD and SVDD,” *Applied Sciences*, vol. 9, no. 18, p. 3706, 2019.
  - [26] Y. Song, Y. Li, L. Jia, and M. Qiu, “Retraining strategy-based domain adaption network for intelligent fault diagnosis,” *IEEE Transactions on Industrial Informatics*, vol. 16, no. 9, pp. 6163–6171, 2020.
  - [27] Y. Lei, Z. He, Y. Zi, and Q. Hu, “Fault diagnosis of rotating machinery based on multiple ANFIS combination with GAs,” *Mechanical Systems and Signal Processing*, vol. 21, no. 5, pp. 2280–2294, 2007.
  - [28] X. H. Wu, S. J. Niu, C. O. Wu, and W. Qin, “An improvement on estimating covariance matrix during cluster analysis using mahalanobis distance,” *Journal of Applied Sport Management*, vol. 30, no. 2, pp. 240–245, 2011.
  - [29] W. Sun, J. Chen, and J. Li, “Decision tree and PCA-based fault diagnosis of rotating machinery,” *Mechanical Systems and Signal Processing*, vol. 21, no. 3, pp. 1300–1317, 2007.
  - [30] A. D. Bull, “Convergence rates of efficient global optimization algorithms,” *Journal of Machine Learning Research*, vol. 12, no. 10, 2011.
  - [31] R. Liu, B. Yang, E. Zio, and X. Chen, “Artificial intelligence for fault diagnosis of rotating machinery: a review,” *Mechanical Systems and Signal Processing*, vol. 108, pp. 33–47, 2018.