

Multiclass Fault Taxonomy in Rolling Bearings at Interpolated and Extrapolated Speeds Based on Time Domain Vibration Data by SVM Algorithms

Purushottam Gangsar · Rajiv Tiwari

Submitted: 10 June 2014 / Published online: 11 November 2014
© ASM International 2014

Abstract The multiclass fault taxonomy of rolling bearings based on vibrations through the support vector machine (SVM) learning technique has been presented in this paper. The main focus of this article is the prediction and taxonomy of bearing faults at the angular speed of measurement as well as innovatively at the interpolated and extrapolated angular speeds. Five different bearing fault conditions, i.e., the inner race fault, outer race fault, bearing element fault, combination of all faults, and a healthy bearing have been considered. Three different statistical feature parameters, namely, the standard deviation, the skewness, and the kurtosis have been obtained from time domain vibration data for bearing fault predictions. The Gaussian RBF kernel and one-against-one multiclass fault classification technique has been used for the taxonomy of bearing fault by the SVM. Also the study of the selection of SVM parameters, like gamma (RBF kernel parameter), best datasets, and the best training and testing percentages have been presented in this paper. The present work observes a near perfect prediction accuracy of the SVM prediction performance when the training and testing are done at a higher rotational speed. It shows a better fault prediction accuracy at the same rotational speed than that of measurement as compared to the interpolated and extrapolated rotational speeds. Also the SVM capability of fault taxonomy decreases with increase in the range of interpolation and extrapolation speeds.

Keywords Rolling bearing · Support vector machine (SVM) · Multi-fault classification · RBF kernel · Interpolation and extrapolation

Introduction

Rolling bearings are crucial machine elements present in a wide field of rotating machine applications. Bearing failures is one of the main causes of breakdowns in such machinery. There have been huge production losses and human casualties due to unexpected bearing failures. The early detection of bearing defects, therefore, is crucial to prevent the system from malfunctions that could cause damages and halt of the entire system. Hence, the fault detection and its taxonomy in rolling bearings is very important [4].

Vibration analysis is one of the most effective methods for the bearing fault diagnosis. Many vibration approaches for the fault diagnosis of rolling bearings have been developed in time, frequency, time-frequency domains with the help of artificial intelligence (AI) methods. In time domain, various statistical parameters, such as the peak, mean, root-mean-square, crest factor, standard deviation, variance, kurtosis, and skewness have been used. In frequency domain, various statistical parameters such as average, center, root-mean-square, and standard deviation of frequency spectrum have been utilized that are obtained from the fast Fourier transform (FFT). Frequency-time domain methods include the short time Fourier transform (STFT), the Wigner–Ville distribution, and the wavelet packet transform [12, 14, 15, 17].

For the automated bearing fault diagnosis, AI methods such as artificial neural network (ANN), immune genetic algorithm (GA), and fuzzy logic have been applied

P. Gangsar · R. Tiwari (✉)
Department of Mechanical Engineering, Indian Institute of
Technology Guwahati, Guwahati, Assam 781 039, India
e-mail: rtiwari@iitg.ernet.in

extensively [8, 19]. These methods are based on the empirical risk minimization (ERM) principle, and there are some common shortcomings, such as relapsing into local minimum easily and slow convergence velocity and over-fitting. Especially, the generalization ability is very low for a limited number of samples. In comparison with these methods, the support vector machine (SVM) can perform significantly well even for a limited number of samples [11].

The SVM has been introduced into rolling bearing fault diagnosis due the fact that it is difficult to obtain enough fault samples in practice and the excellent fault prediction performance of it. Meng et al. [13] introduced the comparison between the RBF neural network and the SVM for cases where only limited training samples were available for the fault diagnosis. The result showed that the SVM had better performance than the RBF neural network both in the training time and the prediction accuracy. The binary-class and multiclass classifier performances of the SVM were discussed, and the diagnosis precision was less dependent on the kernel function and its parameter. Rojas and Nandi [16] proposed the development of SVM for the detection and classification of rolling bearing faults. The training of SVM was done using sequential minimal optimization algorithm, and a mechanism for selecting adequate training parameter was also proposed. Li et al. [11] proposed rolling element bearing fault detection using the SVM with an improved ant colony optimization (IACO). It was used to select a suitable parameter for the SVM. They compared it with the cross-validation and GA methods. Experimental results indicated that the IACO was feasible to optimize parameters for the SVM, and the improved algorithm was effective with the high classification accuracy and in less time.

Sui and Zhang [22] illustrated rolling element bearings fault classification based on the feature evaluation with the SVM. The feature evaluation based on the class separability criterion was discussed in this work. Jack and Nandi [9] illustrated the fault detection of roller bearings using the SVM and the ANN. They defined and estimated statistical features based on moments and cumulants, and selected the optimal features using the GA. In the classification process, they employed the SVM using the RBF kernel with a constant kernel parameter. Zheng and Zhou [26] illustrated the rolling element bearing fault diagnosis based on the SVM. In this paper, firstly, the wavelet packet analysis was utilized to extract the features from the vibration signal and the principal component analysis was performed for features reduction. Secondly, the multiclass SVM as a classifier was employed to diagnose bearing faults. A grid search method in combination with 10-fold cross-validation was applied to find optimal parameters for the multiclass SVM model.

Sugumaran and Ramachandran [21] illustrated the effect of a number of features on the classification of roller bearing faults using the SVM and the proximal support vector machine (PSVM). A set of statistical features and histogram features were extracted from time domain signals and the order of importance was found using a decision tree. It showed the classification efficiency of features, and histogram features were found to be better as compared to statistical features. And the PSVM with the histogram performed marginal better than the SVM. Yang et al. [25] illustrated a fault diagnosis approach for roller bearing based on the intrinsic mode function (IMF) envelope spectrum and the SVM. Liu et al. [10] illustrated the multi-fault classification based on the wavelet SVM (WSVM) with the particle swarm optimization (PSO) algorithm to analyze vibration signals from rolling element bearings. And they showed that the WSVM can achieve a greater accuracy than the SVM. Zhu and Song [27] illustrated the roller bearing fault diagnosis method based on the hierarchical entropy and the SVM with the PSO algorithm.

In this paper, firstly the selection of optimum parameters such as the RBF kernel parameter (γ), the optimum number of datasets, and the optimum percentage of data for the training and the testing have been obtained. After selecting the optimum parameter, i.e., which is very effective for fault classification, it is further used for the multiclass classification of rolling element fault at the same speed, the interpolated speed, and the extrapolated speed as that of measurement based on time domain features by the SVM algorithm using the RBF kernel. Conclusions are drawn from the fault prediction accuracies so obtained.

Support Vector Machine Algorithm

The original SVM algorithm was developed by Vapnik in [24]. The SVM is a machine learning algorithm that analyzes data and recognizes patterns. Initially, it was used for the classification problem but recently they have been extended to the domain of regression analyses [2, 5, 7]. The basic SVM deals with the binary classification. The SVM constructs a hyperplane or a set of hyperplanes in a high or infinite dimensional space, which could be used for the classification and the regression analysis. A good separation is achieved by a hyperplane that has a largest distance to the nearest training datasets of any class. Noticeably, the larger the margin, lower the generalization error of the classifier. Now a set of training data are given to the SVM each marked as belonging to one of two classes, an SVM training algorithm builds a model that predicts whether new data falls into one class or the other.

The SVM is a relatively new machine learning technique based on the structural risk minimization (SRM) principle,

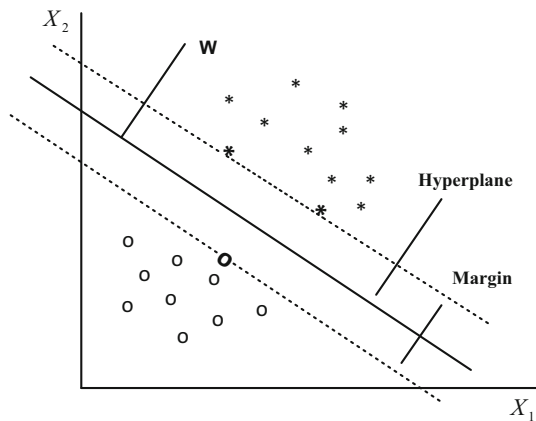


Fig. 1 Data classification by SVM

which is used to solve classification problems by maximizing the margin between two different classes. It can solve the problem of over-fitting, nonlinear, model selection, the curse of dimensionality, and the local minimum in a significant way [23]. It embodies the SRM principle that has achieved higher generalization performance with a small number of samples and is shown superior to the ERM that neural networks use [4, 24]. The SRM minimizes an upper bound on the expected risk, as opposed to ERM that minimizes the error on the training data. This is the reason SVM has greater ability to generalize, which is the aim in statistical learning [13, 17, 18, 20]. The introduction of kernel function is a major advantage of the SVM. It endows with the SVM the ability to deal with nonlinear classification problems by mapping the nonlinear feature space to a high dimensional feature space to solve finally the linear problem. The training of SVM is a convex optimization problem and can always be used to find a global minimum [4].

For finding out the optimal margin classifier, the following optimization problem is defined [1]. The minimization of

$$\tau(w) = \frac{1}{2} \|w\|^2 \quad (\text{Eq 1})$$

subject to

$$y^{(i)} (w^T x^{(i)} + b) \geq 1; \quad i = 1, 2, \dots, m; \quad w^T \in R^N; \\ b \in R, \quad (\text{Eq 2})$$

where the vector w^T (a normal vector to the hyperplane) defines the boundary, x is the input vector of dimension, N and b are scalar thresholds. The constraint could be written as

$$g_i(w) = -y^{(i)} (w^T x^{(i)} + b) + 1 \leq 0. \quad (\text{Eq 3})$$

This optimization problem is a convex quadratic objective function problem with only linear constraints. In Fig. 1,

three dark points lie on the dashed lines parallel to the decision boundary (or the hyperplane). These typical three points are called *support vectors*. When the Lagrangian is constructed for the optimization problem then the following expression is obtained

$$L(w, b, \alpha) = \frac{1}{2} \|w\|^2 - \sum_{i=1}^m \alpha_i [y^{(i)} (w^T x^{(i)} + b) - 1]. \quad (\text{Eq 4})$$

Now the dual optimization problem is defined as

$$\max_{\alpha} W(\alpha) = \sum_{i=1}^m \alpha_i - \frac{1}{2} \sum_{i,j=1}^m y^{(i)} y^{(j)} \alpha_i \alpha_j \langle x^{(i)} x^{(j)} \rangle \quad (\text{Eq 5})$$

subject to

$$\alpha_i \geq 0, \quad i = 1, \dots, m \quad (\text{Eq 6})$$

and

$$\sum_{i=1}^m \alpha_i y^{(i)} = 0. \quad (\text{Eq 7})$$

For solving the nonlinear case, SVMs map the N -dimensional input vector into a higher or Q -dimensional feature space, by choosing a nonlinear mapping function $\phi(x)$, the SVM can construct an optimal hyperplane in new feature space. $k(x, x_i)$ is an inner product kernel performing the nonlinear mapping into feature space and is expressed as

$$k(x, x_i) = k(x_i, x) = \phi(x) \phi(x_i). \quad (\text{Eq 8})$$

The radial basis function is a Gaussian kernel and is expressed as

$$k(x, x_i) = \exp\left(-\gamma \|x - x_i\|^2\right); \quad \gamma = \frac{1}{2\sigma^2} > 0, \quad (\text{Eq 9})$$

where σ denotes the width of RBF kernel. After solving the optimization problem, the SVM classifier is as follows:

$$f(x) = \text{sign} \left\{ \sum_{i=1}^m y_i \alpha_i k(x_i, x) + b \right\}. \quad (\text{Eq 10})$$

The SVM classifiers described above are binary classifiers, and by combining it, it is possible to handle multiclass case. The multiclass SVM classifiers are described below.

Multiclass SVM Classifiers

In the binary classification only two class labels are present, i.e., +1 and -1. But in the real world machinery situation more than two class labels are found, such as the misalignment, mechanical unbalances, shaft cracks, bearing faults, and gear faults. In the bearing fault also several faults appear like the inner race fault, outer race fault, ball element fault, and a combination of these entire faults. For

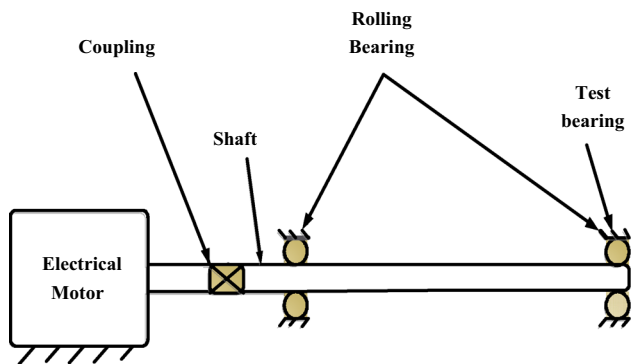


Fig. 2 A schematic line diagram of the MFS



Fig. 3 Machine fault simulator

solving the multiclass problem, methods like one-against-all (OAA), one-against-one (OAO), and direct-acyclic-graph (DAG) are addressed [6]. The earliest used implementation for the SVM multiclass classification is OAA methods. It constructs k SVM models, where k is the number of classes. In OAO methods, if k is the number of classes then $k(k - 1)/2$ classifiers are constructed and each one trains data from two classes. In this method, the training process is similar to the OAO strategy by solving $k(k - 1)/2$ binary SVM. However, in the testing process, it uses a rooted binary DAG, which has $k(k - 1)/2$ internal nodes and k leaves. Hsu and Lin [6] showed a comparison of different multiclass SVM methods and concluded that ‘one-against-one’ is a competitive method for the multiclass classification. Here LIBSVM [3] software has been used for multiclass classification bearing with OAO approach.

Now in the next section, experimental setup and data acquisition are described.

Experimental Setup

Experiments were performed on Machinery Fault Simulator (MFSTM), which was capable of simulating a range of machine faults such as in the gear box, rolling bearings, motors, shaft resonances, and shaft misalignments. The schematic diagram of MFS is shown in Fig. 2. In the MFS experimental setup, a three-phase induction motor was connected to the shaft through a flexible coupling with a healthy bearing near the motor end and a faulty (test) bearing at the other end of the shaft as shown in Fig. 3. This setup allowed the study of bearing defects by introducing faulty bearings in the machine and then studying its vibrational signature. In the study of faults in bearings, five different types of bearing fault condition were considered as shown in Fig. 4 (i) no defect bearing (ND), (ii) inner race fault bearing (IRF), (iii) outer race fault bearing (ORF), (iv) bearing element fault (BEF), and (v) combination fault bearing (CFB). Bearings were fully covered and shielded so close-view of faults could not be shown, separately.



Fig. 4 Rolling bearings with five individual fault conditions

A tri-axial accelerometer (sensitivity: x -axis 100.3 mV/g, y -axis 100.7 mV/g, and z -axis 101.4 mV/g) was mounted on the top surface of the bearing housing, which could record accelerations in time domain in three orthogonal x , y , and z directions. Vibrations were monitored on a Brüel and Kjær FFT analyzer and vibration simulations were processed on the Pulse Lab Shop ver. 7.0.1.298. Datasets were taken for the rotational speed of 10–42.5 Hz in the interval of 2.5 Hz for each of five fault conditions. For each dataset, 320 cycles of data with 2000 samples each were taken. Total 2000×320 data points were collected in each of the three directions x , y , and z .

Statistical Feature Selection

In the present study, statistical moments like the standard deviation, the skewness, and the kurtosis are selected for the effective identification of the bearing fault. These statistical features are briefly explained below.

Standard Deviation (σ)

It is a measure of the energy content in the vibration signal. The standard deviation is the second standardized moment of the data and is expressed as

$$\sigma = \sqrt{\frac{\sum_{i=1}^n (x_i - \mu)^2}{n - 1}}, \quad (\text{Eq 11})$$

where $\{x_1, x_2, \dots, x_n\}$ are the observed values of the sample items and μ is the mean value of these observations, while n stands for the size of the sample.

Skewness (χ)

It is a measure of the asymmetry of the probability distribution of a real-valued random variable about its mean. The skewness value can be positive or negative, or even undefined. It is the third standardized moment of the data and is expressed as

$$\chi = \frac{n}{(n-1)(n-2)} \sum_{i=1}^n \left(\frac{x_i - \mu}{\sigma} \right)^3. \quad (\text{Eq 12})$$

Kurtosis (κ)

It is a measure of the peakedness of the probability distribution of a real-valued random variable. It is the fourth standardized moment of data and is expressed as

$$\kappa = \left\{ \frac{n(n+1)}{(n-1)(n-2)(n-3)} \sum_{i=1}^n \left(\frac{x_i - \mu}{\sigma} \right)^4 \right\} - \frac{3(n-1)^2}{(n-2)(n-3)}. \quad (\text{Eq 13})$$

These three statistical features were calculated from 2000 data points in a sample. Finally, 320 samples were collected in three orthogonal directions for above three features (i.e., total 9 combinations), which means 9×320 feature points were available for each faults. For five different faults (i.e., IRF, ORF, BEF, CFB, and ND) there were total $9 \times 5 \times 320$ feature points available. These total feature points were divided and used for the training, testing of optimizing the parameters, and for the final testing in the simulation.

Results and Discussions

The vibration data were obtained from the machine fault simulator in time domain. After collecting the vibration data, three statistical parameters: the standard deviation, the skewness, and the kurtosis have been extracted in the x , y , and z directions using the MATLABTM. Figure 5 shows

time domain features of acquired signals for the BEF at 10 Hz rotating speed. For the multiclass classification of bearing faults namely, ND, IRF, ORF, BEF, and CFB, the SVM software first needs to be trained with the data comprising these bearing faults, individually. The SVM software consists of various training functions based on different types of kernels. Here, we have used the Gaussian RBF kernel so for this kernel a parameter gamma (γ) is a very important input parameter to be chosen optimally.

The flowchart shown in Fig. 6 summarizes the overall procedure of execution of the training and testing of the SVM algorithm. Now a study on the optimum selection of various SVM and kernel parameters would be performed.

Selection of the Optimum RBF Kernel Parameter

Firstly, we need to select the optimum RBF kernel parameter (γ) for the best prediction accuracy, so arbitrarily we choose rotational speed of 25 Hz, 260 feature datasets, and percentage of training and testing feature data as 80% and 20%, respectively. Later in this work, we will check the optimality of these two parameters also, i.e., the number of feature datasets, and the percentages of the training and testing feature data. Prediction capability performance results are summarized in Table 1 and are shown Fig. 7. The Gaussian RBF kernel parameter (gamma) value of 0.03 gives the highest prediction accuracy at the rotational speed 25 Hz as compared to the other values. For both very less and very higher values of gamma, it gives less prediction accuracy, so for further classification of bearing faults at the same speed, interpolated speed, and extrapolated speed also the value of gamma is chosen as 0.03.

Selection of the Optimum Percentage of Training and Testing Data

Here, we examine which set of the training and testing percentage is the best one for the prediction accuracy. We used the optimum gamma value of 0.03, rotational speed of 25 Hz, and 260 feature datasets. In Figs. 1, 8 represents the training dataset of 65% and the testing dataset of 35%, 2 represents the training dataset of 70% and the testing dataset of 30%, 3 represents the training dataset of 75% and the testing dataset of 25%, 4 represents the training dataset of 80% and the testing dataset of 20%, and 5 represents the training dataset of 85% and the testing dataset of 15%. Prediction performance results are summarized in Table 2 and shown in Fig. 8. It is evident that the training dataset of 80% and the testing dataset of 20%, and also the training dataset of 85% and the testing dataset of 15% give 100% fault prediction accuracy at the rotational speed 25 Hz, so for further classification of bearing fault at same speed, at

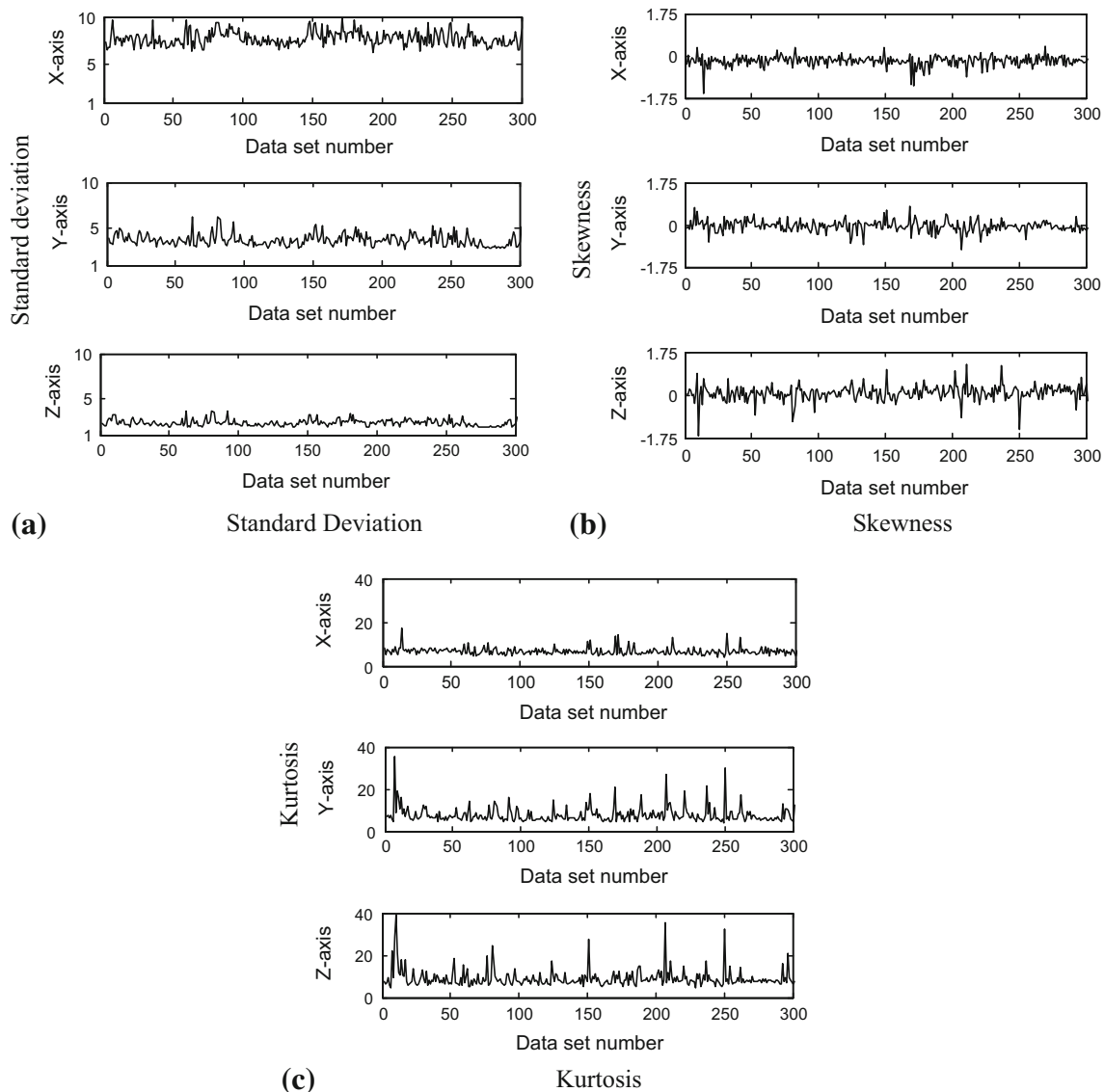


Fig. 5 Time domain features of acquired signals for BEF at 10 Hz rotating speed

interpolated speed, and at extrapolated speed, the training dataset of 80% and testing dataset of 20% have been used.

Selection of the Optimum Number of DataSets

Total 320 datasets were measured for all types of faulty bearing conditions from the MFS. Here the optimum number of datasets has been chosen for the fault prediction accuracy by giving optimum input gamma value of 0.03 as obtained in the above step and for 80% training data and 20% testing data, and also for 75% training data and 25% testing data at a rotational speed of 25 Hz. These prediction results are summarized in Table 3 and are shown in Figs. 9 and 10. The best prediction accuracy occurred for 260 datasets at the rotational speed of 25 Hz than any other number of datasets. Hence, for the further classification of bearing faults at the

same speed, the interpolated speed and the extrapolated speed, 260 datasets have been used. The above selection of parameters is presented for a particular speed, however, similar trends are observed at other speeds also.

Now fault prediction accuracies are obtained for the training and the testing at the *same speed*, *interpolated speed*, and *extrapolated speed* using the optimized parameters obtained above.

Training and Testing at the Same Rotational Speed

For obtaining the prediction accuracy, the training and testing of the SVM is done at the same rotational speed (individually), i.e., 10–40 Hz at the interval of 5 Hz. Table 4 and Fig. 11 show the fault prediction accuracy for the same speed case. It is observed that the percentage of fault

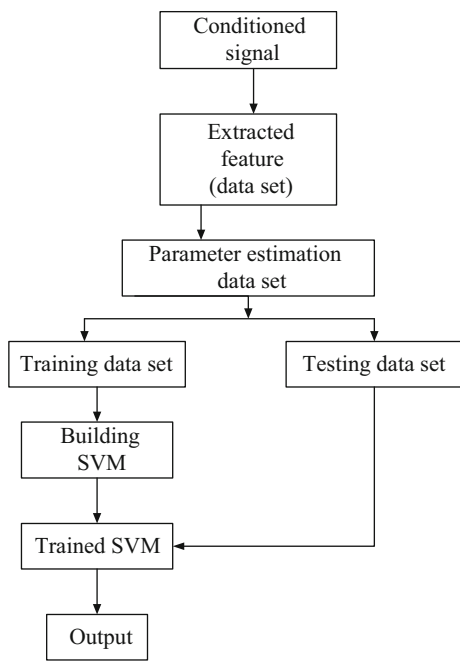


Fig. 6 Flow chart of multiclass fault diagnosis system

Table 1 The optimum selection of the SVM (RBF kernel) parameter gamma

Gamma	ND	IRF	ORF	BEF	CFB
0.003	100	100	100	98.08	100
0.03	100	100	100	100	100
0.3	98.08	98.08	100	98.08	98.08
3	90.38	96.15	90.38	59.62	100
30	98.08	38.46	28.85	0	0
300	0	100	0	0	0
3000	0	100	0	0	0

These bold numbers signify the best prediction accuracy in an individual fault category

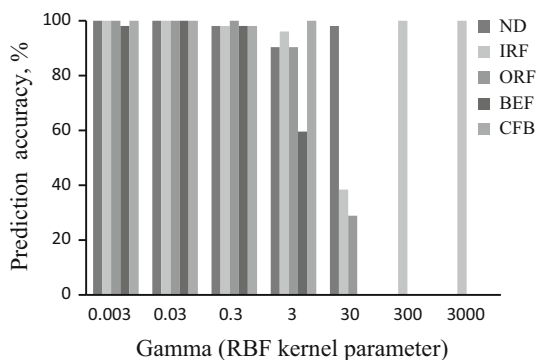


Fig. 7 The selection of the optimum gamma for best fault predictions prediction at 10 Hz rotational speed varies from 76.92% for BEF to 100% for ND and ORF, at 15 Hz rotational speed it varies from 65.38% for BEF to 98.08% for ND and ORF, at

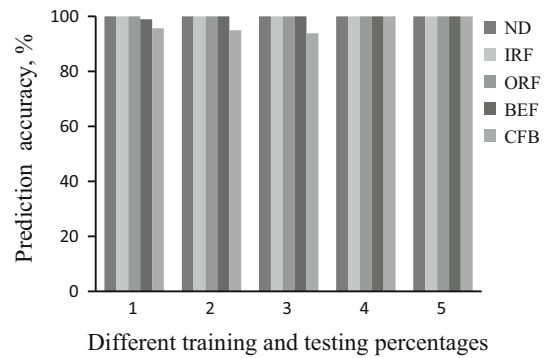


Fig. 8 The selection of the optimum percentages of the training and testing data (1–65%, 35%; 2–70%, 30%; 3–75%, 25%; 4–80%, 20%; 5–85%, 15%)

Table 2 The selection of optimum percentage of the training and testing data

Training % - Testing %	ND	IRF	ORF	BEF	CFB
65%-35%	100	100	100	98.9	95.6
70%-30%	100	100	100	100	94.87
75%-25%	100	100	100	100	93.85
80%-20%	100	100	100	100	100
85%-15%	100	100	100	100	100

These bold numbers signify the best prediction accuracy in an individual fault category

Table 3 The selection of optimum number of feature datasets

Datasets	ND	IRF	ORF	BEF	CFB
<i>Case A: For 75%–25% and for best gamma value 0.03</i>					
100	100	100	100	96	88
120	100	100	100	94.29	85.71
180	100	100	100	93.33	100
220	100	100	100	98.18	94.55
260	100	100	100	100	93.94
320	100	100	100	83.75	97.5
<i>Case B: For 80%–20% and for best gamma value 0.03</i>					
100	100	100	100	100	90
140	100	100	100	96.43	96.43
180	100	100	100	100	100
220	100	100	100	100	90.91
260	100	100	100	100	100
320	100	100	100	79.69	96.88

These bold numbers signify the best prediction accuracy in an individual fault category

20 Hz rotational speed it varies from 69.23% for CFB to 100% for ND, IRF, and ORF, at 25 Hz rotational speed it comes 100% for all, at 30 Hz rotational speed it comes 92.31% for CFB and 100% for all other faults and at 35, 40 Hz rotational speed it comes 100% for all faults.

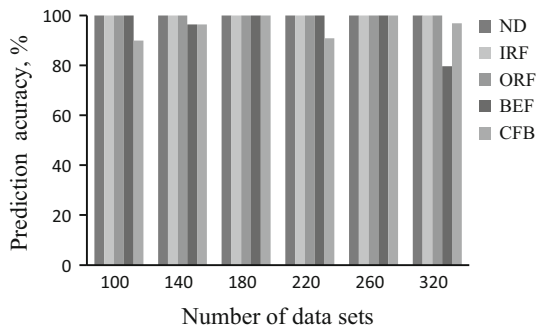


Fig. 9 The selection of the optimum number of feature datasets (for 80% and 20% of the training and testing feature data, respectively)

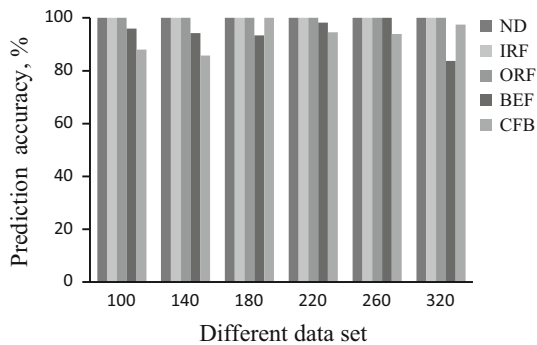


Fig. 10 The selection of the optimum number of feature datasets (for 75% and 25% of the training and testing feature data, respectively)

Table 4 The percentage prediction accuracy for the same speed as measured data

Training speed (Hz)	Testing speed (Hz)	ND	IRF	ORF	BEF	CFB
10	10	100	86.54	100	76.92	78.85
15	15	98.08	92.31	98.08	65.38	94.23
20	20	100	100	100	96.15	69.23
25	25	100	100	100	100	100
30	30	100	100	100	100	92.31
35	35	100	100	100	100	100
40	40	100	100	100	100	100

It is observed a near perfect prediction accuracy at a higher speed beyond 25 Hz since at the higher rotational speed the signal-to-noise ratio is better. This is due to better manifestation of faults at higher speeds. The prediction accuracy of BEF and CBF are lesser than other faults and the healthy bearing. For BEF, this is because it is not necessary that the BEF hit the outer race or inner race in every rotation because of spinning of rolling elements along with the rolling. And this may change the characteristics of vibration signature quite often. The prediction accuracy is very good for the ND and the ORF for higher speeds as well as lower speeds because the vibration signature that comes from no

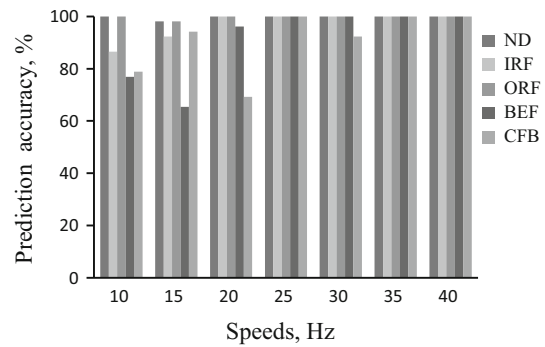


Fig. 11 The fault prediction accuracy at same speed as measurements

Table 5 Interpolation speed ranges

No.	Training speed (Hz)	Testing speed (Hz)
<i>Range 5 Hz</i>		
1	10–15	12.5
2	15–20	17.5
3	20–25	22.5
4	25–30	27.5
5	30–35	32.5
6	35–40	37.5
<i>Range 10 Hz</i>		
1	10–20	15
2	15–25	20
3	20–30	25
4	25–35	30
5	30–40	35
<i>Range 15 Hz</i>		
1	10–25	17.5
2	15–30	22.5
3	20–35	27.5
4	25–40	32.7
<i>Range 20 Hz</i>		
1	10–30	20
2	15–35	25
3	20–40	30

defect bearing is always good. Similarly, rolling elements are always going to hit ORF at every ball pass and it would provide better characteristics in the vibration signature and results in better prediction for the ORF. Similarly, for IRF beyond 10 Hz gives higher prediction accuracy.

Training Done at Two Different Speeds and Testing at an Intermediate Speed

For obtaining the fault prediction accuracy, the training of SVM is done at two different speeds and the testing is done

Table 6 The fault prediction accuracy for different interpolated speeds

Training speeds (Hz)	Testing speed (Hz)	ND	IRF	ORF	BEF	CFB
<i>Range 5 Hz</i>						
10, 15	12.5	23.08	51.92	92.31	50	92.31
15, 20	17.5	65.38	100	100	63.46	65.38
20, 25	22.5	96.15	88.46	100	98.08	98.08
25, 30	27.5	100	100	100	84.62	92.31
30, 35	32.5	96.15	71.15	100	100	92.31
35, 40	37.5	98.08	100	100	100	100
<i>Range 10 Hz</i>						
10, 20	15	1.92	3.85	96.15	75	40.38
15, 25	20	94.23	7.69	100	5.77	88.46
20, 30	25	61.54	21.15	100	98.08	75
25, 35	30	61.54	73.08	100	78.85	94.23
30, 40	35	98.08	100	100	100	100
<i>Range 15 Hz</i>						
10, 25	17.5	13.46	0	69.23	0	25
15, 30	22.5	0	0	90.38	57.69	1.92
20, 35	27.5	1.92	0	100	94.23	94.23
25, 40	32.7	84.62	0	100	57.69	96.15
<i>Range 20 Hz</i>						
10, 30	20	0	0	0	90.38	0
15, 35	25	0	0	100	9.62	90.38
20, 40	30	0	0	100	94.23	57.69

Table 7 Extrapolation speed ranges

No.	Training speed (Hz)	Testing speed (Hz)
<i>Range 5 Hz</i>		
1	10–12.5	15
2	15–17.5	20
3	20–22.5	25
4	25–27.5	30
5	30–32.5	35
6	35–37.5	40
<i>Range 10 Hz</i>		
1	10–15	20
2	15–20	25
3	20–25	30
4	25–30	35
5	30–35	40
<i>Range 15 Hz</i>		
1	10–17.5	25
2	15–22.5	30
3	20–27.5	35
4	25–32.5	40
<i>Range 20 Hz</i>		
1	10–20	30
2	15–25	35
3	20–30	40

Fig. 12 The fault prediction accuracy for different interpolated speeds (for 1, 2, ..., in abscissa refer to Table 5)

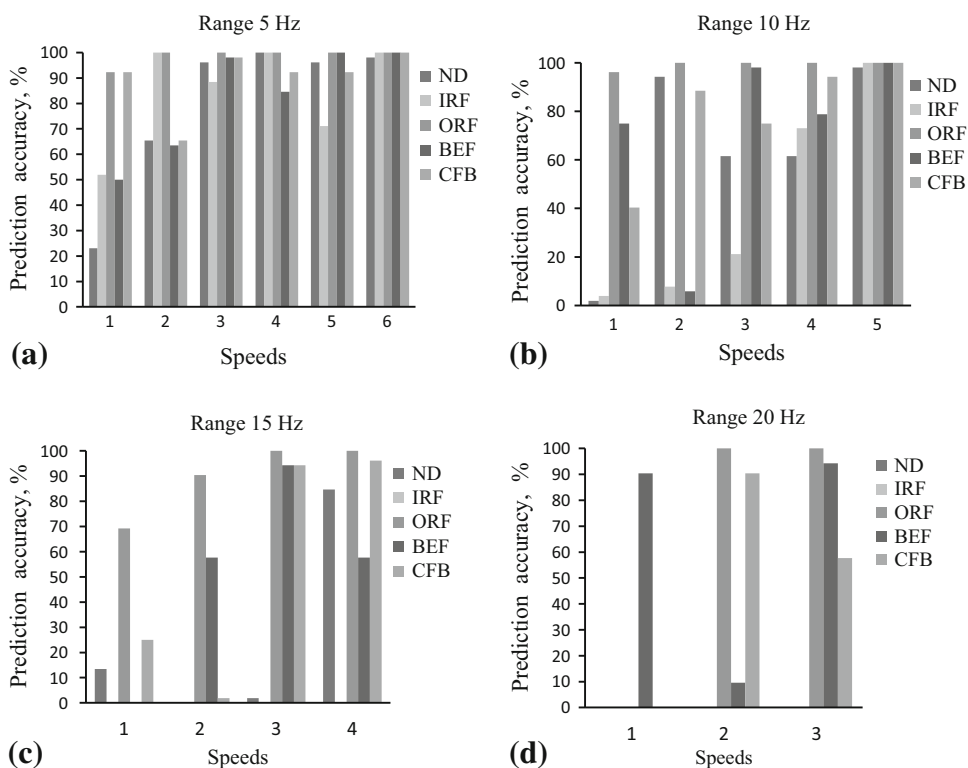


Table 8 The fault prediction accuracy for different extrapolated speeds

Training speeds (Hz)	Testing speed (Hz)	ND	IRF	ORF	BEF	CFB
<i>Range 5 Hz</i>						
10, 12.5	15	5.77	5.77	84.62	65.38	96.15
15, 17.5	20	1.92	1.92	100	63.46	98.08
20, 22.5	25	57.69	23.08	100	65.38	96.15
25, 27.5	30	0	0	100	21.15	100
30, 32.5	35	0	96.15	100	98.08	100
35, 37.5	40	100	98.08	96.15	100	100
<i>Range 10 Hz</i>						
10, 15	20	1.92	3.85	100	1.92	98.08
15, 20	25	61.54	92.31	98.08	5.77	98.08
20, 25	30	0	0	94.23	28.83	100
25, 30	35	0	98.08	100	94.23	100
30, 35	40	53.85	73.08	94.23	100	98.08
<i>Range 15 Hz</i>						
10, 17.5	25	7.69	0	100	0	100
15, 22.5	30	0	0	0	0	100
20, 27.5	35	0	0	67.31	0	100
25, 32.5	40	84.62	0	100	57.69	96.15
<i>Range 20 Hz</i>						
10, 20	30	0	0	0	9.62	25
15, 25	35	0	0	100	9.62	90.38
20, 30	40	0	0	100	94.23	57.69

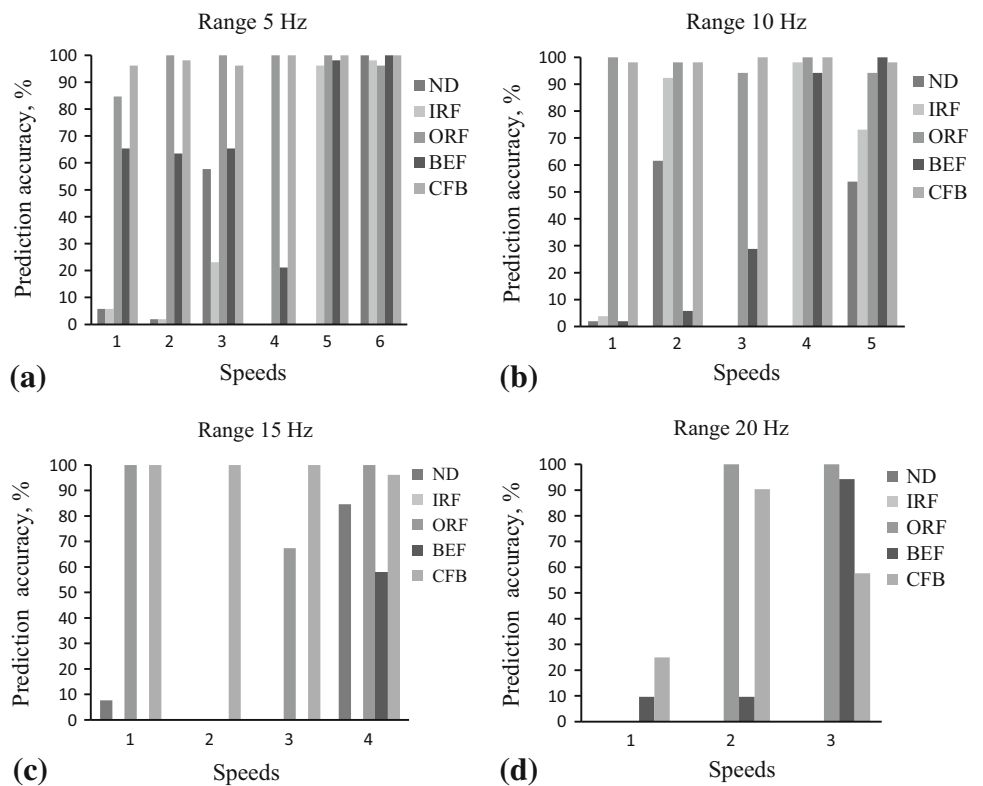
at an intermediate speed for which training of SVM was not done, this is known as the *frequency interpolation*. Here, four different interpolation ranges are considered, i.e., 5, 10, 15, and 20 Hz which is shown in Table 5. In Table 6 and Fig. 12a–d, the fault prediction accuracy is obtained using the frequency interpolation for four different ranges. It is observed that for 5 Hz interpolated speed range, the percentage prediction accuracy varies from 23.08% to 100%, for 10 Hz it varies from 1.92% to 100%, and for 15 and 20 Hz it varies from 0% to 100%. For 5 Hz range the lowest prediction accuracy is 23.08% for the ND, for 10 Hz range the lowest prediction accuracy is 1.92% for the ND, for 15 and 20 Hz range the lowest prediction accuracy is 0% for the ND, IRF, and BEF.

It is examined that the SVM capability of the fault classification decreases with increasing the range of interpolation. It results in very less prediction accuracy beyond 10 Hz interpolation speed range, so the fault prediction is not very useful for the interpolated speed range beyond 10 Hz.

Training Done at Two Different Speeds and Testing Done at an Extrapolated Speed

For obtaining the prediction accuracy, the training of SVM is done at two different speeds and the testing is done at an extrapolated speed for which training of SVM was not

Fig. 13 The fault prediction accuracy for different extrapolated speeds (for 1, 2, ..., in abscissa refer to Table 7)



done, this is known as the *frequency extrapolation*. Here, four different extrapolation ranges are considered, i.e., 5, 10, 15, and 20 Hz, and are shown in Table 7. In Table 8 and Fig. 13e–h, the fault prediction accuracy is obtained using the frequency extrapolation for four different frequency ranges. For the extrapolation, the fault prediction accuracy varies from 0% to 100% for all speed ranges. The lowest prediction accuracy comes 0% for ND and IRF for 5 Hz as well as 10 Hz speed range. The lowest prediction accuracy comes 0% for all other faults except CFB for 15 Hz speed range. The lowest prediction accuracy comes 0% for ND, IRF, and ORF for 20 Hz speed range. It is observed that the SVM capability of fault classification decreases with increasing the range of extrapolation frequency. It results in very less prediction accuracy beyond 10 Hz extrapolation speed range and the fault prediction is not very useful for extrapolated speed beyond 10 Hz.

The prediction accuracy increases with the rotational speed, i.e., at a higher speed beyond 25 Hz the classification of faults are better at the same speed, the interpolated speed, and the extrapolated speed. The worst prediction accuracy 65.38% for the BEF at 15 Hz and 69.23% for the CFB at 20 Hz comes for the same speed case. The interpolation and extrapolation techniques for the multiclass classification of bearing fault are very useful when vibration data are not available for particular speeds but beyond the 10 Hz range interpolation as well as extrapolation techniques are not useful for fault predictions. By this exercise, it is concluded that nearly 100% classification of fault can be achieved by the same speed prediction case, somewhat good for the interpolated speed case and marginal for extrapolated speed case. It is observed that fault classifications by the interpolation technique are better than the extrapolation technique for the same range of speed and at a higher speed as well as lower speed of operation.

Conclusions

In this paper, the fault diagnosis of rolling bearings using the SVM is presented. First, we extracted features from the vibration signal in time domain and then applied *one-versus-one* technique for the multiclass fault classification of rolling bearings. The selection of optimum SVM parameters, such as gamma value (RBF kernel parameter), the percentage of training and testing data, and the number of datasets for further fault classification of bearing faults have been illustrated. Finally, the capability of SVM to classify bearing faults at the same rotational speed, at the interpolated speed, and at the extrapolated speed has been presented. In the present paper, it has been observed that the SVM prediction performance when the training and testing is done at a higher rotational speed a near perfect

prediction accuracy is found. This is because at the higher rotational speed the noise does not affect so much due to better signal-to-noise ratio in the measurement data. And it shows that a better prediction accuracy for the same rotational speed as compared to the frequency interpolation and the frequency extrapolation. Also the SVM capability of fault classification decreases with increase in the range of interpolation and extrapolation frequencies. The interpolation and extrapolation techniques could be applied to measurement data in the frequency and time-frequency (wavelet) domains.

Acknowledgments Authors would like to thank Mr. D. J. Bordoloi Vibration & Acoustic Laboratory, Department of Mechanical Engineering, IIT Guwahati for his timely help during the experimentation. And this research supported by LIBSVM tool (Version 3.1, 2011) is used for the present work, which is freely available online at <http://www.csie.ntu.edu.tw/~cjlin/libsvm>.

References

1. S. Abbasiona, A. Rafsanjania, Rolling element bearings multi-fault classification based on the wavelet denoising and support vector machine. *Mech. Syst. Signal Process.* **21**, 2933–2945 (2007)
2. C.J.C. Burgess, A tutorial on support vector machines for pattern recognition. *Data Min. Knowl. Disc.* **2**, 955–974 (1998)
3. C.C. Chang, C.J. Lin, LIBSVM: A library for support vector machines. *ACM Trans. Intell. Syst. Technol.* **2**, 27:1–27:27 (2011)
4. K.C. Gryllias, I.A. Antoniadis, A Support Vector Machine approach based on physical model training for rolling element bearing fault detection in industrial environments. *Eng. Appl. Artif. Intell.* **25**, 326–344 (2012)
5. S.R. Gunn, *Support vector machines for classification and regression* (Department of Electronics and Computer Science of University of Southampton, Southampton, 1998), pp. 1–28
6. C.W. Hsu, C.J. Lin, A comparison of methods for multi-class support vector machines. *IEEE Trans. Neural. Netw.* **13**(2), 415–425 (2002)
7. C.W. Hsu, C.C Chang, C.-J. Lin, *A Practical Guide to Support Vector Classification* (2003)
8. Y.C. Huang, C.M. Huang, K.Y. Huang, Fuzzy logic applications to power transformer fault diagnosis using dissolved gas analysis. *Proc. Eng.* **50**, 195–200 (2012)
9. L. Jack, A. Nandi, Fault detection using support vector machines and artificial neural networks, augmented by genetic algorithms. *Mech. Syst. Signal Process.* **16**(2–3), 373–390 (2002)
10. Z. Liu, H. Cao, X. Chen, Z. He, Z. Shen, Multi-fault classification based on wavelet SVM with PSO algorithm to analyze vibration signals from rolling element bearings. *Neurocomputing* **99**, 399–410 (2012)
11. X. Li, A. Zheng, X. Zhang, C. Li, L. Zhang, Rolling element bearing fault detection using support vector machine with improved ant colony optimization. *Measurement* **46**, 2726–2734 (2013)
12. K.F. Martin, P. Thrope, Normalized spectra in monitoring of rolling bearing elements. *Wear* **159**, 153–160 (1992)
13. L. Meng, W. Miao, W. Chunguang, Research on SVM classification performance in rolling bearing diagnosis. in *the international conference on intelligent computation technology and automation (ICICTA)*, vol. 3, pp. 132–135, IEEE, 2010

14. N.G. Nikolaou, I.A. Antoniadis, Rolling element bearing fault diagnosis using wavelet packets. *NDT E Int.* **35**(3), 197–205 (2002)
15. H. Prasad, The effect of cage and roller slip on the measured defect frequency response of rolling element bearings. *ASLE Trans.* **30**(3), 360–367 (1987)
16. A. Rojas, A.K. Nandi, Detection and classification of rolling-element bearing faults using support vector machines. in *the IEEE workshop on machine learning for signal processing*, pp. 153–158, 2005
17. B. Samanta, K.R. Al-Balushi, S.A. Al-Araimi, Artificial neural network and support vector machine with genetic algorithm for bearing fault detection. *Eng. Appl. Artif. Intell.* **16**, 657–665 (2003)
18. B. Samantha, K.R. Al-Balushi, Artificial Neural Networks based fault diagnostics of rolling element bearings using time domain features. *Mech. Syst. Signal Process.* **17**, 317–328 (2003)
19. N. Saravanan, K.I. Ramachandran, Incipient gear box fault diagnosis using discrete wavelet transform (DWT) for feature extraction and classification using artificial neural network (ANN). *Expert Syst. Appl.* **37**(6), 4168–4181 (2010)
20. V. Sugumaran, G.R. Sabareesh, K.I. Ramachandran, Fault diagnostics of roller bearing using kernel based neighborhood score multi-class support vector machine. *Expert Syst. Appl.* **34**(4), 3090–3098 (2008)
21. V. Sugumaran, K. Ramachandran, Effect of number of features on classification of roller bearing faults using SVM and PSVM. *Expert Syst. Appl.* **38**(4), 4088–4096 (2011)
22. W.T. Sui, D. Zhang, Rolling element bearings fault classification based on SVM and feature evaluation. in *International Conference on Machine Learning and Cybernetics, IEEE*, vol. 1, pp. 450–453, 2009
23. V. Vapnik, An Overview of Statistical Learning Theory. *IEEE Trans. Neural Netw.* **10**(5), 988–998 (1999)
24. V. Vapnik, V. Levin, E. Le, Y. Cun, Measuring the VC-dimension of a learning machine. *Neural Comput.* **6**, 851–876 (1994)
25. Y. Yang, D. Yu, J. Cheng, A fault diagnosis approach for roller bearing based on IMF envelope spectrum and SVM. *Measurement* **40**(9), 943–950 (2007)
26. H. Zheng, L. Zhou, Rolling element bearing fault diagnosis based on support vector machine. in *2nd International Conference on Consumer Electronics, Communications and Networks (CEC-Net)*, IEEE, pp. 544–547, 2012
27. K. Zhu, X. Song, A roller bearing fault diagnosis method based on hierarchical entropy and support vector machine with particle swarm optimization algorithm. *Measurement* **47**, 67–669 (2014)