Fault Diagnosis Method of Motor Bearing Based on Bayesian Optimized Quadratic Discriminant Analysis

Jinshan Lin

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

Abstract: Motor bearing is an important part of the motor, and timely troubleshooting of motor bearing faults is of great significance to the safe and stable operation of the motor. This paper proposes a fault diagnosis method for motor bearings based on Bayesian optimized quadratic discriminant analysis (QDA). This method takes the QDA model as the main diagnostic model, with gaussian processes as the probabilistic surrogate model and expected improvement function as the collection function. The hyperparameter set of the model is optimized by using Bayesian optimization (BO). In addition, the diagnosis results of the support vector machine (SVM) and k-nearest neighbor (KNN) models are compared with QDA based on the same data set. The experimental results show that: Bayesian optimized QDA has a better performance.

1. Introduction

The normal operation of the motor is of great significance to industrial production. At the same time, due to the complex structure and operating environment of the motor, after it is put into use, under long-term operation, the equipment will age, and the internal faults may occur [1]. Therefore, the accurate diagnosis of various latent faults of the motor is of great significance to the stable operation of the motor. Motor bearings are most prone to get faults during motor operation [2], so how to improve the efficiency and accuracy of motor bearing fault diagnosis and identification has become a hot spot for researchers at home and abroad.

In recent years, machine learning and artificial intelligence technologies such as artificial neural networks (ANN) [3], support vector machines (SVM) [4], Bayesian networks, fuzzy theory, etc. have been widely used in motor bearing fault diagnosis. However, single machine learning algorithm has some limitations in actual motor fault diagnosis. The training speed of artificial neural network is slow, and it is difficult to find the global optimal solution; fuzzy theory is subjective and the training accuracy is not high; support vector machine has the best diagnosis effect, but it has higher requirements for the selection of kernel function [5]. At present, a variety of optimization algorithms are applied to improve machine learning methods and used in motor bearing diagnosis. Some scholars extract the vibration signal of the motor bearing, and combine it with the particle swarm optimized SVM to achieve good fault diagnosis result [6]; some scholars use wavelet decomposition [7] to extract features of the signal, and use the adaptive neural network model for diagnosis to achieve nice effect.

The result of single machine learning algorithm applied to diagnosis is often determined by its own parameter selection. If a suitable hyperparameter set can be found, the diagnosis result will be more accurate. However, the current researches on finding the optimal parameter set mostly adopt random searching approach and grid search method for parameter optimization. The random searching approach adopts random optimization of parameters, which is often difficult to find the optimal solution; the grid search method will traverse the parameter set to find the optimal set, so that the optimal solution can be found, but it takes a lot of time. When encountering a huge amount of data and a large number of parameters obviously the grid search method is not feasible any more[8]. When encountering large-scale parameter set optimization, Bayesian Optimization (BO) [9] is the best choice. It uses a priori function and acquisition function to find the optimal parameter set with less

computing time. At present, BO has been widely used in navigation, sensors, service interaction and other fields, and has achieved good results.

This article builds a quadratic discriminant analysis (QDA) model at the beginning. Then use Bayesian optimization to optimize the hyperparameter set of the model, and validates the algorithm based on the data set obtained from the Bearing Data Center of Case Western Reserve University in the United States. Comparing the results of the proposed method with other machine learning diagnosis, the experimental results show that the use of Bayesian optimized quadratic discriminant analysis method can achieve high accuracy and low missive judgement rate, which is of practical value for the fault diagnosis of motor bearings.

2. Bayesian optimization of quadratic discriminant analysis algorithm

The discriminant analysis method is based on certain discriminant criteria, establishes one or several discriminant functions, uses the large-scale data of the research object to determine the relevant parameters in each discriminant function, and then compares with the discriminant indicators, so that the samples can be classified one by one. QDA belongs to Bayesian discrimination algorithm. According to the order of discriminant function, Bayesian discriminant is divided into linear discriminant analysis (LDA) and quadratic discriminant analysis (QDA)[10]. The Bayesian discriminant principle is based on the sample data that has been mastered, summarizes the classification law of the research object, and then establishes a discriminant function to classify the sample.

1.1 Principle of Quadratic Discriminant Analysis

Suppose that the analysis data is divided into m categories: ω_1, ω_2, ... , ω_m, and the prior probability of each category is $P(\omega_1)$, $P(\omega_2)$, ... , $P(\omega_m)$. Given an input vector X of unknown class label, $X = \{X_1, X_2 ... , X_t \}$. t is the number of feature quantities of the input vector. If the probability $P(\omega_i)$ of X belonging to the class ω_i is the largest, X will be classified into the class i according to the Bayesian judgment rule. That is, $P(\omega_i) > P(\omega_j), j = 1,2,...,m ; j \neq i$, then X is classified into category i. It can be known from Bayes' theorem:

$$P(\omega_i|X) = \frac{P(X|\omega_i)P(\omega_i)}{P(X)}$$

If the probability of ω_i in the entire class $P(\omega_i)$ is known, in order to make a decision, the conditional density $P(X|\omega_i)$ must be estimated.

It is usually assumed that the conditional density is a Gaussian normal distribution function, namely:

$$P(X|\omega_i) = \frac{1}{(2\pi)^{t/2}|\Sigma_i|^{1/2}} exp\left\{-\frac{1}{2}(X - u_i)^T \Sigma_i^{-1}(X - u_i)\right\}$$

In the formula, $u_i = \frac{1}{n_i}\sum_{j=1}^{n_i} X_{ij}$, $\Sigma_i = \frac{1}{n_i}\sum_{j=1}^{n_i}(X_{ij} - u_i)(X_{ij} - u_i)^T$ respectively the mean vector and covariance matrix of the i-th type; t is the number of eigenvalues of the input vector X

Then take the logarithm of both sides of the formula and simplify it to get:

$$lg\big(P(\omega_i|X)\big) = lg(P(X|\omega_i)) + lg\big(P(\omega_i)\big) - lg\big(P(X)\big) = -\frac{1}{2}(X - u_i)^T \Sigma_i^{-1}(X - u_i) - \frac{1}{2}lg(|\Sigma_i|)$$

$$-\left(\frac{P}{2}lg(2\pi)\right) + lg\big(P(\omega_i)\big) - lg\big(P(X)\big)$$

if $g_i(X) = lg\big(P(\omega_i|X)\big)$, then the formula can be further simplified to

$$g_i(X) = lg\big(P(\omega_i)\big) - \frac{1}{2}(X - u_i)^T \Sigma_i^{-1}(X - u_i) - \frac{1}{2}lg(|\Sigma_i|)$$

The above formula is the judgment equation of the quadratic discriminant method, which can be used to estimate the corresponding probability. The discriminant equation obtained above is a quadratic polynomial about x, and the classification boundary is a curve, which is different from a

linear discriminant. In some calculations, we also use the variance reduction method to process the estimation of the covariance matrix.

1.2 QDA considering Bayesian optimization

Bayesian optimization is a global optimization algorithm based on probability distribution [11], which has a wide range of applications and is extremely efficient. Bayesian optimization first uses the given objective function to be optimized, randomly sampling in the parameter set, initially searching for the objective function distribution, and then searching for the optimal solution that can make the objective function reach the maximum or minimum according to the data information, and continuously Iterate until the distribution fitted by the sampling points is close to the true objective function.

Bayesian optimization mainly has two core processes, Prior Function (PF) and Acquisition Function (AC). The acquisition function can also be called Utility Function.

PF can be divided into parametric model and non-parametric model according to the parameters of the model. Compared with the parametric model in the optimization process, the non-parametric model has the characteristics of the constant number of parameters, which has stronger flexibility and scalability, so the non-parametric model is better used to describe the unknown objective function. Among the non-parametric models, Gaussian Processes [12] (Gaussian Processes, GP) can theoretically achieve the fitting effect of countless multilayer neural networks, so it is the most widely used. The collection function is the basis for determining how to search for the next optimal solution from the parameter set [13]. There are currently three types: Upper confidence bound (UCB), probability of improvement (PI), Expected improvement (EI). EI can solve the problem that PI does not consider how much the unknown point is larger than the known point. The EI function seeks the expectation that the unknown point function value is greater than $f(x^+)$, that is, the following formula:

$$EI(x) = \begin{cases} \left(u(x) - f(x^+)\right)\phi(z) + \sigma(x)\phi(z) & \sigma > 0 \\ 0 & \sigma > 0 \end{cases}$$
$$z = \frac{u(x) - f(x^+)}{\sigma}$$

Where $\phi(\cdot)$ is the normal cumulative distribution function, $\Psi(\cdot)$ is the normal probability density function, and $f(x^+)$ represents the existing maximum value.

The corresponding optimization process of the quadratic discrimination based on Bayesian optimization is:

1) Randomly generate an initial solution within the set range of the model hyperparameter set, and obtain the initial distribution and initial sampling set of the objective function by optimizing the secondary discriminant model test.

2) Select the next most potential evaluation point x_t according to certain criteria, which can maximize the EI acquisition function. Then it is substituted into the model calculation to obtain the objective function value f_t corresponding to x_t .

3) Add new (x_t, f_t) to the historical sampling set samples, and perform Gaussian process iteration. After correction, a Gaussian model closer to the true distribution of the objective function is obtained.

4) When the maximum number of iterations M (this article is set to 30 times), output the optimal sampling point and the corresponding optimal value of the objective function and stop the iteration.

Model parameter optimization flowchart:
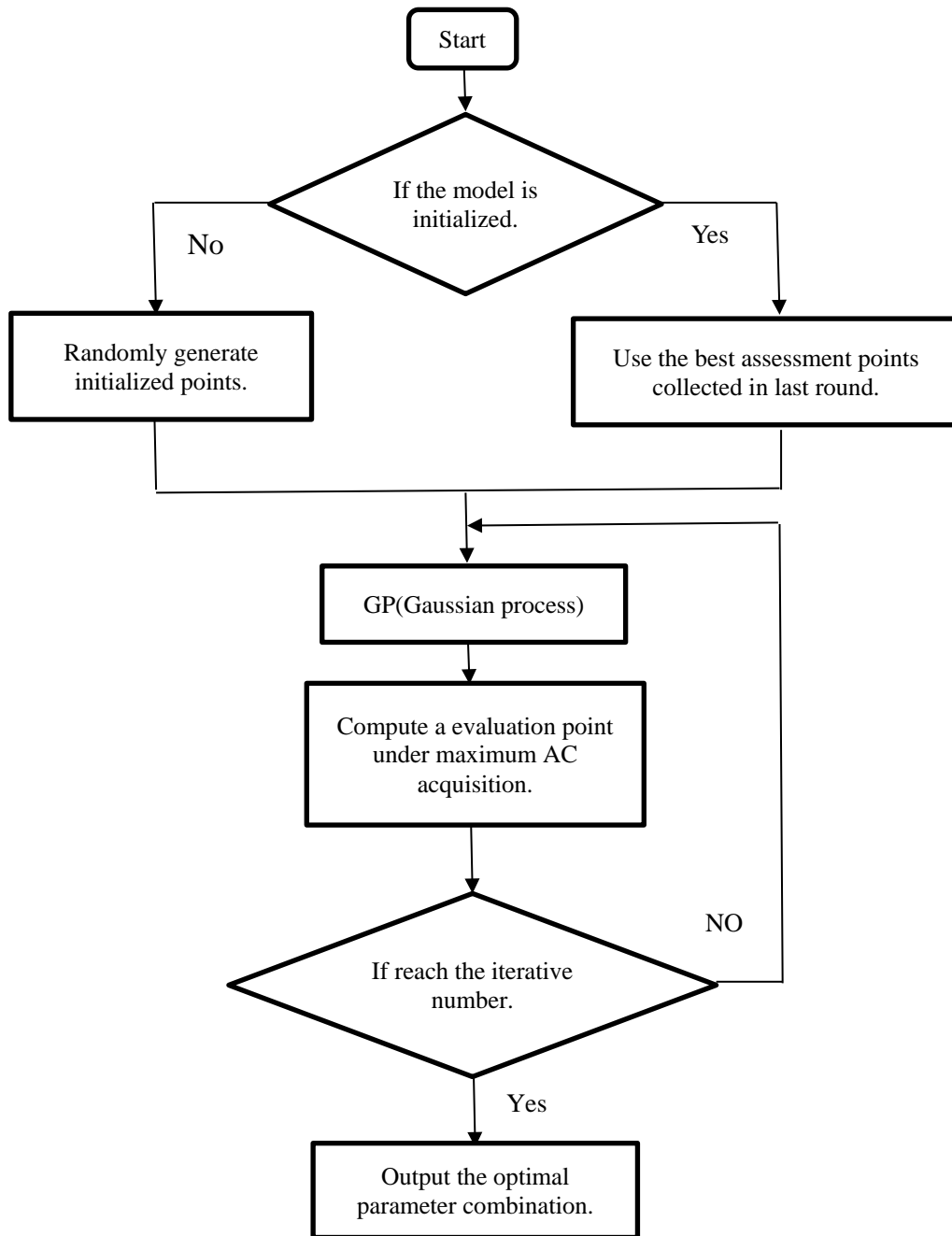
3. Failure case analysis

2.1 Feature parameter selection and sample distribution

This article uses the experimental data obtained from the Bearing Data Center of Case Western Reserve University in the United States, measured by the bearing test system shown in the Figure 1.
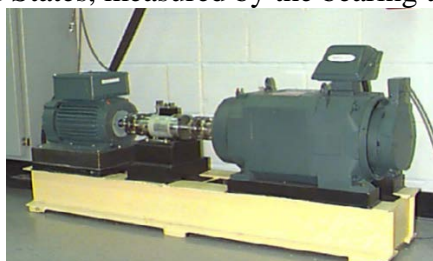

Fig.1 Bearing test system

The bearing model is SKF6205-RS, and the data is measured at a sampling frequency of 48kHz and a speed of 1750r/min; it includes 4 kinds of bearing states, which are normal bearings, bearings with inner ring damage diameters of 0.1778mm, 0.3556mm, and 0.5334mm and the damage depth is 0.2794mm. The collected vibration signals are divided into a group of 20 sample points, and the group is used as the feature vector, and input into the quadratic discriminant analysis model based on Bayesian optimization for training and verification.

This article uses a 5-fold cross-validation method to train the motor bearing data. N-fold cross-validation is to divide a data set into N parts, in which the N-1 parts are as the training set, and one part is as the test set, until each part has been the test set so as to prevent the occurrence of overfitting.

2.2 Result analysis

The training data and test data is divided by five-fold cross-validation, the maximum number of iterations is 30, and the confusion matrix is shown in Figure 2:

Fig.2 QDA model test data confusion matrix

Tab.1 Data distribution of the fault samples

Sample type	normal operation	Fault type 1	Fault type 2	Fault type 3
Accuracy Rate	99%	96%	83%	94%
False Negative Rate	1%	4%	17%	6%

The overall diagnostic accuracy of the model reached 92.9%, and the iteration time is 165.89s. In real life, if the diagnosis method classifies the bearing of the fault type incorrectly, it can be re-diagnosed manually, which has little impact on production safety; but if the faulty bearing is diagnosed as a normal bearing by the diagnosis method, manual intervention will not be performed again, and the motor may have a safety accident due to a faulty diagnosis. Therefore, if we only consider the misjudgment rate of the faulty bearing as a normal bearing, the correct diagnosis rate of the proposed method can reach 99.78%.

4. Comparison of diagnosis effects of different machine learning methods

In order to analyze the diagnostic performance of the Bayesian-optimized quadratic discriminant analysis model, this paper uses the same data set to train and test two machine learning fault diagnosis models: SVM and weighted KNN. The SVM uses a quadratic polynomial kernel function to make high-order nonlinear functions separable. The number of neighbors of the weighted KNN is set to 10, and the data is standardized, using Euclidean as the distance measurement, and inverse distance square as the distance weight. The confusion matrice of the diagnosis results of the SVM and

weighted KNN are shown in Figure 3 and Figure 4, respectively. The performance comparison of them is shown in Table 2.
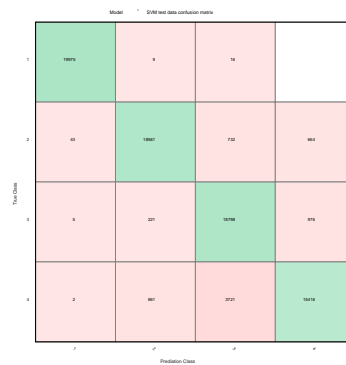

Fig.2 KNN model test data confusion matrix   Fig.3 SVM model test data confusion matrix

Tab. 2 Comparison between the QDA, SVM and KNN

Model	Accuracy Rate	False Negative Rate	Run Time
QDA Based on OA	92.9%	0.22%	115.36s
SVM	90.9%	0.06%	4397.3s
KNN	83.1%	5.67%	1885.5s

Comparing the diagnostic effects of the three models in Table 2, it can be seen that the Bayesian optimization-based quadratic discriminant analysis model has the highest fault diagnosis accuracy, which is greatly improved compared with the other traditional machine learning methods SVM and KNN models. Although the SVM model has the lowest misjudgment rate, its running time is 37 times that of the Bayesian-optimized quadratic discriminant analysis model. Therefore, under the background of the era of big data, the Bayesian optimization quadratic discriminant model has a better application prospect in the field of motor bearing fault diagnosis due to its high diagnostic efficiency and high accuracy.

5. Conclusion

Aiming at the shortcomings of the traditional single algorithm based motor bearing fault diagnosis model, this paper proposes a quadratic discriminant analysis diagnosis model based on Bayesian optimization. In order to test the effectiveness of the quadratic discriminant analysis algorithm based on Bayesian optimization in motor bearing fault diagnosis, the data set of the Bearing Data Center of Case Western Reserve University was used to test the model. Then, it was also compared with the effects of the other machine learning methods SVM and KNN in the motor bearing fault diagnosis.

The conclusion is as follows:

1) Diagnosis results of motor bearings are compared by using machine learning methods SVM and KNN, the quadratic discriminant model based on Bayesian optimization. And it's proven that the quadratic discriminant model has higher fault diagnosis efficiency and accuracy.

2) Adopting the Gaussian process as the probabilistic proxy model and the Bayesian optimization model with EI as the acquisition function, the final diagnosis accuracy of the quadratic discriminant analysis of Bayesian optimization can reach 92.9%. If only when the faulty bearing is considered as a normal bearing, the diagnosis method is wrong, then the classifier can reach an amazing accuracy of 99.78%.

3) The quadratic discriminant analysis method based on Bayesian optimization proposed in this paper has a wide range of applications, not only for fault diagnosis of motor bearings, but also for fault diagnosis of large rotating machinery and small electronic components [14].

References

[1] XU AI-hua, DU Yang, YUAN Tao. Research on Fault Diagnosis of Motor Bearing Based on Deep Learning [J]. Modular Machine Tool & Automatic Manufacturing Technique,2020(03):45-48+54.

[2] Zhao Jinguo, Zhang Xiaojuan. Control Method Based on Hub Motor Fault Diagnosis and R econstruction,2020,53(07):75-78.

[3] Gong Maofa, Liu Yanni, Wang Laihe, et al. Fault diagnosis of power transformers based on chaos particle swarm optimization BP neuralnetwork[J]. Electrical Measurement & Instrumentation, 2016,53(15):13-16, 32.

[4] Li Ying, Shu Naiqiu, Transformer Fault Diagnosis Based on Fuzzy Clustering and Complete Binary Tree Support Vector Machine[J].Transactions of China Electrotechnical Society, 2016, 31(4)：64-70.

[5] POLIKAR R. Ensemble based systems in decision making[J]. IEEE Circuits & Systems Magazine, 2006, 6(3)：21-45.

[6] Ding Ruicheng，Huang Yourui，Chen Zhenping，et al．Research on motor bearings fault diagnosis based on LMD and SVM[J]．Modular Machine Tool & Automatic Manufacturing Technique，2016(8)：81-84(in Chinese)．

[7] Zheng Lantian ， Li Ansheng ， Du Wenliao ． Self-organization map neural network diagnosis based motor bearing fault visualization[J]．Machinery Design & Manufacture，2013(12)：73-75，78(in Chinese)．

[8] Wang Xue, Han Tao. Transformer fault diagnosis based on Bayesian optimized random forest [J/OL].Electrical Measurement & Instrumentation:1-8[2020-08-08].http://kns.cnki.net/kcms/detail/23.1202.TH.20200106.1605.028.html.

[9] Liu Xiao-xiong ,Wu Yan,Shi Jing-ping,Zhang Wei-guo .Method一of fault diagnosis using Bayesian optimization algorithm [J].Appljcation Research of Computers,2009,26(01):129-131.

[10] YANG Hui-zhong,JIA Shu-kuang.Multi一model Modeling Method Based on Quadratic Discriminant Analysis[J].Control Engineering of China,2010,17(05):662-664+685.

[11] Tong Zhaojing ,Shi xiuhua,Wang Wenbing,Du Xiangdang,Tian Jing.Generator Rotor Vibration Fault Diagnosis Based on Optimizition Clustering Bayesion Network [J].Journal of Vibration,Measurement & Diagnosis,2014,34(02):237-241+394.

[12] LI Congcong ,WANG Tong ,XIANG Yuwei,WANG Zengping, SHI Bonian, ZHANG Yan .Analytical method based on improved Gaussian mixture model for probabilistic load flow [J]. Power System Protection and Control ,2020,48(10):146-155.

[13] Ding Ruicheng ,Huang Yourui ,Chen Zhenping，et al．Research on motor bearings fault diagnosis based on LMD and SVM[J].Modular Machine Tool & Automatic Manufacturing Technique，2016(8)：81-84(in Chinese)．

[14] LI BING, Han Rui, HE Yigang , ZHANG Xiaoyi ,HOU Jinbo. Applications of the Improved Random Forest Algorithm in Fault Diagnosis of Motor Bearings[J]. Proceedings of the CSEE,2020,40(04):1310-1319+1422.