Research Article

Rolling Bearing Fault Diagnosis Based on SVM Optimized with Adaptive Quantum DE Algorithm

Yuanyuan Li, Qichun Sun, Hua Xu, Xiaogang Li, Zhijun Fang, and Wei Yao

1Shanghai University of Engineering Science, Songjiang, Shanghai 201620, China
2Kunfeng Quantum Technology Co. Ltd., Pudong, Shanghai 201203, China

Correspondence should be addressed to Yuanyuan Li; liyuanyuanedu@163.com

Received 16 January 2022; Revised 26 March 2022; Accepted 25 April 2022; Published 17 May 2022

Copyright © 2022 Yuanyuan Li 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 optimize traditional fault diagnosis models for practical applications, a fault diagnosis model based on support vector machines optimized with the adaptive quantum differential evolution of (AQDE-SVM) is proposed in this study. First, the traditional differential evolution is rewritten based on real number encoded into a qubit encoding. Second, this study proposes an adaptive quantum rotation gate and uses this gate to update the probability amplitude of the qubits. Finally, compared with quantum genetic algorithm support vector machines (QGA-SVM) and differential evolution-support vector machines (DE-SVM), etc., the results show that the algorithm proposed in this study has a higher diagnosis accuracy and shorter running time, providing great practical engineering value in the application of rolling bearing fault diagnosis.

1. Introduction

Rolling bearing is one of the most widely used mechanical parts in the modern industrial system for power transmission, in the meantime, it is also one of the most prone components to have failures. Rolling bearings account for more than 51% of failures in induction motors alone [1]. Failures of rolling bearings will cause damage to other parts and even the entire mechanical system. In severe cases, they can lead to catastrophic accidents. Therefore, it is of great practical significance to carry out fault diagnosis research on rolling bearings [2].

With the rapid improvement of Artificial Intelligence, intelligent fault diagnosis based on machine learning has gradually become the mainstream of research. Compared with traditional methods, the intelligent diagnosis method with higher accuracy and more reliable diagnostic performance does not depend on human experience. There are many relevant research achievements promoting the development of fault diagnosis methods such as convolutional neural network (CNN) [3], random forest [4], back propagation network [5], generative adversarial network [6], fuzzy C-means [7], and other machine learning algorithms.

Normally, it is difficult to collect a large amount of data on the failed rolling bearings since the machinery generally works in good condition. Therefore, fault diagnosis research on rolling bearings is usually considered as the small-sample problem. Support vector machine (SVM), one of the most classic algorithms in machine learning, classifies data by solving the maximum margin hyperplane of training samples. SVM is not only widely used in portrait recognition, text classification, handwritten character recognition, and other classification tasks but also has a considerable degree of superiority when it is used to deal with small sample problems. In recent years, SVM has become one of the most used classification methods [8], moreover, researchers have extended the application field of SVM for the rolling bearing fault diagnosis with good empirical results.

Zhang et al. proposed a comprehensive strategy of combining mixed kernel-support vector machine with grasshopper optimization algorithm to identify typical faults of rotating machinery. The mixed kernel can balance the learning ability and generalization ability of SVM, and the optimization algorithm can find the proper parameters, thereby improving the accuracy of SVM rolling bearing fault diagnosis [9]. Chen et al. improved the traditional SVM to
adboost-SVM with ensemble learning method and used the grid search method to traverse the search space to find the best parameters. The ensemble learning method can effectively improve the performance of the original algorithm, but the weights need to be adjusted repeatedly, which will be expensive for the computation, and the grid search method will further increase the time of model training [10]. Huo et al. considered that the vibration of rolling bearing involves nonlinear and unbalanced characteristics. Therefore, an adaptive multiscale weighted permutation entropy is proposed to process the original signals of rolling bearing, then the processed signals are fed into SVM for fault diagnosis. The preprocessed signals of rolling bearings can improve the correlation between features and labels, thereby increasing classification accuracy [11]. Zhu and Xiong used the quantum genetic algorithm to optimize SVM for fault diagnosis. Quantum-based genetic algorithm takes advantage of the parallelism of quantum computing and has a faster running speed than traditional genetic algorithm. However, only one rotation gate is used in the review to update the probability amplitude of the qubit, resulting in only one search direction of the genetic algorithm, which reduces the performance of the algorithm to a certain extent [12].

The research of the abovementioned papers shows that the use of SVM for fault diagnosis usually involves two optimization directions. One is to improve the accuracy of SVM itself, and the other one is to adjust the parameters of SVM adaptively using optimization algorithms. Both optimization methods are quite useful. Moreover, because the data of the rolling bearing itself determine the upper limit of the SVM fault diagnosis, and it is more effective to preprocess the original data rather than train the SVM directly. Second, since the parameters used in the SVM model are usually data-specific, it is necessary to use an optimization algorithm to adjust the parameters adaptively. However, traditional optimization algorithms have their limitations and are easy to fall into a local optimum, which leads to a decrease in the diagnosis accuracy of SVM. On the other hand, it takes more time in the optimization process, and in real applications of fault diagnosis, a quick response time is generally required.

In order to improve the accuracy and speed of fault diagnosis, we propose a fault diagnosis model based on SVM optimized with the adaptive quantum differential evolution. The differential evolution is a swarm-based intelligent optimization algorithm proposed by R. Storn and K. Price. Its operation mode and mathematical principles are relatively simple, and it also has a good convergent performance, which is suitable for modification to a quantum form. The research content of this study includes the following three aspects:

(1) An adaptive quantum differential evolution (AQDE) is proposed in this study. AQDE is used to optimize the parameters of SVM and proposed a fault diagnosis model based on AQDE-SVM. Compared with QGA-SVM and DE-SVM, etc., AQDE-SVM has significant advantages in the application of fault diagnosis.

(2) Ensemble empirical mode decomposition (EEMD) combined with time domain and frequency domain was used to extract effective features from the original bearing data in this paper, and then input these features into the fault diagnosis model.

(3) We use Case Western Reserve University (CWRU) bearing datasets for standard fault diagnosis experiment and use Xi’an Jiaotong University-SumYoung (XJTU-SY) bearing datasets for early fault diagnosis experiment. Especially for XJTU-SY bearing datasets, this study uses Pauta criterion to judge the time of fault occurrence and realize early fault diagnosis experiment.

The main contribution of this study is to propose AQDE algorithm. AQDE is improved based on quantum theory and can control the search direction of the DE according to the current number of iterations, thus improve AQDE’s running speed and search performance. As an optimization algorithm, AQDE can be applied to all kinds of search problems, so as to improve the efficiency of the solution.

The rest of the review is organized as follows: We first introduce the fundamentals of quantum theory, SVM and AQDE in Section 2. We then review the methods of data preprocessing and two experiments of fault diagnosis in Section 3. Finally, in Section 4, we summarize the work in this study and conclude with a brief discussion of our future research.

2. Fault Diagnosis Model Based on AQDE-SVM

2.1. Quantum Theory. Quantum theory is one of the greatest achievements of modern physics, emerging technologies such as quantum communication and quantum computers are expected to profoundly change our way of life in the near future. Quantum has many special properties, such as superposition and entanglement, and AQDE proposed in this paper was inspired by quantum superposition. In the field of quantum computing, the quantum superposition refers to the quantum bit (qubit) that will contain bit 0 and bit 1 at the same time before measurement, and will randomly collapse toward bit 0 or bit 1 with a certain probability after measurement. This review proposes a series of strategies to control this randomness so as to help AQDE get better diversity and convergence.

2.2. Fundamentals of SVM. Selecting an appropriate classification algorithm to complete the identification of the different operating states of the rolling bearing is essential in the field of fault diagnosis. In practical engineering applications, it is difficult to obtain enough samples. Therefore, this study selects a support vector machine that has significant advantages in small sample learning problems of rolling bearing fault diagnosis. Assume that training data contains \( n \) samples, and the types of training data are 2, denoted as follows:

\[
(x_i, y_i)_{i=1}^{n}, \quad y_i = \{1, -1\}. \tag{1}
\]

In equation (1), \( x_i \) represents the training data, \( y_i \) is the type of data, including “1” and “-1.”
In order to achieve better robustness and noise resistance of SVM, certain restrictions on the hyperplane are required. When we determine the hyperplane, the first step is to calculate the distances of all training data to the hyperplane and find the closest data to the hyperplane:

\[ d_{\text{min}} = \min_{i=1,2,...,n} y_i \frac{1}{\|w\|} (wx_i + b), \] (2)

where \( w \) and \( b \) are parameters of the hyperplane. Then we search for the \( d_{\text{min}} \) that is the farthest away from the hyperplane, and we can get

\[
\begin{align*}
\max_{w,b} d_{\text{min}} \\
\text{s.t. } y_i (wx_i + b) \geq d_{\text{min}}.
\end{align*}
\] (3)

To facilitate the solution, we can rewrite equation (3) as follows:

\[
\begin{align*}
\min_{w,b} & \frac{1}{2} w^T w \\
\text{s.t. } & y_i (wx_i + b) \geq 1.
\end{align*}
\] (4)

Using the Lagrange multiplier method to solve equation (4), and we can get the parameters of the hyperplane.

2.3. Parameters in SVM. For outliers and nonlinear separable cases, SVM introduces a penalty factor and kernel function. Specifically, if the sample contains outliers, the slack variable \( \xi \) and the penalty factor \( C \) are introduced, and the equation (4) is rewritten as follows:

\[
\begin{align*}
\min_{w,b} & \frac{1}{2} w^T w + C \sum_{i=1}^{N} \xi_i, \\
\text{s.t. } & \begin{cases} 1 - y_i (wx_i + b) \leq \xi_i, \\ \xi_i \geq 0. \end{cases}
\end{align*}
\] (5)

In equation (5), the slack variable \( \xi \) represents the distance between the outlier and the hyperplane. The penalty factor \( C \) represents the importance that SVM attaches to the loss caused by outliers. The larger the \( C \) is, the higher the accuracy is, but the robustness will be reduced accordingly, and vice versa.

If the sample is nonlinearly separable, SVM will map the sample from the \( R^d \)-dimensional original space to the \( R^e \)-dimensional feature space (\( e > d \)) through a mapping \( \varphi \), so that the sample is linearly separated in the new high-dimensional space, and the equation (5) is rewritten as follows:

\[
\begin{align*}
\min_{w,b} & \frac{1}{2} w^T w, \\
\text{s.t. } & 1 - y_i (w^T \varphi(x_i) + b) \leq 0.
\end{align*}
\] (6)

When we solve equation (6), it is necessary to calculate the inner product \( \varphi(x_i)^T \varphi(x_j) \) of the sample in the feature space. For the case of either ultrahigh dimensional or even infinite dimensions, calculating this inner product will be very expensive. After introducing the kernel function, the two steps of function mapping and inner product calculation are reduced to one step. Moreover, the computational complexity is reduced from \( O(e) \) to \( O(d) \). Specifically, a kernel function \( \kappa(x_i, x_j) \) is constructed as follows:

\[
\kappa(x_i, x_j) = \varphi(x_i)^T \varphi(x_j).
\] (7)

The sample inner product of the feature space can be calculated in the original space in equation (7). In real applications, the most commonly used kernel function of SVM is the radial basis kernel function, and its mathematical expression is

\[
\kappa(x_i, x_j) = \exp \left( -\gamma \|x_i - x_j\|^2 \right).
\] (8)

In equation (8), \( \gamma \) is a kernel function parameter. The larger the \( \gamma \) is, the higher the dimension of feature space is. Generally, the higher the dimensionality of the feature space is, the better effect is obtained after training the model; however, it is likely to result in over-fitting and reducibility of the generalization ability of the model.

The penalty factor and the kernel function parameter play a key role in SVM, and the conventional method is to use the heuristic optimization algorithm to adjust these parameters, such as genetic algorithm [13], ant colony [14], and differential evolution [15]. However, these optimization algorithms tend to fall into local optimum in the training process, and the traditional improvement strategy will increase the computational complexity. Therefore, this combines the differential evolution with quantum theory; moreover, through the experimental comparison, it can be proved that quantum differential evolution has significant advantages in accuracy and computational efficiency.

2.4. Build AQDE-SVM. The standard DE [16] adopts real number coding and contains four steps: swarm initialization, mutation, crossover, and selection. In the AQDE, we have improved on all of these steps. First, qubits are used to encode the initial swarm. Second, this review proposes an adaptive quantum rotation gate and uses this gate to update the probability amplitude of every qubit in each generation of the swarm. Finally, the qubits collapse into binary bits after measurement, and we convert the binary bits to decimal numbers and substitute them into SVM. The steps are as follows.

2.4.1. Generating Initial Swarm. The initial swarm in this study contains 50 individual vectors, and the number of dimensions of the individual vector is 2, which corresponds to two parameters of SVM. Each parameter is represented by eight bits. Each bit is a qubit before measurement. The qubit set in this study can be expressed as follows:

\[
|\phi\rangle = \sin \theta |0\rangle + \cos \theta |1\rangle.
\] (9)

In equation (9), \( |\phi\rangle \) is the Dirac symbol, it mathematically represents the column vector in Hilbert space. \( |\phi\rangle \) represents the qubit, which will collapse into classical bit 0 with
probability \( \sin^2 \theta \) and collapse into classical bit 1 with probability \( \cos^2 \theta \) after measurement. Moreover, the probability satisfies \( \sin^2 \theta + \cos^2 \theta = 1 \). \( \theta \) is quantum angle, and satisfies the condition as follows:

\[
\theta \in \text{rand} \( 0, 1 \) \times \frac{\pi}{2}\]

According to equation (9), each qubit needs to generate a quantum angle. Therefore, we need to generate a total of \( 50 \times 2 \times 8 \) quantum angles in the initial swarm.

Combine equations (9) and (10), the mathematical expression of the initial swarm can be obtained

\[
P_{i,g} = \begin{bmatrix} C_{i,g} \\ y_{i,g} \end{bmatrix} = \begin{bmatrix} \phi_{i,g,C,1}, \ldots, \phi_{i,g,C,j}, \ldots, \phi_{i,g,C,8} \\ \phi_{i,g,y,1}, \ldots, \phi_{i,g,y,j}, \ldots, \phi_{i,g,y,8} \end{bmatrix}. \tag{11}\]

In equation (11), \( P_{i,g} \) is the \( i \)th individual vector in the swarm of the \( g \) generation. \( C_{i,g} \) is the first dimension of \( P_{i,g} \), corresponding to the penalty factor of SVM, \( y_{i,g} \) is the second dimension of \( P_{i,g} \), corresponding to the kernel function parameters of SVM. \( \phi_{i,g,C,j} \) is the \( j \)th qubit of \( C_{i,g} \), \( \phi_{i,g,y,j} \) is the \( j \)th qubit of \( y_{i,g} \), and \( j \in [1, 8] \).

2.4.2. Update the Probability Amplitude of Qubit. Differential evolution is an optimization algorithm based on the swarm. We usually require richer swarm diversity in the early stage of iteration so that the individual vector can be fully distributed in the solution space. At the later stage of iteration, the convergence rate is requested to be faster, thus accelerating the convergence of the algorithm. Based on the above considerations, this review proposes an adaptive quantum rotation gate \( G_r \) which is used to update the probability amplitude of qubit, so as to achieve the effect described above. Because the qubit will collapse to \( |0\rangle \) or \( |1\rangle \) with a certain probability after measurement; therefore, in the case of the swarm diversity of AQDE, the probability of qubit collapse to \( |0\rangle \) or \( |1\rangle \) needs to be closer on the original probability. In the case of accelerated convergence of AQDE, the probability of qubit collapse to \( |0\rangle \) or \( |1\rangle \) state needs to be higher than the original probability. The steps are as follows.

The mathematical formula for the collapse of the qubit in the superposition state after the measurement is

\[
\begin{align*}
|\phi\rangle = |1\rangle, \quad \sin^2 \theta > \text{rand} \( 0, 1 \), \\
|\phi\rangle = |0\rangle, \quad \text{else}.
\end{align*}
\]

In equation (12), \( \theta \in (0, \pi/2) \), \( \sin \theta \in (0, 1) \). The larger the \( \sin^2 \theta \) is, the more likely the qubit is to collapse into \( |1\rangle \), and vice versa. Therefore, in the early stage of iteration, the quantum rotation gate controls \( \sin \theta \) to rotate in the direction of 0.707 on the original probability. The schematic diagram is shown in Figure 1.

According to Figure 1, when \( \sin \theta \in (0, 0.707) \), then the rotation direction of \( G_r \) is 1; otherwise, the rotation direction of \( G_r \) is 2.

At the later stage of iteration, \( G_r \) controls \( \sin \theta \) to rotate in the direction of coordinate axis on the original probability. The schematic diagram is shown in Figure 2.

As a similar demonstration is shown in Figure 1, the direction of rotations 3 and 4 is shown in Figure 2. \( \sin \theta \in (0, 0.707) \) is the condition for direction 3; otherwise, the rotation direction of \( G_r \) is 4.

Therefore, the mathematical expression of \( G_r \) proposed in this study is

\[
\begin{bmatrix}
\sin \theta \\
\cos \theta
\end{bmatrix} = \begin{cases}
\begin{bmatrix}
\sin \theta \\
\cos \theta
\end{bmatrix} \cdot \text{rand} \( (0, 1) \times \frac{\pi}{2} \), & \text{if } \theta \in (0, \pi/2), \\
\begin{bmatrix}
\sin \theta \\
\cos \theta
\end{bmatrix} \cdot \text{rand} \( (0, 1) \times \frac{\pi}{2} \), & \text{if } \theta \in [\pi/2, \pi),
\end{cases}
\] \tag{13}

In equation (13), \( f(r) \) represents the counterclockwise quantum gate, and the mathematical formula is

\[
f(r) = \begin{bmatrix}
\cos r - \sin r \\
\sin r \cos r
\end{bmatrix}. \tag{14}\]
In equation (13), $g(r)$ represents the clockwise quantum gate, and the mathematical formula is

$$g(r) = \begin{bmatrix} \sin r \cos r \\ \cos r - \sin r \end{bmatrix}. \quad (15)$$

In equations (14) and (15), $r$ is the rotation angle, which affects the convergence speed of AQDE. A too-large rotation angle will cause the AQDE to converge prematurely, and a too-small rotation angle will cause the AQDE to converge slowly or even fail to converge. Han [17] pointed out that it is appropriate to set the interval of the rotation angle within $[0.001\pi, 0.05\pi]$.

In equation (13), where $G$ represents the maximum number of iterations, $g$ represents the current number of iterations, rand(0, 1) is a randomly generated number from 0 to 1. As can be seen from the formula, in the early stage of the AQDE iteration, the smaller the $g$ is, the easier the inequality rand(0, 1) < (G-g)/G is established, and vice versa. $\sin \theta$ and $\cos \theta$ are the original probability amplitude of the qubit, and $\sin \theta$ and $\cos \theta$ are the updated qubit probability amplitude.

### 2.4.3. Obtain the Diagnostic Accuracy of SVM

In this review, each parameter is represented in eight qubits. After measurement, we got eight classical binary bits, and the corresponding decimal range of eight binary is $[0, 2^8]$. Eitrich [18] points out that it is advisable to control the parameter value range of SVM within the interval of $(0, 100)$; therefore, the decimal needs to be scaled. The mathematical formula for scaling is

$$d^* = \frac{a-b}{2^a} \times d. \quad (16)$$

In equation (16), $a$ and $b$ represent the upper limit and lower limit of SVM parameter values, respectively, $n$ is the number of the qubit, $d$, and $d^*$ are parameters of SVM before and after scaling. According to the parameters introduced by Eitrich [18], this study sets $a = 100$, $b = 0.01$. Finally, we use AQDE-SVM for fault diagnosis, so as to obtain the diagnosis accuracy.

### 3. Fault Diagnosis Experiment

The dataset used in this study comes from the CWRU bearing datasets [19] and XJTU-SY bearing datasets [20]. In order to evaluate the effectiveness of AQDE-SVM more comprehensively, we use CWRU bearing datasets for standard fault diagnosis experiment, and use XJTU-SY bearing datasets for early fault diagnosis experiment. The description of the CWRU datasets is shown in Table 1.

The computer used in the experiments is configured with an i5-9300H CPU, clocked at 2.4 GHz, memory 16 GB, and the programming language used is Python.

#### 3.1. Data Preprocessing

In traditional Fourier transform, the frequency is defined according to a complete cycle, which leads to the poor effectiveness of Fourier transform in processing bearing signals. Therefore, EEMD [21] is used to preprocess the rolling bearing data in this study. EEMD overcoming the limitation of Fourier transform in the constant frequency and solving the phenomenon of modal aliasing in Empirical Mode Decomposition (EMD) [22].

In order to intuitively understand the effect of EEMD, Bearing1_1 in Table 2 is considered as an example in this study. The Bearing1_1 data reconstructed into a signal matrix of 100×1024 (the number of samples is 100, and each sample is composed of 1024 continuously sampled data points), and we take the first sample as an example for EEMD. The decomposition results are shown in Figure 3.

In Figure 3, IMF represents intrinsic mode function. However, not all IMFs contain effective features, so screening out suitable IMF is the key to extracting the fault information of rolling bearings.

Kurtosis is one of the most widely used statistical parameters in the field of fault diagnosis. During the normal operation of the rolling bearing, the kurtosis value is approximately equal to 3. When the fault occurs, causing the amplitude distribution of the vibration signals to be skewed or scattered, then the corresponding kurtosis value will increase accordingly. Therefore, the greater the kurtosis value is, the more serious the failure of the rolling bearing is. Similarly, if the kurtosis value of the IMF is large, the fault information contained in this IMF is mostly complete. The mathematical formula for calculating kurtosis is

$$X_{kurt} = \frac{n{\sum_{i=1}^{n} (x_i - x_{mean})^4}}{[\sum_{i=1}^{n} (x_i - x_{mean})^2]^2}, \quad (17)$$

where $x_i$ represents the sample, $x_{mean}$ represents the average value of the sample over $n$ seconds, $n$ represents the total number of samples.

In this review, equation (17) is used to screen the IMF with the largest kurtosis value and use the IMF as the data sample for subsequent analysis. In order to directly reflect the screening process, this review takes the IMFs decomposed in Figure 3 as an example, and the kurtosis of each IMF is shown in Figure 4.
Table 1: Data introduction (CWRU bearing datasets).

| Dataset | Fault diameter | Fault type | Label |
|---------|----------------|------------|-------|
| Bearing1_1 | 0.1778 inch | Inner race | H1    |
| Bearing1_2 | 0.1778 inch | Rolling element | H2    |
| Bearing1_3 | 0.1778 inch | Outer race (6:00) | H3    |
| Bearing1_4 | 0.1778 inch | Outer race (3:00) | H4    |
| Bearing1_5 | 0.1778 inch | Outer race (12:00) | H5    |
| Bearing1_6 | - | Normal | H6    |

Table 2: Data introduction (XJTU-SY bearing datasets).

| Dataset   | Actual life      | Fault type | Number of samples |
|-----------|------------------|------------|-------------------|
| Bearing2_1 | 8 h11 min       | Inner race | 491               |
| Bearing2_2 | 2 h41 min       | Outer race | 161               |
| Bearing2_3 | 8 h53 min       | Cage       | 533               |

It could be seen from Figure 4 that the kurtosis value of IMF1 is the largest, so IMF1 is selected as the data sample for subsequent analysis.

Furthermore, kurtosis is more sensitive to the early level of faults, but as the level of faults increases, the kurtosis parameter gradually saturates, and the diagnostic ability decreases. Therefore, crest factor and power spectrum are used to improve the effectiveness of data preprocessing in this review.

The crest factor belongs to time domain parameter which is more sensitive to faults related to surface damage and wear. The mathematical formula of the crest factor is

\[ X_{crest} = \frac{\max(x_1, x_2, \ldots, x_n)}{x_{rms}}, \] (18)

where \( x_{rms} \) represents the root mean square value of the sample over \( n \) seconds.

The power spectrum belongs to frequency domain analysis, which can extract effective features from the signals. The mathematical formula of the power spectrum is

\[ PS = \frac{|FFT[\text{cor}(x, x), m]|}{\max(FFT[\text{cor}(x, x), m])} \] (19)

where \( FFT \) is fast Fourier transform, \( \text{cor}() \) is autocorrelation function, \( x \) represents the samples and \( m \) represents the length of Fourier transform.

The kurtosis, crest factor, and power spectrum of IMF1 were calculated and their input were used in AQDE-SVM to realize fault diagnosis.

3.2 Fault Diagnosis of CWRU Bearing Datasets. The method of data preprocessing in Section 3.1 is used to process CWRU datasets. In order to simulate real scenarios, where the fault data is much less compared to the normal data, 40 samples are selected from the failure data of Bearing1_1 to Bearing1_5, and 200 samples are selected from the normal data of Bearing1_6 in this review. A total of 70% of them are used as training data and the rest of them are used as test data.

First, in order to verify the advantages of AQDE in optimization efficiency, this review used AQDE-SVM, QGA-SVM, and DE-SVM for comparison. The number of iterations set as 50, and the number of individual vectors or chromosomes is set as 50. Each model is run independently ten times, and we calculate the average accuracy of 0, 10, 20, 30, 40, and 50 iterations, respectively. The accuracy curves are shown in Figure 5.

It can be observed from Figure 5 that since the 20th iteration, the average accuracy of AQDE-SVM is slightly higher than QGA-SVM, and significantly higher than DE-SVM, indicating that the AQDE has better optimization performance.

Second, the comparison experiment of CNN and SVM is added and each model is run independently 10 times. CNN contains four convolutional layers and pooling layers, and SVM is set \( C = 0.6, \gamma = 0.8 \). As the fault diagnosis experiment set in this review is multiclassification, so macroaverage and microaverage are selected to evaluate each model. The comparison results are shown in Table 3.

From the comparison of results in Table 3, it could be seen that the average accuracy, macroaverage, and microaverage of the AQDE-SVM is the highest, and the improved DE based on quantum theory that has a significant improvement in running time compared to the traditional DE. Compared with CNN, AQDE-SVM has the same running time but higher diagnostic accuracy. Although CNN can further improve their diagnostic accuracy with the deepening of network layers, they need to consume more computing resources and running time. Therefore, the AQDE-SVM fault diagnosis model proposed in this review is more cost-effective.

Finally, in order to verify the anti-noise ability of each algorithm in the case of noise interference, three groups of Gaussian white noise were added to the original signal of rolling bearing with \( SNR = 6, SNR = 4, SNR = 2, SNR = 0, SNR = -2, \) and \( SNR = -4 \), respectively. The data preprocessing method is the same as before, AQDE-SVM, QGA-SVM, DE-SVM, SVM, and CNN are used for fault diagnosis, respectively. Macroaverage was selected for evaluation and the comparison results are shown in Figure 6.

It can be observed from Figure 6 that AQDE-SVM achieves the highest macroaverage regardless of strong noise interference or weak noise interference, indicating the excellent anti-noise ability of AQDE-SVM.

3.3 Fault Diagnosis of XJTU-SY Bearing Datasets. XJTU-SY datasets contain the whole process data of rolling bearings from normal to damaged and the fault types. Denote that the inner race fault in the XJTU-SY dataset is I1, the outer race fault is I2, the cage fault is I3, and the normal data is I4. The description of the XJTU-SY datasets is shown in Table 2.

We use XJTU-SY bearing datasets for early fault diagnosis experiment, and the data do not contain labels that can be identified by the model, so the first step is to determine the time point of failure. Take Bearing2_1 as an example and its vibration data as are shown in Figure 7.

As the normal vibration data and the time-domain feature of rolling bearings satisfies Gaussian distribution, the...
Figure 3: EEMD decomposition results.
Root mean square is used as the feature to calculate the fault threshold in this study. Root mean square (RMS) is calculated in equation (20):

$$RMS_m = \sqrt{\frac{x_1^2 + x_2^2 + \cdots + x_m^2}{m}}$$

Equation (20)

$RMS_m$ represents the root mean square of the first $m$ seconds. $x_m$ represents the data of the first $m$ seconds. Then we calculate the average and variance of the $RMS_m$.

$$\mu_m = \frac{1}{m} \sum_{i=1}^{m} RMS_i,$$

$$\sigma_m = \sqrt{\frac{1}{m} \sum_{i=1}^{m} (RMS_i - \mu_i)^2}.$$ 

Equation (21)

Set the threshold as follows:

$$\text{Threshold}_m = \mu_m + 3\sigma_m.$$ 

Equation (22)
If

$$RMS_j > \text{Threshold}_j$$  \hspace{1cm} (23)

Then it determines that the rolling bearing has failed in the $j$ second.

In order to intuitively understand the process of Pauta criterion, Bearing2_1 dataset is considered as an example in this paper.

Figure 8 shows that Bearing2_1 failed when RMS exceeded the threshold. We repeat the above steps to calculate the threshold of Bearing2_2 and Bearing 2_3, and obtained that Bearing2_1 failed at 28,038 seconds, Bearing2_2 failed at 3204 seconds, Bearing2_3 failed at 19,991 seconds. The first 60 seconds of fault occurrence were selected as early fault samples, 60 seconds of normal data from Bearing2_1, Bearing2_2, and Bearing2_3 were selected as normal samples.

The sampling frequency of the sensor is 32.5 kHz/min, so normal, outer race failure, inner race failure, and cage failure contain 97,500, 32,500, 32,500 and 32,500 continuously sampled data points, which were reconstructed into a 180 * 1024 signal matrix. The method of data preprocessing is the same as Section 3.1, and 70% of them are used as training data, and the rest of them are used as test data. AQDE-SVM, QGA-SVM, DE-SVM, SVM, and CNN have been run independently on the data 10 times, and the accuracy obtained for each time is shown in Figure 9.

According to Figure 9, AQDE-SVM achieves the highest accuracy in all experiments, proving the stability of its performance. For CNN, it is a black-box algorithm whose output is very unstable and in order to improve the stability of CNN, deepening network layers need to be added. Moreover, CNN contains a large number of hyperparameters that need to be set manually which further improves the running time.

The average accuracy, macroaverage, microaverage, and average time of each model are shown in Table 4.

From Table 4, it could be seen that compared with other algorithms, AQDE-SVM has achieved the best performance in all evaluation indicators, and also has significant
advantages in running time. Therefore, the AQDE-SVM proposed in this chapter satisfies the requirements of speed, stability, and accuracy in actual scenes, and thus has higher engineering application value.

4. Conclusion

The novel fault diagnosis model proposed in this review is based on SVM optimized with the quantum DE. Compared with other classical optimization algorithms, the main advantage of AQDE is running speed, thus improving the efficiency of fault diagnosis. Moreover, in order to comprehensively verify the effectiveness of AQDE-SVM, standard fault diagnosis and early fault experiments were set up in this review; accuracy, macroaverage, microaverage, and running time were used for evaluation. The evaluation results show that AQDE-SVM achieve the highest accurate fault diagnosis results in the same running time, indicating that AQDE-SVM is more suitable for the engineering application of fault diagnosis. In our future work, we will expand fault diagnosis research on self-test data.

Data Availability

The CWRU dataset can be downloaded at: https://engineering.case.edu/bearingdatacenter and XJTU-SY dataset can be downloaded at: http://biaowang.te-ch/xjtu-sy-bearing-datasets.

Conflicts of Interest

The authors declare that they have no known competing financial interest or personal relationships that could have appeared to influence the work reported in this paper.

Authors’ Contributions

Conceptualization was done by Y.L. and Z.F. Methodology was developed by Y.L. and H.X. and software was provided by Q.S. Validation was performed by Q.S. and X.L. and formal analysis was performed by W.Y. Investigation was done by, Z.F., resources were provided by Y.L. Data curation was done by Q.S. Original draft was written by Q.S. Reviewing and editing were performed by W.Y and X.L. Visualization was done by W.Y and X.L. Supervision was done by W.Y and X.L. Project administration was done Y.L. and H.X. Funding acquisition was done by H.X. All authors have read and agreed to the published version of the manuscript.

| Algorithm     | Average accuracy (%) | Macro-average (%) | Micro-average (%) | Average time (s) |
|---------------|----------------------|-------------------|-------------------|-----------------|
| AQDE-SVM      | 98.88                | 96.51             | 96.43             | 7.42            |
| QGA-SVM       | 97.41                | 94.71             | 94.64             | 7.48            |
| DE-SVM        | 94.46                | 91.95             | 92.86             | 20.39           |
| SVM           | 86.39                | 84.60             | 87.50             | 0.11            |
| CNN           | 96.59                | 93.68             | 94.61             | 7.45            |

Acknowledgments

The authors thank Shanghai University of Engineering Science and Kunfeng Quantum Technology Co., Ltd for assisting with the measurement setups. This research was funded by the National Key R&D Program of China, grant no. 2020AAA0109300 and Technology Development Fund, grant nos. PKX2020-R17.

References

[1] T. Viet, K. Jaeyoung, K. Ali, and J. M. Kim, "bearing fault diagnosis under variable speed using convolutional neural networks and the stochastic diagonal levenberg-marquardt algorithm," Sensors, vol. 17, no. 12, 2017.
[2] W. Mao, W. Feng, and Y. Liu, "A new deep auto-encoder method with fusing discriminant information for bearing fault diagnosis," Mechanical Systems and Signal Processing, vol. 150, Article ID 107233, 2021.
[3] S. Tang, S. Yuan, and Y. Zhu, "Data preprocessing techniques in convolutional neural network based on fault diagnosis towards rotating machinery," IEEE Access, vol. 8, Article ID 149487, 2020.
[4] G. Xu, M. Liu, Z. Jiang, and D. Soffker, "bearing fault diagnosis method based on deep convolutional neural network and random forest ensemble learning," Sensors, vol. 19, no. 5, 2019.
[5] Y. Sun, A. Xu, K. Wang, H. Xiaoja, and G. Haifeng, "A novel bearing fault diagnosis method based on principal component analysis and BP neural network," in Proceedings of the IEEE International Conference on Electronic Measurement & Instruments (ICEMI), pp. 1125–1131, Changsha, China, November 2019.
[6] X. Xiong, H. Jiang, X. Li, and M. Niu, "A Wasserstein gradient-penalty generative adversarial network with deep auto-encoder for bearing intelligent fault diagnosis," Measurement Science and Technology, vol. 31, no. 4, Article ID 045006, 2020.
[7] S. Wan and X. Zhang, "Bearing fault diagnosis based on teager energy and mean-shift fuzzy C-means," Structural Health Monitoring, vol. 19, no. 14, pp. 1976–1988, 2020.
[8] J. Cervantes, F. Garcia, L. Rodriguez, R. M. Lisbeth, and A. Lopez, “A comprehensive survey on support vector machine classification: applications, challenges and trends,” Neurocomputing, vol. 408, pp. 189–215, 2020.
[9] J. Zhang, J. Zhong, M. Zhong, and J. Zhang, “A Goa-MSVM based strategy to achieve high fault identification accuracy for rotating machinery under different load conditions,” Measurement, vol. 163, Article ID 108067, 2020.
[10] F. Chen, M. Cheng, B. Tang, B. Chen, and W. Xiao, “Pattern recognition of a sensitive feature set based on the orthogonal neighborhood preserving embedding and adaboost SVM algorithm for rolling bearing early fault diagnosis,” Measurement Science and Technology, vol. 31, no. 10, Article ID 105007, 2020.
[11] Z. Huo, Y. Zhang, G. Jombo, and L. Shu, “Adaptive multiscale weighted permutation entropy for rolling bearing fault diagnosis,” *IEEE Access*, vol. 8, Article ID 87529, 2020.

[12] X. Zhu and J. Xiong, "Fault diagnosis of rotation machinery based on support vector machine optimized by quantum genetic algorithm," *IEEE Access*, vol. 6, Article ID 33583, 2018.

[13] Y. Cao, K. Yin, C. Zhou, and B. Ahmed, “Establishment of landslide groundwater level prediction model based on GA-SVM and influencing factor Analysis,” *Sensors*, vol. 20, no. 3, 2020.

[14] M. Pan, C. Li, R. Gao, Y. Huang, and H. You, "Photovoltaic power forecasting based on a support vector machine with improved ant colony optimization," *Journal of Cleaner Production*, vol. 277, Article ID 123948, 2020.

[15] K. Li, J. Luo, Y. Hu, and S. Li, "A novel multi-strategy DE algorithm for parameter optimization in support vector machine," *Journal of Intelligent Information Systems*, vol. 54, no. 3, pp. 527–543, 2020.

[16] R. Storn and K. Price, "Differential Evolution-A simple and efficient adaptive scheme for global optimization over continuous spaces," *Journal of Global Optimization*, vol. 11, no. 4, pp. 341–359, 1997.

[17] K. Han and J. Kim, "On setting the parameters of quantum-inspired evolutionary algorithm," *The 2003 Congress on Evolutionary Computation*, vol. 1, pp. 178–194, 2003.

[18] T. Eitrich and B. Lang, "Efficient optimization of support vector machine learning parameters for unbalanced datasets," *Journal of Computational and Applied Mathematics*, vol. 196, no. 2, pp. 425–436, 2006.

[19] Csegroups, "The case western reserve university bearing data center [EB/OL]," vol. 11, p. 15, 2012, http://csegroups.case.edu/bearingdatacenter/pages/download-data-file.

[20] B. Wang, Y. Lei, N. Li, and N. Li, "A hybrid prognostics approach for estimating remaining useful life of rolling element bearings," *IEEE Transactions on Reliability*, vol. 69, no. 1, pp. 401–412, 2018.

[21] Z. Wu and N. Huang, "Ensemble empirical mode decomposition: a noise assisted data analysis method," *Advances in Adaptive Data Analysis*, vol. 1, no. 1, pp. 1–41, 2009.

[22] N. Huang, Z. Shen, and S. Long, "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis," *Proceedings of Royal Society of London Series A*, vol. 454, pp. 903–995, 1998.

[23] A. Duan, L. Guo, H. Gao, X. Wu, and X. Dong, "Deep focus parallel convolutional neural network for imbalanced classification of machinery fault diagnostics," *IEEE Transactions on Instrumentation and Measurement*, vol. 69, no. 11, pp. 8680–8689, 2020.