

Motor Bearing Fault Diagnosis based on Multi-domain Feature Extraction and Improved Optimized Decision Tree

Jinshan Lin

Baoding Department of Electrical Engineering, North China Electric Power University, Baoding, Hebei, 071003, China

Keywords: Fault diagnosis Multi-domain feature extraction Decision tree Bayesian optimization CART.

Abstract: In view of the large noise of the collected signals in the current fault diagnosis methods of motor bearings, the incomplete feature extraction, and the low efficiency and accuracy of fault diagnosis, this paper proposes a method based on multi-domain signal feature extraction and Bayesian Optimized decision tree fault diagnosis method for motor bearings. The method mainly includes signal feature extraction and optimized decision tree fault diagnosis. First, the bearing vibration signal data set is used to extract feature values in the time domain, frequency domain, and time-frequency domain respectively; then Bayesian optimization is used to supervise the decision tree. The parameters are optimized, and the training data is used to train and predict the optimized decision tree classifier. At the same time, the same data set is used to train the decision tree model without feature value extraction, the optimized decision tree model without feature value extraction, and the decision tree model with feature value extraction. The diagnosis results of the four models are compared, and the results show that the method can effectively reduce the dimensionality of the signal data, and can greatly improve the fault diagnosis rate and the fault diagnosis accuracy rate.

1. Introduction

When the motor is running, due to the impact of the environment and the voltage quality, it is often accompanied by an impact signal, resulting in the majority of motor faults in bearing faults. The frequent occurrence of motor failures seriously affects the safety of national industrial production. Therefore, it is very necessary to diagnose the fault of the motor bearing, and how to improve the accuracy of the fault diagnosis is a problem that needs to be solved urgently. Motor bearing fault diagnosis mainly includes signal feature vector extraction and fault diagnosis. Because motors often work under conditions of time-varying speeds, how to achieve effective signal feature extraction is an important basis for bearing fault diagnosis.

The actual operating state of the motor is complex, and the collected bearing vibration signals often have nonlinear characteristics. At present, for data processing, time-frequency methods are generally used to extract features of signals, such as Fourier transform [4]. However, Fourier transform is difficult to extract effective eigenvalues of non-stationary signals; wavelet transform is not sensitive to the high frequency part of the signal, and it is easy to cause the missing of the extracted signal eigenvalues. In terms of fault type recognition, deep learning is widely used in various fault diagnosis due to its powerful feature extraction and learning capabilities. BP neural network is widely used because of its high-dimensional function nonlinear mapping ability and complex pattern classification ability. The wavelet packet decomposition decomposes the motor bearing vibration signal in the full frequency band, which is beneficial to highlight the fault impact signal and improve the accuracy of bearing fault diagnosis.

Therefore, this article will perform full-angle and high-precision extraction of data signal features in multiple domains such as time domain, frequency domain, and time-frequency domain, and use Bayesian optimized decision trees for motor bearing fault diagnosis.

2. Feature parameter selection

2.1. Time domain statistical characteristics

Signal time-domain statistical analysis is to extract relevant features of the signal in the time domain. Time-domain features can be divided into two types: dimensional eigenvalues and non-dimensional eigenvalues. Dimensional eigenvalues are related to the operating state of the equipment, the result of feature extraction is often noisy, the effect of feature extraction is not ideal, and it is rarely used in engineering practice. Compared with the dimensional characteristic index, the dimensionless index has the following three characteristics: (1) In the face of noise interference, the dimensionless eigenvalue has a good identification ability, and its value is basically unchanged. (2) The impact failure occurs with a high probability during the operation of the bearing. The kurtosis index, margin index, and impulse index have good characteristics for it. (3) The root mean square value is very stable.

This paper extracts 6 dimensional eigenvalues including peak value and 5 dimensionless eigenvalues including waveform index from vibration signal.

Table 1 Number and calculation formula of dimensional time-domain indicators

Numbering	Characteristic index	formula
T ₁	Peak	$x_{max} = \max x(t) $
T ₂	Average amplitude	$\bar{ x } = \frac{1}{N} \sum_{i=1}^N x_i $
T ₃	Root mean square	$x_{rms} = \sqrt{\frac{1}{N} \sum_{i=1}^N x_i^2}$

Table 2 Dimensionless time domain index number and calculation formula

Numbering	Characteristic index	formula
T ₇	Margin Index	$L = \frac{x_p}{x_r}$
T ₈	Impulse indicator	$I = \frac{x_p}{x}$
T ₉	Kurtosis index	$K = \frac{a}{x_{rms}^4}$

Note: where x represents vibration signal

2.2. Frequency domain statistical characteristics

In general, the processing of data in the time domain is simpler and the results are more intuitive. The representation of data in the frequency domain is more concise, and observing the signal in the frequency domain also makes the analysis of the problem more in-depth, and the data feature extraction is more comprehensive. In this paper, Fourier transform (FFT) is used to display the time domain signal in the frequency domain, and to perform correlation analysis and feature extraction on the frequency spectrum. This paper selects five important frequency domain characteristics of the center of gravity frequency, and the calculation formula is shown in Table 3.

Table 3 Frequency domain index number and calculation formula

Numbering	Characteristic index	formula
T ₁₂	Center of gravity frequency	$FC = \frac{\int_0^{+\infty} fS(f)df}{\int_0^{+\infty} S(f)df}$
T ₁₃	Rms frequency	$RMSF = \sqrt{\frac{\int_0^{+\infty} f^2 S(f)df}{\int_0^{+\infty} S(f)df}}$
T ₁₄	Frequency variance	$VF = \frac{\int_0^{+\infty} (f - FC)^2 S(f)df}{\int_0^{+\infty} S(f)df}$

2.3. Time-frequency domain statistical characteristics

Bearing fault vibration signals are often mixed with impact signals and noise, showing non-stationary characteristics. Therefore, on the basis of time domain and frequency domain feature selection, this paper combines time domain information and real-time frequency domain analysis on the basis of frequency domain data processing and analysis.

2.3.1 Wavelet packet decomposition

Wavelet packet decomposition is a time-frequency analysis theory developed in recent years. In the signal Fourier transform analysis, the signal data is decomposed in the frequency domain and cannot reflect the time-frequency characteristics of the signal, which may cause the loss of the signal characteristics. Compared with wavelet transform, wavelet packet decomposition has higher sensitivity to high-frequency signals and higher detection accuracy. Because the change of the bearing state is very complicated, it is difficult to accurately diagnose the motor bearing state only from the low-frequency part or the high-frequency part of the signal. Therefore, the wavelet packet decomposition with the characteristics of full frequency domain and high precision has a broad application prospect in the fault diagnosis of motor bearings.

Perform wavelet packet decomposition on the faulty bearing signal to obtain multiple wavelet packets, and calculate the energy of each wavelet packet separately. The formula for wavelet packet decomposition is:

$$\begin{cases} w_{2n}(t) = \sqrt{2} \sum_k h(k) w_n(2t - k) \\ w_{2n+1}(t) = \sqrt{2} \sum_k g(k) w_n(2t - k) \end{cases}$$

Where $g(k)$ is a low-pass filter, $h(k)$ is a high-pass filter

The formula for calculating the energy of each node of the wavelet packet is

$$E_m = \sum_{k=1}^N (d_m(k))^2$$

By analyzing and processing the bearing data, two time-frequency domain eigenvalues can be obtained, numbered T17~T18.

2.3.2 Entropy feature

Entropy is a physical quantity used to measure the distribution of energy in space. The larger the entropy value, the more chaotic the signal distribution, the smaller the entropy value, and the more uniform the signal distribution. The information entropy in the time domain and frequency domain

are used as signal time-frequency domain features to extract T19 and T20 respectively. The information entropy calculation formula is:

$$H(x) = \sum_{i=1}^N P(x_i) \frac{1}{\log(P(x_i))}$$

Combine wavelet packet decomposition, calculate wavelet packet energy entropy, and use it as feature T21:

$$H(m) = \sum_{m=1}^2 P_m \frac{1}{\log(P_m)}$$

This paper analyzes the bearing fault data mainly in the time domain, frequency domain, and time-frequency domain, and extracts the relevant 21 fault data features.

3. Decision tree and CART algorithm

Decision tree is a theory that is widely used in data processing and analysis. There are many theoretical methods for decision tree generation. This paper adopts the CART algorithm to establish a motor bearing fault diagnosis decision tree model.

3.1. Decision tree theory

Since the decision tree theory was proposed in 1960, decision trees have been quickly applied to various fields. Before starting to create a decision tree, you first need to divide the data set into a training set and a detection set. Then use the training set to build a decision tree, pruning the decision tree according to the established branch threshold, and then use the detection set to adjust and optimize the parameters of the decision tree model to establish the final decision tree model.

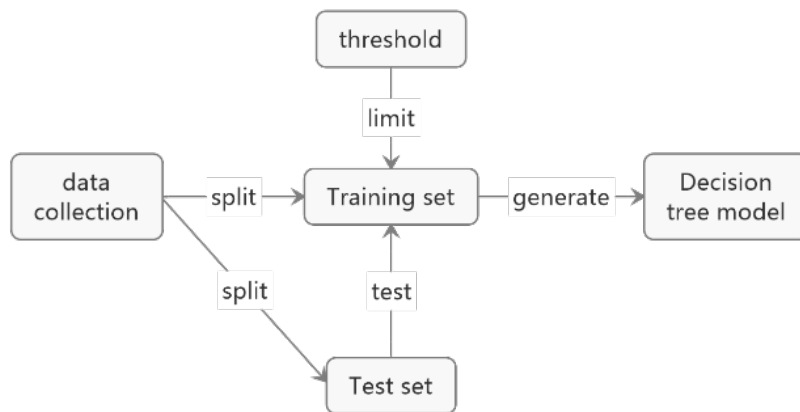


Fig.1: Decision tree model creation process

3.2. CART (Classification and Regression Trees) algorithm

The CART algorithm uses a recursive dichotomous feature space, thereby dividing the feature space into a limited number of different units, so as to obtain the predicted probability distribution on different units; the CART algorithm uses the coefficients of each influencing factor to select features.

And use it to build a decision tree, and calculate the factor with the smallest index as the best split point. Assuming that there are N classification categories in the existing data set M, the number of classification categories is, then the index of the data set M is calculated as:

$$Gini(M) = 1 - \sum_{i=1}^N \frac{|m_i|}{|M|}$$

If according to the attribute A_i in feature A, according to the CART dichotomy theory, the sample set is divided into two categories: A1 and A2, then the index of the sample set under the condition of feature A is:

$$Gini(M|A) = \frac{|M_1|}{|M|} Gini(M_1) + \frac{|M_2|}{|M|} Gini(M_2)$$

4. Bayesian optimization

4.1. Basic principles of Bayesian optimization

Parameters are an important factor in determining whether the diagnostic model is accurate [18]. The parameters of the model are obtained through the training data set, while the hyperparameters are manually specified before training. In today's big data era, once the hyperparameters are determined, long training is required to obtain the final model and evaluate the effect of the model. The determination of hyperparameters has a great influence on the effect of the model, so how to quickly obtain the optimal hyperparameters has become an urgent problem to be solved.

At present, most models use grid search or random search to find the optimal hyperparameters, but the results are not very satisfactory. Random search often fails to find optimal hyperparameters, and grid search takes too long. The advantage of Bayesian optimization is that the optimal parameter set of the model can be found in a short time.

Assuming a set of hyperparameter combinations, different results will be obtained by selecting different hyperparameters. Bayesian optimization assumes that there is a functional relationship between the hyperparameters and the final loss function we need to optimize.

4.2. Acquisition function

The collection function is an important basis for how the model finds the next optimal solution from the parameter set. Currently, there are mainly three collection functions: PI (probability of improvement), EI (Expected improvement), and UCB (Upper confidence bound).

This article mainly adopts PI.

First assume that $f' = \min f$ and f' represent the minimum value of f currently known. Then define the utility function, you can get

$$u(x) = \begin{cases} 0, & f(x) > f' \\ 1, & f(x) \leq f' \end{cases}$$

$$a_{PI}(x) = E[u(x)|x, D] = \int_{-\infty}^{f'} N(f; u(x), K(x, x)) df = \Phi(f' ; u(x), K(x, x))$$

Only the maximum value of 1 is needed to find the optimal hyperparameter.

5. Failure case analysis

5.1. Feature parameter selection and sample distribution

This paper adopts a data set of vibration signals of bearings in different health states of Canada-University of Ottawa. The data is the vibration data measured by the accelerometer, including vibration signals collected from bearings of different health conditions under time-varying increasing speed conditions. The data are all sampled at 200,000 Hz, and the sampling duration is 10 seconds. The bearing status in the data includes five categories: (1) healthy, (2) faults with inner race defects, (3) faults with outer race defects, (4) faults with ball defects, (5) combined defects The faulty inner ring, outer ring and ball. The working speed condition is speed increase.

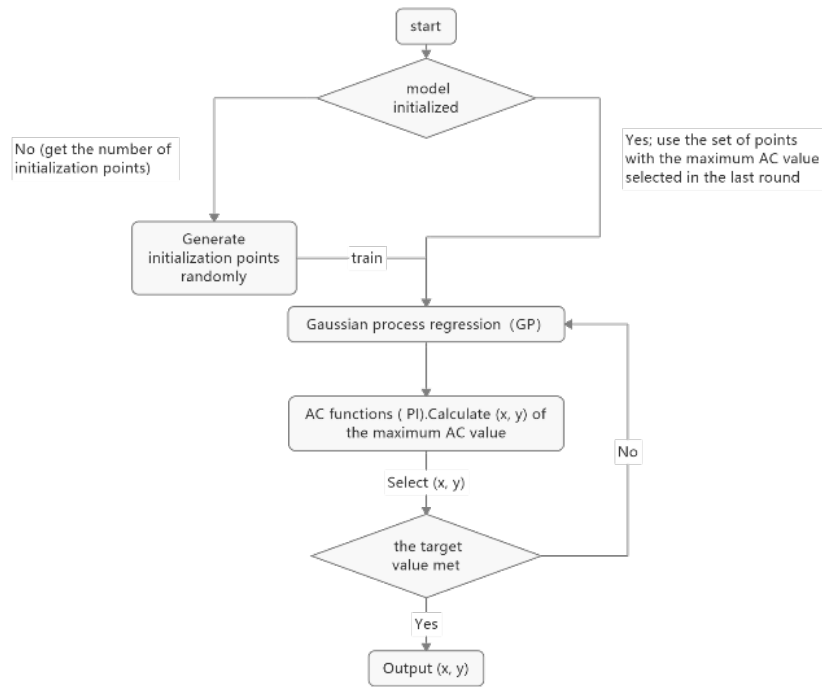


Fig.2: Bayesian optimization process

The collected time-varying bearing vibration signals are grouped into groups of 200 sample points, and the 21 feature extraction values of each group are calculated, and the combination is used as the feature vector of each group of samples, and they are used in optimization decision-making. The training of fault category diagnosis is carried out in the tree model.

5.2. Result analysis

Based on the training data and test data, the resulting bearing fault diagnosis confusion matrix is shown in Figure 3:

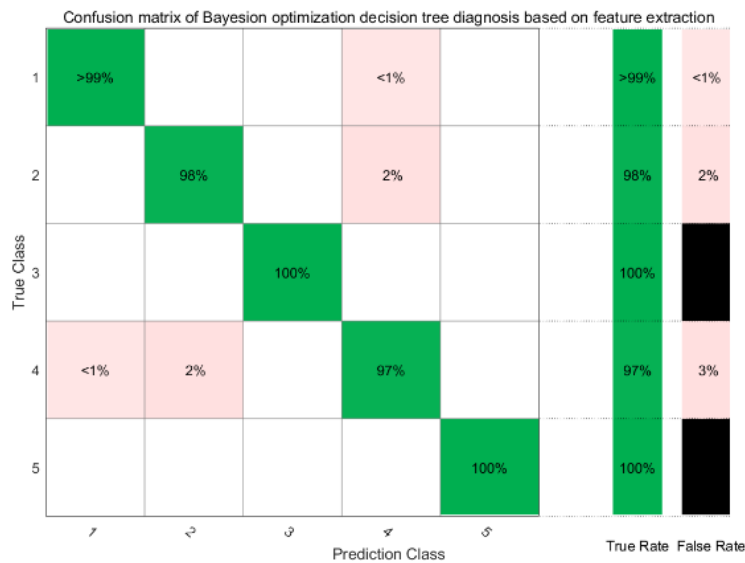


Fig.3 Optimizing the confusion matrix of decision tree model text date (Feature extraction)

5.3. Result analysis

In order to better analyze the performance of the model, this paper uses the same training set and test set to establish an optimized decision tree model with feature extraction, a decision tree model

with feature value extraction, and a decision tree model without feature value extraction. Comparative analysis of results.

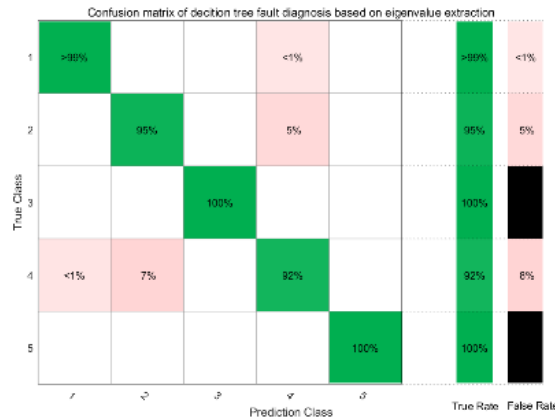


Fig.4 the confusion matrix of decision tree model text date

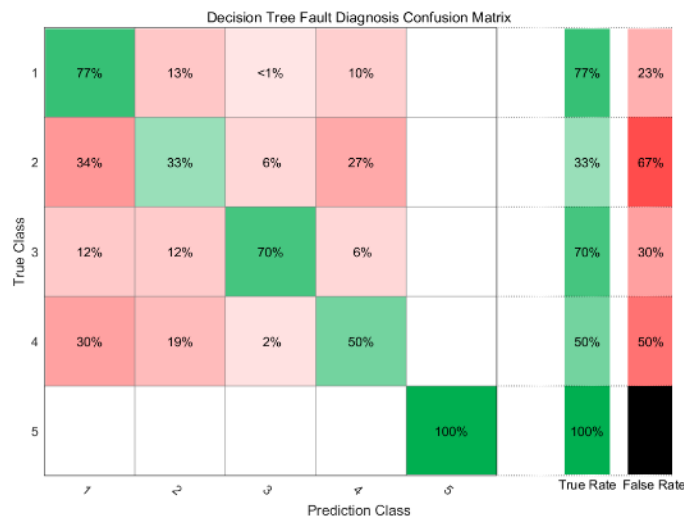


Fig.5 the confusion matrix of decision tree model text date

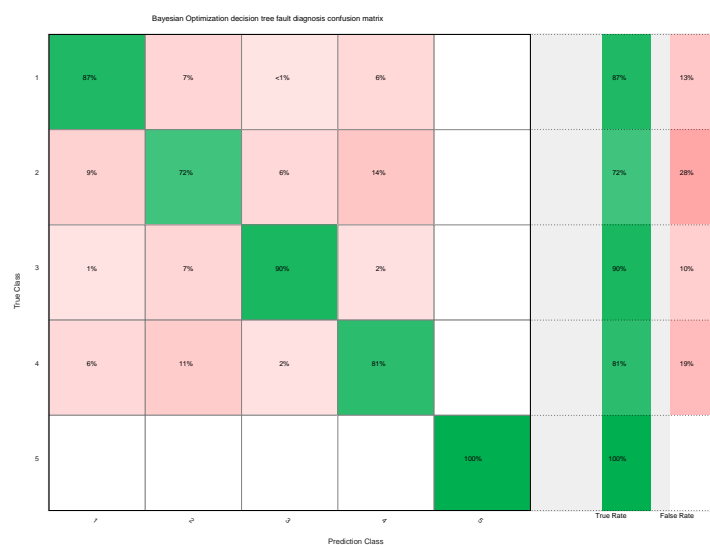


Fig.6 Optimizing the confusion matrix of decision tree model text date

Through the comparison of the fault diagnosis results of the above different models, it can be seen that the extraction of eigenvalues has greatly improved the accuracy of the fault diagnosis of motor bearings. At the same time, the eigenvalue extraction of the data set can reduce the dimensionality of the data and effectively improve the fault diagnosis speed.

6. Conclusion

Aiming at the shortcomings of the traditional motor bearing fault diagnosis model for insufficient data feature extraction, this paper uses the full-angle and high-precision extraction of data features in the time domain, frequency domain, and time-frequency domain. A decision tree diagnosis model based on Bayesian optimization. In order to test the superiority of data feature extraction and optimized decision tree model in motor bearing fault diagnosis, the data set of Ottawa Bearing Health Data Set, Ottawa, Canada was used to test the model. Then it was compared with the bearing fault diagnosis results of the decision tree model with feature value extraction, the optimized decision tree model without feature value extraction, and the decision tree model without feature value extraction.

The diagnosis model constructed based on the optimized decision tree extracted from the special diagnosis of the data performs the fault diagnosis of the motor bearing, and the following conclusions are made:

1) Compared with the decision tree model without feature value extraction, the diagnosis result has greatly improved the accuracy of fault category diagnosis. At the same time, the data feature extraction is used to reduce the data dimension, and the fault diagnosis rate is greatly improved.

2) The Bayesian hyperparameter optimization decision tree model using PI as the acquisition function, the final diagnosis accuracy of the model can reach 98.9%, which is better than the decision tree model without parameter optimization.

3) The Bayesian optimization decision tree model based on multi-domain feature extraction proposed in this paper has the characteristics of high accuracy and high diagnostic speed, and has a good application prospect in the field of motor bearing fault diagnosis.

References

- [1] XU Bo¹, HUANG Chenrong², LU Ali². Induction Motor Common Faults and Key Fault Diagnostics [J]. MICROMOTORS,2015,48(05):106-110.
- [2] ZHAO Hui-min, ZHANG Zhi-qiang. Study on Selection of Characteristic Parameters for Bearing Fault Diagnosis [J]. Internal Combustion Engine & Parts,2019(19):49-51.
- [3] WANGJin-f u, LIFu-cai. Review of Signal Processing Methods in Fault Diagnosis for Review of Signal Processing Methods in Fault Diagnosis for Machinery Using Time-frequencyAnalysis[J]. Noise and Vibration Control,2013,33(03):198-202.
- [4] SHI Guangyu¹, XU Jian², YANG Qiang². Intelligent Fault Diagnosis on Wind Turbine Bearing Based on Convolutional Neural Network [J]. Journal of North China Electric Power University,2020,47(04):71-7