Indian Journal of Engineering & Materials Sciences
Vol. 27, August 2020, pp. 878-888

Diagnosis of bearing faults using multi fusion signal processing techniques and mutual information

V. Dave[a], S. Singh[b] & V. Vakharia[a*]

[a]Department of Mechanical Engineering, School of Technology, Pandit Deendayal Petroleum University, Gandhinagar 382 007, Gujarat, India

[b]Machinery Fault Diagnostics Laboratory, Department of Mechanical Engineering
Guru Nanak Dev University, Amritsar, 143005, Punjab, India

Received: 26 May 2020

Bearing is a widely used rotating component in most of the industrial machinery. Failure of bearings can incur substantial losses in the industries. During operation, to prohibit failure in bearing, it becomes necessary to identify faults that occur in bearings. In the present work, bearing vibration signals have been taken for the detection of faults in bearings. In the next step, features obtained from various signal processing techniques such as ensemble empirical mode decomposition (EEMD), walsh hadamard transform (WHT) and discrete wavelet transform (DWT) have been used to detect bearing faults (inner race defect, outer race defect, and ball defects). To select the mother wavelet, the maximum energy to entropy ration criteria has been used. Mutual Information feature ranking algorithm is used to select the relevant features. Machine learning techniques such as Random Forest, Support Vector Machine, Artificial Neural Network, and IBK are used. Training and tenfold cross-validation procedures applied to all ranked features. Results reveal that random forest gives 100 % training accuracy with one ranked feature and 98.43 % ten-fold cross-validation accuracy with seven features. From the results, it is observed that the proposed methodology can be reliable and it may serve as an effective tool for fault diagnosis of bearing.

Keywords: Fault diagnosis, Walsh hadamard transform, Ensemble empirical mode decomposition, Discrete wavelet transform, Support vector machine, Mutual information

1 Introduction

Detection of bearing faults is an essential subject for extensive research. Due to friction, heavy loads and complicated operating conditions, various kinds of faults are developed in bearings. Accurate diagnosis of faults helps to make reasonable maintenance decisions and ensure the safety of machinery. However, diagnosis of bearing faults is still a challenging task for researchers, since dynamic characteristics of bearings are much more complicated. Bearings are attached with the rotating shaft and, further the shaft is attached with motors, and due to operating conditions signals are masked by noise. Vibration signals carry necessary information about the status of bearings and hence are considered as one of the principal tools to diagnose faults in bearing. Various signal processing methods are used to detect faults such as fast fourier transform (FFT), wavelet transform (WT), empirical mode decomposition (EMD), hilbert transform (HT) *etc.*

Wavelet transform is a type of signal processing technique, applied in various applications like fault diagnosis[1], acoustic, motor current and vibration signature analysis of induction motors[2–5], EEG[6], *etc.* Signals acquired from machinery are non-stationary in nature, therefore FFT is not considered as a reliable method. Wavelet transforms emerged as one of the popular time-frequency techniques which is used for non-linear and non-stationary signals. Due to its multi-resolution potential, wavelet transforms is one of the methodologies used in fault identification. The wavelet transform is non- adaptive and better results can be obtained after the selection of wavelet base function. Walsh hadamard transform (WHT) is a signal processing technique that is applied for many applications such as shape-based image retrieval[7], detect faults in pumps[8], bearing fault diagnosis[9,10], detection and segmentation of image[11] *etc.* As compared to other transforms, computation of the Walsh transformis faster because of the matrix obtained through WHT (having values 1 and -1) requiring simpler mathematical operations[12]. EMD is

*Corresponding author (E-mail: vinayvakharia4343@gmail.com)

a modified form of hilbert-huang transform (HHT), which is a popular time-frequency analysis method applicable for machinery components. Intrinsic mode function (IMF) based signal decomposition is the main function of EMD. Mode mixing in EMD[13] has been solved by modified ensemble empirical mode decomposition (EEMD)[14]. In the process of decomposition using signal processing techniques, raw signals are transformed and useful information regarding machinery fault can be extracted after analyzing the statistical features. Once suitable signal processing techniques are identified, feature extraction, feature selection/ranking and feature classification can be done to identify the faults. In the feature extraction process, statistical features such as kurtosis, skewness, crest factor, form factor, *etc.* are calculated and the feature vector is formed. However, all the extracted features may not be appropriate and are redundant, whichin turn reduces the fault diagnosis accuracy, therefore, feature selection also known as a feature ranking technique emerges as a likely tool to correctly diagnose the faults. To enhance the fault detection, feature-ranking method is a popular method in which it is observed that the ranked feature conveys useful information about the signal acquired and are important for enhancing the functioning and exactness of the fault identification system[15]. In an article, genetic algorithm was used for feature selection for correctly identify bearing faults[16]. Feature ranking technique such as Fisher Score and Mahalanobis Distance was used by Wu *et al.*[17] to take into account the relevant feature which helps in improved accuracy to detect bearing faults.Various authors have also used feature ranking in other applications. Recently, Vakharia and Gujar[18] used Relief technique to demonstrate the advantage of feature ranking in regression modelling to predict the Portland cement composition.

To diagnose the bearing faults, signals collected from multiple sensors mounted at several positions of rotor bearing system, are employed by many authors. Safizadeh and Latifi[19] presented a methodology based on data acquired from more than one sensor for detection of faults in bearings. A study, by Zhi *et al.*[20] proposed a fusion model (with data acquired from number of sensors) for condition monitoring and parameter optimization in a grinding operation. Statistical features were extracted by using wavelet decomposition[21] from the acquired load signals and acoustic signals. It was based on the fusion of features extracted from three accelerometers and with varying fault size and shaft rotational speed. The authors mentioned that the exploratory results indicate that the suggested method shows better fault diagnosis as compared to other methods. Even though multi-sensor based fusion techniques are applied by various authors and are reported inthe literature for various applications, the multi fusion signal processing technique to detect bearing faults is reported less. Integrating the features extracted from various signal processing techniques and to identify the relevant features, can strengthen the fault detection abilities. The contribution of proposed methodology is to first combine the features extracted from EEMD, WT, and WHT and to apply Mutual Information Criterion for selecting the exceptional attribute subset. Afterwards, machine learning techniques are compared to recognize the bearing faults. Ten-fold cross-validation is considered as a parameter to assess the fault identification accuracy because of the accurate results obtained. Fig. 1 shows the methodology to diagnose bearing faults using multi fusion signal processing techniques.

2 Signal processing methods

Pre-processing of the acquired signals has been done by using signal processing techniques like EEMD, WT andWHT. Brief descriptions of techniques used are discussed as below:

2.1 Ensemble Empirical Mode Decomposition (EEMD)

EEMD is considered a prevailing technique for unsteady and non-uniform signals. IMFs (Intrinsic Mode Decomposition) are the key part of this method that decomposes the signals. For the proper disintegration of the signal, significant conditions need to be satisfied (1) in the given data set, the maximum and zero-crossing value must be equal or one. (2) The mean value of a given information set is maximum for the local envelope and it is zero for minimum envelope. IMF shows the normal oscillatory mode inherent in indicator as the main function, thus EEMD is a self improved process that can work satisfactorily for nonstationary signals[22]. Following footsteps need to be followed for EEMD implementation:

Step 1: For the n^{th} trial, by adding a white noise timeseries $u_n(t)$ to a given signal *x(t)*, a new time series isgenerated which is represented as:

$$y_n(t) = X(t) + u_n(t) \quad \ldots (1)$$

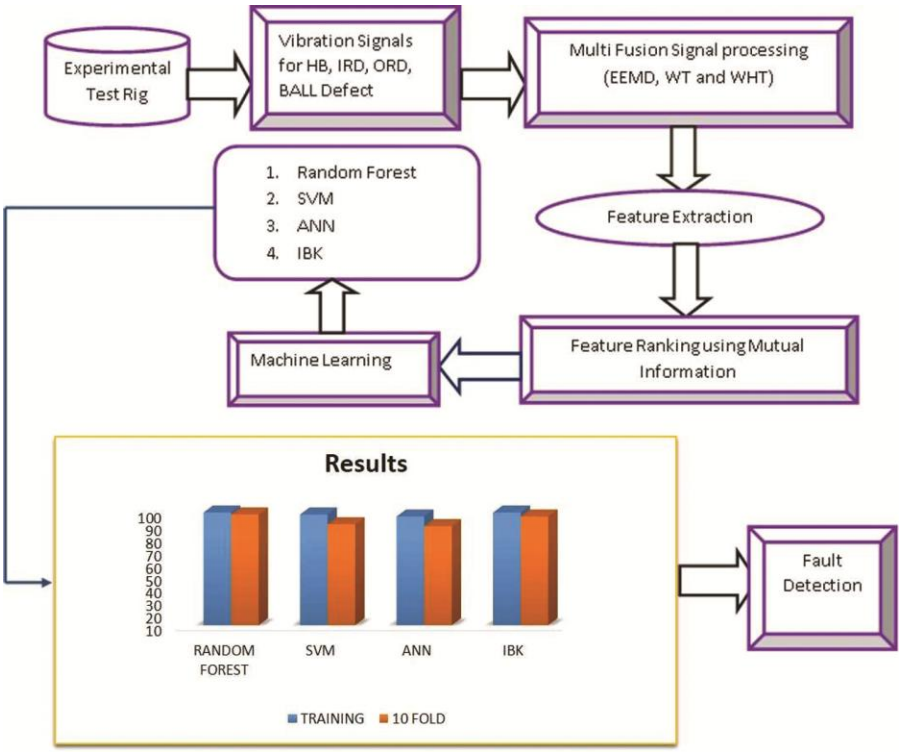


Fig. 1 — Flow chart of the methodology used for bearing fault diagnosis using multi fusion features.

For *n= 1*;*2*... *N*; *N* represents the ensemble number. Step 2: Established on the original EMD, the noise contaminated signal $Y_n(t)$is decomposed into a set of IMFs andresidual.

$$y_n(t) = \sum_{m=1}^{M-1} IMF_m^{(n)}(t) + r_M^{(n)}(t) \quad \ldots (2)$$

where, *M-1* represents the total number of the IMFs derived in each decomposition of $y_n(t)$, $IMF_m^{(n)}$ is the m^{th} IMF and $r_M^{(n)}$ is the residual obtained in the n^{th}trial.

Step 3: The steps (1) and (2) are recapitulate for all trials. In every iteration, a dissimilar white noise series $u_n(t)$ is inserted to the main signal.

Step 4: The concluding IMF from EEMD (IMF_m^{avg}) is gained by the average value of the entirem IMF related to N trials:

$$IMF_m^{avg}(t) = \frac{1}{N}\sum_{n=1}^{N} IMF_m^{(n)}(t) \quad \ldots (3)$$

The outcomes attained by the EEMD rests on the selection of the ensemble number (N) and the added noise (A)'s magnitude. Following relation should be met:

$$\varepsilon = \frac{A}{\sqrt{N}} \quad \ldots (4)$$

ε= final standard deviation error, which is the difference of the original signal from the addition of IMFs ensuing from the EEMD and ensemble number, *N* = 100 is set.

2.2 Discrete Wavelet Transform (DWT)

Wavelets provide time-scale information of a signal, aiding the abstraction of non-linear features. This property formulates the wavelets as an excellent medium for analyzing signals of a transient or non-stationary nature.DWT can be formulated as:

$$DWT\ (j,k) = \frac{1}{\sqrt{2^j}}\int_{-\infty}^{\infty} x(t)\psi^*\left(\frac{t-2^j k}{2^j}\right) \quad \ldots (5)$$

Here ψ^* reflects the complex conjugate of the scaled and shifted wavelet function. When signal x(t) passes through the wavelet filters, it is decomposed in to low and high frequency components as

$$\begin{cases} a_{j,k} = \sum_m h(2k-m)a_{j-1,m} \\ d_{j,k} = \sum_m g(2k-m)a_{j-1,m} \end{cases} \quad \ldots (6)$$

In Eq. (6) $a_{j,k}$ is the approximate coefficients, giving small amplitude frequency components of the acquired signal, and $d_{j,k}$ describes detail coefficient, which conform to signal's high frequency components.

2.3 Walsh Hadamard Transform (WHT)

The implementation of Walsh functions needed to executeon the signal depends on the fourier transform. Walshfunction is a vital parameter for the implementation of WHT, where the value of the Hadamard matrix depends on mathematical operations like addition and subtraction. The problems observed from discrete fourier transform (DFT), like distortion and leakage of frequencies information can be solved by WHT. The Hadamard transform can be defined by

$$\mathrm{H}_1 = \frac{1}{\sqrt{2}}\begin{pmatrix} H_{m-1} & H_{m-1} \\ H_{m-1} & -H_{m-1} \end{pmatrix} \quad \ldots (7)$$

where the $\frac{1}{\sqrt{2}}$= normalization factor (value depends on applications). Variations of Hadamard matrices are as follows:

$$H_0 = +1 \quad \ldots (8)$$

$$H_1 = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \quad \ldots (9)$$

3 Fault Identification Techniques

Machine learning techniques are widely categorized as classification and regression. For classification, labels are predicted whereas in regression numerical values are predicted. To perform classification and regression, the feature vector is needed. Since bearing fault identification requires labels *i.e.* fault in the inner race, faults in outer race, fault in rolling element, therefore classification algorithms such as support vector machine (SVM), artificial neural network (ANN), random forest (RF) and IBk were used in present work.

3.1 Support Vector Machine (SVM)

It is a frequently applied method for classification and regression in fault diagnosis since the dataset is small. For a linear and nonlinear dataset, a hyperplane is formed in such a way that splits the data centred on margin. Adjoining data points used to describe margin are called support vectors. A hyperplane is formed, which separates the feature vector belonging to various fault conditions, based on optimization equations. (refer Fig. 2)

$$\mathrm{Min}\frac{1}{2}\|w\|^2 + C\sum_{i=1}^{K}\xi_i \quad \ldots (11)$$

$$\text{Subjected to}\{\ y_i(w^T x_i + b) \geq 1 - \xi_i \quad \ldots (12)$$
$$\xi_i \geq 0, i = 1,2, \ldots K$$

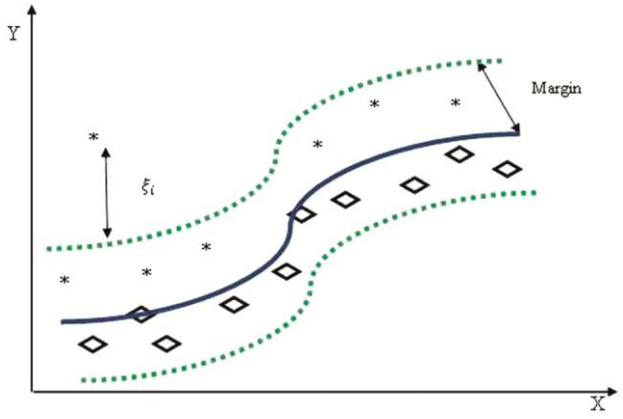


Fig. 2 — Nonlinear SVM classification.

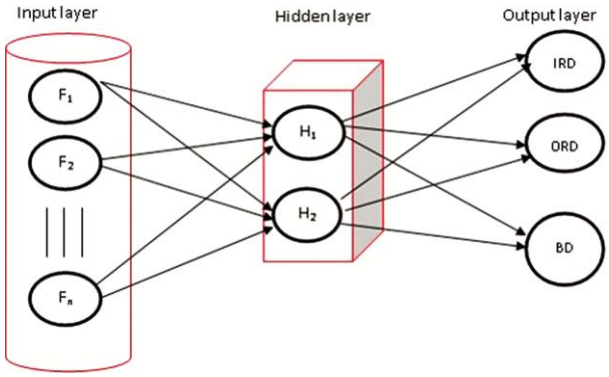


Fig. 3 — ANN architecture for classification.

where, C is a constant which represents error penalty and $\xi_i$represents slack variable in Fig. 2.

3.2 Artificial Neural Network (ANN)

The artificial neural network is considered as an analogy of the human brain and is a popular machine learning method. It is considered as an excellent tool to determine the patterns in feature set which helps in differentiating various fault conditions. An input layer, hidden layer, and output layer comprise ANN which is linked by a node. A weighted sum of every single input is computed, which is constructed by the kind of function used and the obtained worth is progressed to the next layer and the process recurring for all the inputs[24]. As seen from the Fig. 3, ANN neuron is an important component that is present in the biological brains and signals can be transmitted from one neurons to another. The association amongst the artificial neurons are called 'edges'. The groupings of artificial neurons and edges have a weight which fine-tunes as the learning ensues. One input layer (f), hidden layer (H)and output layer are present in feed forward neural networkwith twenty-seven input and four output with radial basis function(RBF).

3.3 Random Forest

Random forest is a type of ensemble learning algorithm which uses a tree known as CART (Classification and Regression Tree) for prediction. Prediction of fault cases is made with the help of Bagging or Bootstrap Aggregating which consists of random subsets of the feature vector. Instead of searching greedily, it randomly samples features from the data set and allows several instances to be used repeatedly in the training stage and the process is repeated continuously until the final prediction[25]. Figure 4 shows the methodology of random forest, when it is used for fault diagnosis of bearings.

3.4 Instance Based Learner (IBk)

In machine learning, instance-based learning is a popular algorithm that is used for both classification and regression. In the instance-based learning algorithms when applied to fault diagnosis, distances or similarity between various instances are computed and finally, predictions are made. K-nearest neighbour algorithm is an example of the instance-based learner in which predictions are obtained for an instance x after searching through the entire feature vector for the K most similar instances which represent the neighbours and gives results for all the instances. Authors used Euclidean distance to determine the fault diagnosis accuracy.

4 Experimental data

The experiments were conducted in a test facility[25] in which bearings are fitted on fan end and drive end of motors. The bearings are attached to the housings and accelerometers are attached to record the vibration signals under various fault conditions. The inner race of bearings is attached with shaft and the rotational speed varied as 1730, 1750, 1772 and 1797 rpm with sampling frequency 12 kHz. Defects of varying size are created individually in the inner race, outer race and ball with the following configurations:

a) Inner race defect (IRD) with diameter of 0.1778 mm, 0.3556 mm and0.5334 mm.
b) Outer race defect (ORD) with diameter of 0.1778 mm, 0.3556 mm and 0.5334 mm.
c) Ball defect (BD) with diameter of 0.1778 mm, 0.3556 mm and 0.5334 mm.

Drive end vibration signals had been chosen to conduct the study due to the availability of a large number of signals against the fan end vibration signals. 6205-2RSL JEM SKF deep groove ball bearing used at both ends whose overall dimensions are given in Table 1. The arrangement of the testing rig is shown in Fig. 5.

In total 64 samples belonging to IRD, ORD, BD and healthy cases with various fault sizes and variations in the rotational speed of the shaft have been used. To correctly identify the fault identification nine features namely RMS, Standard deviation, Skewness, Kurtosis, Form factor, Peak to peak value, Crest factor, RSSQ and Mean are extracted from three signal processing techniques EEMD, WT, and WHT.The feature vector formed consists of 64 fault cases and 27 features. To extract features from EEMD, the selection of IMFs is required. The authors calculated the energy and cross-correlation for all the sixteen modes considered. Table 2 shows the sample cross-correlation value of 16 IMFs. It is observed that for healthy bearing, the

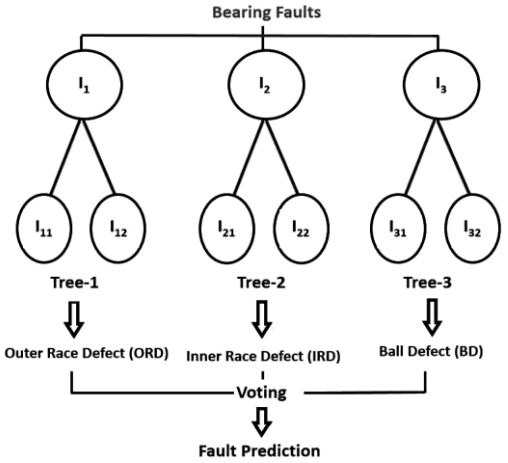


Fig. 4 — Random forest classification.

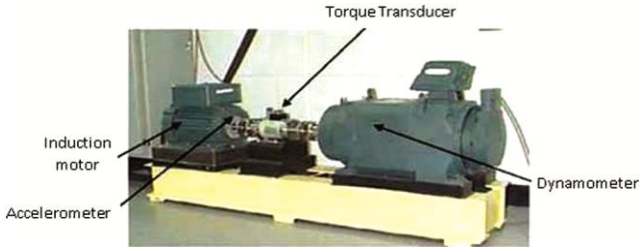


Fig. 5 — Layout of bearing test rig.

Table 1 — Bearing specifications.

Bearing type	6205(SKF) Deep groove ball bearing
Outer race diameter (mm)	52
Inner race diameter (mm)	25
Ball diameter (mm) (d)	7.94
Bearing pitch diameter (mm) (D)	39
Ball number	9
Contact angle	0°

Table 2 — Cross-correlation value of IMFs for all the four classes of bearing.

IMFs	HB	IRD	ORD	BD
1	0.71	0.88	0.99	0.89
2	0.78	0.42	0.09	0.33
3	0.36	0.26	0.05	0.28
4	0.38	0.13	0.02	0.17
5	0.34	0.04	0.01	0.10
6	0.19	0.03	0.00	0.03
7	0.12	0.00	0.00	0.00
8	0.01	0.00	0.00	0.00
9	0.00	0.00	0.00	0.00
10	0.00	0.00	0.00	0.00
11	0.00	0.00	0.00	0.00
12	0.01	0.00	0.00	0.00
13	0.01	0.00	0.00	0.00
14	0.01	0.00	0.00	0.00
15	0.01	0.00	0.00	0.00
16	0.00	0.00	0.00	0.00

Table 3 — IMFs selection using EEMD.

Bearings	IMFs	Energy	Cross-correlation
HB	2	347.41	0.78
HB	2	478.72	0.64
IR	1	7979.11	0.88
IR	1	4134.98	0.95
IR	1	3278.73	0.92
IR	1	27750.80	0.98
IR	1	44531.60	0.75
IR	1	53911.80	0.99
OR	1	71729.70	0.99
OR	1	6718.77	0.98
OR	1	1020.96	0.89
OR	1	37349.50	0.95
OR	1	5904.22	0.71
OR	1	4571.54	0.70
OR	1	12929.30	0.95
BD	1	2138.01	0.95
BD	1	2380.83	0.92
BD	1	1602.31	0.86
BD	1	1149.85	0.89
BD	1	1177.61	0.83

values of second cross correlation are maximum and for the remaining three faulty conditions, the first value of cross-correlation is maximum.So for healthy bearing conditions, features are extracted from IMF 2 and for other fault conditions, IMF 1 is considered for feature extractions. Table 3 shows the maximum value of energy and cross correlation from the random sample vibration signals. To extract the features from the DWT we have chosen Coiflet wavelet at level 1 with detail coefficients. More detail about the feature extraction process can be referred from[26].

Figures (6–9) show time domain and FFT plot for bearing condition at 1772 RPM. Figure 6 refers to the healthy bearing which corresponds to varying compliance frequency (87Hz). Fig. 7 shows the ball pass frequency at inner race (159.95 Hz), similarly, Fig. 8 represent ball pass frequency at outer race (105.4 Hz) and Fig. 9 represent two times ball pass frequency (69.52 Hz) which represent faults in rolling element.

5 Results and Discussion

To verify the effectiveness of methodology proposed, training and ten-fold cross-validation of fusion features by classifiers: Random Forest, SVM, ANN, and IBk have been used. In the initial stage, all fused features are required for training. Once the training of classifiers is done then ten-fold cross-validation is performed to assess the effectiveness of extracted features for fault identifications. In the process of ten-fold cross-validation, the fusion features vector is partitioned into ten equal-sized folds and then ten iterations are completed. One of the ten-fold is used for testing and the rest nine-fold are used for training. The feature vector formed consists of 64 instances and 27 features. To identify the relevant features, authors performed feature ranking using Mutual Information and the rankedfeature are shown in Table 4.

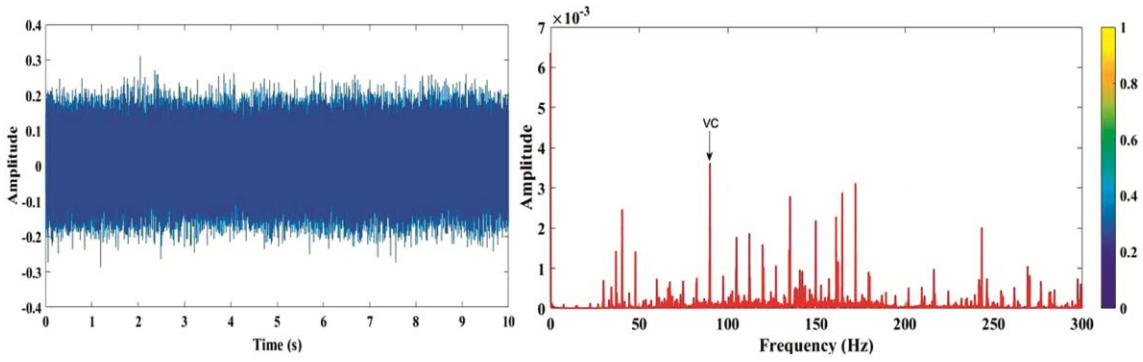


Fig. 6 — Time domain and FFT plot for healthy bearing at 1772 RPM.

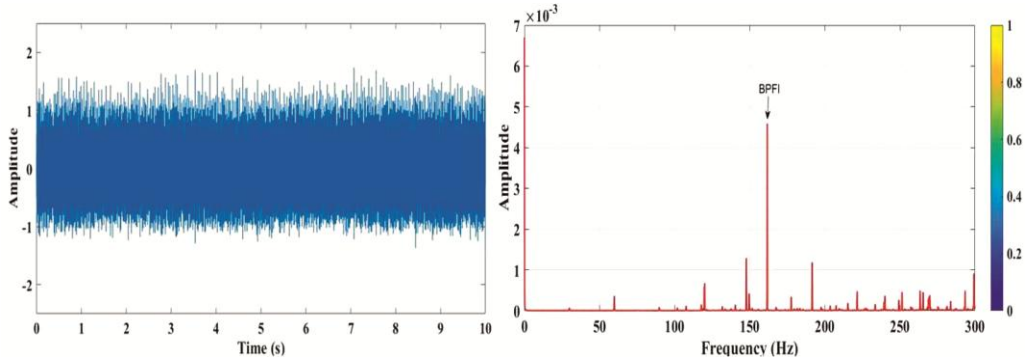


Fig. 7 — Time domain and FFT plot for bearing with inner race defect at 1772 RPM.

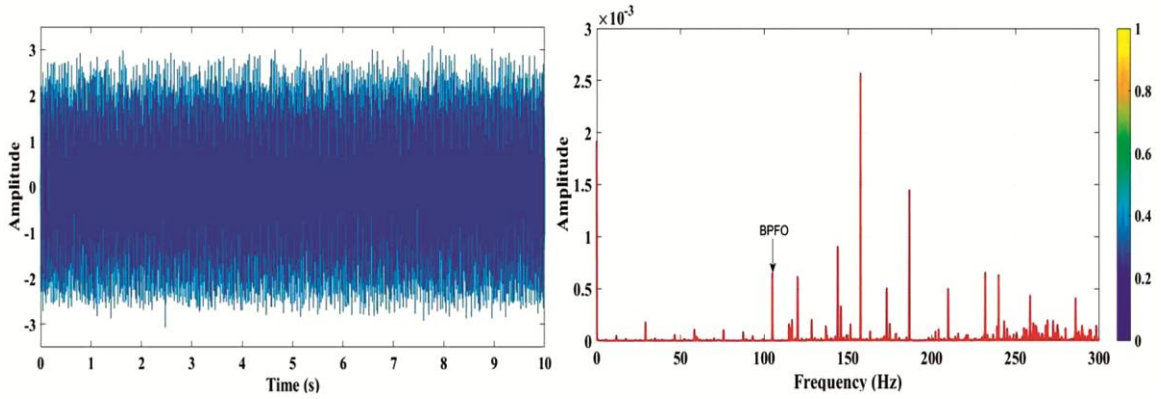


Fig. 8 — Time domain and FFT plot for bearing with outer race defect at 1772 RPM.

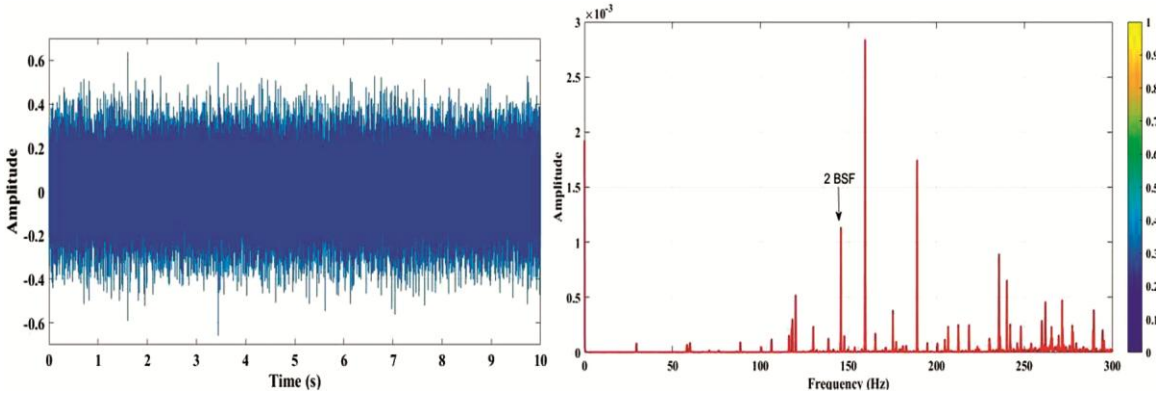


Fig. 9 — Time domain and FFT plot for bearing with ball defect at 1772 RPM.

It is observed that the RMS feature extracted from the WHT technique is the most significant feature, out of all 27 features followed by Standard deviation extracted from WHT, Skewness extracted from DWT and so on. These 27 features are used for fault identification of bearing faults using machine learning techniques. Figures (10 & 11) show the fault identification accuracy when training and ten-fold cross-validation is implemented on the machine learning algorithms. As observed from Fig. 10, Random Forest gives 100 % accuracy to identify all the bearing faults when the only one ranked feature *i.e.* RMS extracted from WHT is used. SVM can identify bearing faults with a maximum 98.43 % accuracy, with fifteen ranked features. ANN identifies faults with a maximum 96.87 % accuracy, with seventeen ranked features. IBK comparatively gives better accuracy of 100 % with only four ranked features as competed to SVM and ANN respectively. Thus, from Fig. 10, it can be inferred that Random

Table 4 — Feature ranked using MI.

Feature ranking	Feature name	Value
1	RMS (WHT)	5.45
2	Standard deviation(WHT)	5.45
3	Skewness (DWT)	5.21
4	Form factor (DWT)	5.20
5	Peak to peak (WHT)	5.19
6	Crest factor (DWT)	5.14
7	Crest factor(EEMD)	5.12
8	Peak to peak (EEMD)	4.99
9	Kurtosis(EEMD)	4.94
10	Kurtosis (DWT)	4.85
11	Root sum of square(WHT)	4.84
12	Crest factor(WHT)	4.69
13	Skewness(EEMD)	4.65
14	Average(WHT)	4.65
15	RMS (EEMD)	4.65
16	Standard deviation (EEMD)	4.65
17	Form factor(WHT)	4.58
18	Average(EEMD)	4.57
19	Root sum of square (EEMD)	4.50
20	RMS (DWT)	4.49
21	Root sum of square (DWT)	4.39
22	Peak to peak (DWT)	4.38
23	Standard deviation (DWT)	4.06
24	Form factor(EEMD)	3.99
25	Skewness(WHT)	3.91
26	Kurtosis(WHT)	3.82
27	Average(DWT)	3.61

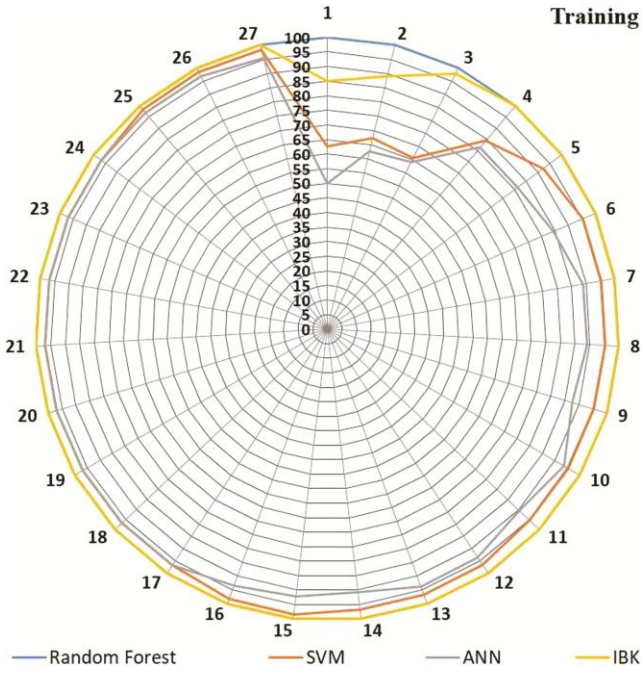


Fig. 10 — Fault identification accuracy based on training.

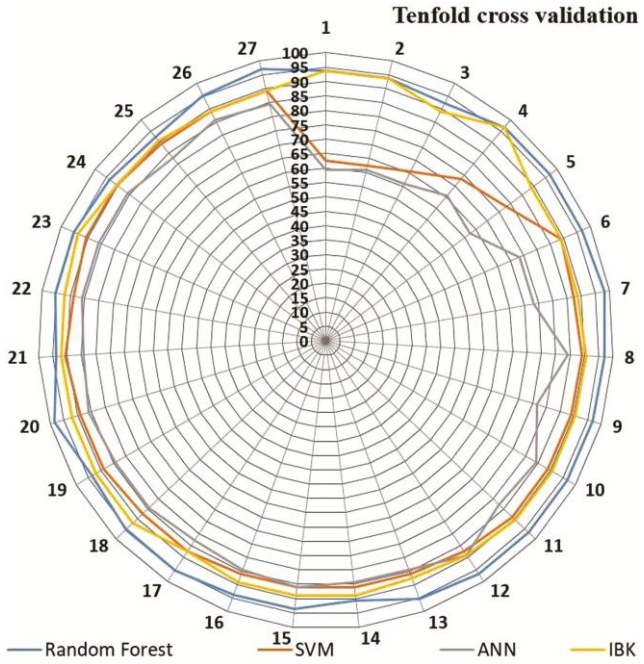


Fig. 11 — Fault identification accuracy based on ten-fold cross validation.

Forest is the best classifier to identify bearing faults with Mutual Information based ranking of multi fusion features extracted from EEMD,WHT, and DWT when training is done. To confirm the generality of the projected methodology, ten-fold cross-validation results are shown in Fig. 11.

It is observed that the accuracy of fusion ranked features increases as the numbers of features are added. The maximum ten-fold cross-validation accuracy achieved is 98.43 % with seven ranked features when Random Forest is used, whereas the maximum achievable ten-fold cross validation accuracy with IBk is 96.87 % with just four ranked features. However, the accuracy to identify bearing faults decreases when SVM and ANN are used for faultidentification giving a maximum of 90.62 % and 89.06 % with twenty-one and twelve ranked features, as observed from Fig.11. For the training the lowest accuracy observed is 50 % with ANN and only one ranked feature followed by 62.5 % with SVM, one ranked feature and 85 % with IBk, one ranked feature respectively. When cross-validation is performed then the minimum accuracy observed is 59.3 %, 62.5 % and 87.5 % with ANN, SVM, and IBkrespectively with one, one and thirteen features respectively. The minimum accuracy observed with Random Forest is 92.18 % with three ranked features. Training and tenfold cross-validation accuracy with all the four

classifiers are mentioned in Table 5. To assess accuracy and effectiveness of the algorithm when types of faults are more; confusion matrix is one of the effective techniques, which summarizes the performance of a classification algorithm to detect the type of fault. In a confusion matrix actual fault and predicted fault arerepresented by row and column respectively. Table 6 shows the confusion matrix for all the four classifiers considered in present study. It is observed that Random Forest givessuperior fault identification accuracy for both training and tenfold-cross validation as compared to SVM, ANN and IBk. It is able to detect all faults when training is performed whereas in case of tenfold-cross validation it correctly identifies HB, IRD and ORD. Table 7 shows the maximum training and tenfold-cross validation accuracy with different cases of signal

Table 5 — Training and tenfold cross-validation accuracy for all twenty seven features.

Feature rank	Random forest		SVM		ANN		IBk	
	Training	Tenfold	Training	Tenfold	Training	Tenfold	Training	Tenfold
1	100	93.75	62.5	62.5	50	59.37	100	93.75
2	100	93.75	67.18	62.5	62.5	60.93	100	93.75
3	100	92.18	65.62	65.62	64.06	60.93	100	89.06
4	100	96.87	84.37	73.43	81.25	65.62	100	96.87
5	100	96.87	92.18	78.125	81.25	62.5	100	89.06
6	100	96.87	95.31	89.06	84.37	73.43	100	89.06
7	100	98.43	95.31	87.5	89.06	73.43	100	89.06
8	100	96.87	95.31	89.06	89.06	84.37	100	90.62
9	100	96.87	95.31	89.06	87.5	76.56	100	90.62
10	100	96.87	95.31	89.03	93.75	84.37	100	90.62
11	100	96.87	95.31	89.06	90.62	82.81	100	90.62
12	100	96.87	96.87	87.5	93.75	89.06	100	89.06
13	100	95.31	96.87	85.93	93.75	84.37	100	87.5
14	100	90.62	96.87	85.93	90.62	84.37	100	89.06
15	100	93.75	98.43	85.93	92.18	85.93	100	89.06
16	100	93.75	98.43	85.93	93.43	84.37	100	89.06
17	100	95.31	96.87	87.5	96.87	82.81	100	87.5
18	100	95.31	96.87	87.5	96.87	84.37	100	92.18
19	100	93.75	96.87	89.06	96.87	84.37	100	92.18
20	100	98.43	96.87	89.06	96.87	85.93	100	92.18
21	100	93.75	96.87	90.62	96.87	84.37	100	92.18
22	100	95.31	96.87	89.06	96.87	85.93	100	92.18
23	100	95.31	96.87	90.62	96.87	85.93	100	93.75
24	100	93.75	96.87	90.62	96.87	85.93	100	90.62
25	100	92.18	98.43	89.06	96.87	82.81	100	90.62
26	100	95.31	98.43	89.06	96.87	85.93	100	89.06
27	100	96.87	98.43	89.06	95.31	84.37	100	89.06

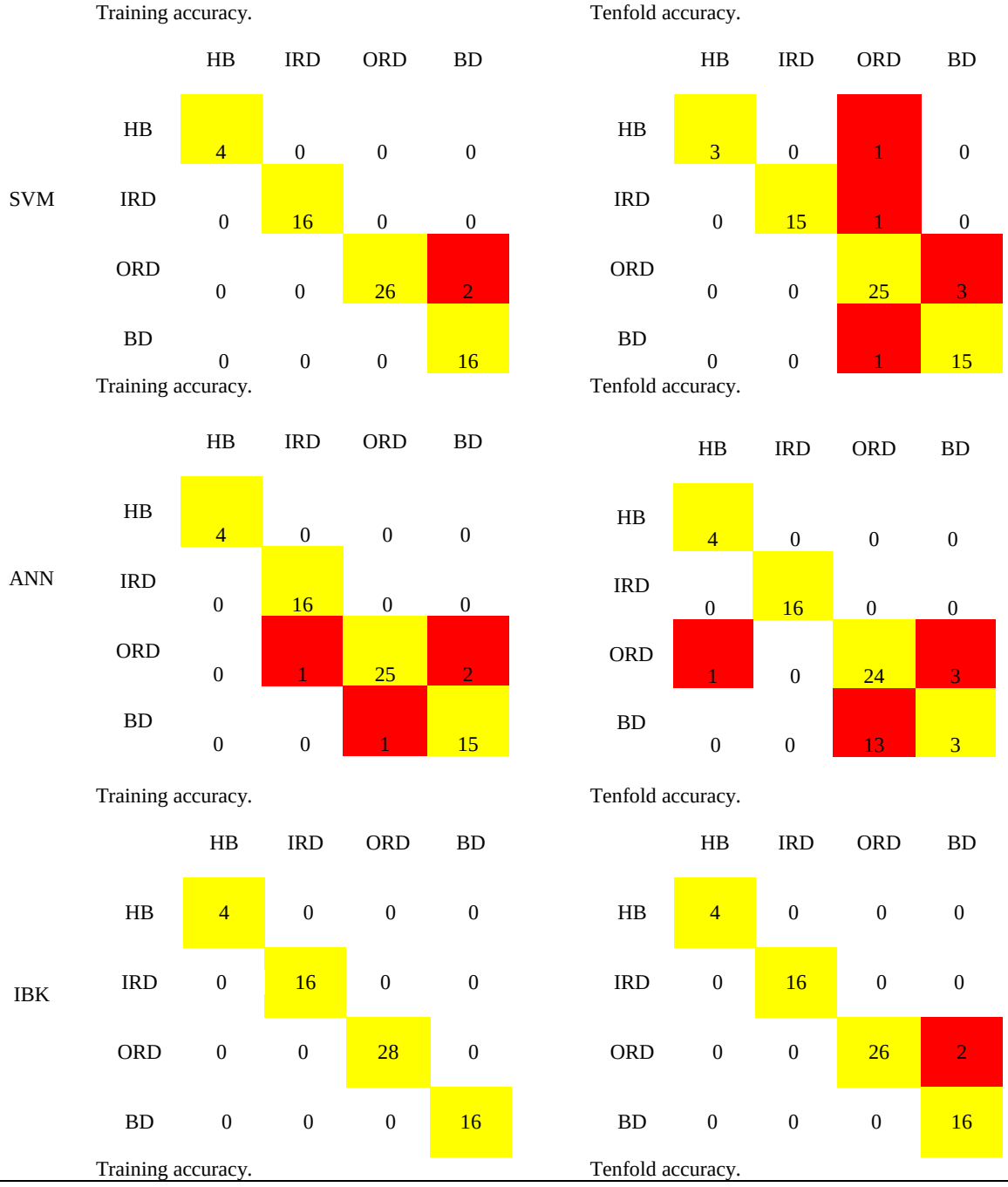
Table 6 — Confusion matrix for all four classifiers.

		HB	IRD	ORD	BD
Random forest	HB	4	0	0	0
	IRD	0	16	0	0
	ORD	0	0	28	0
	BD	0	0	0	16

		HB	IRD	ORD	BD
	HB	4	0	0	0
	IRD	0	16	0	0
	ORD	0	0	28	0
	BD	0	0	1	15

(*Contd*.)

Table 6 — Confusion matrix for all four classifiers. (*Contd*.)



processing techniques used. Considering ranked fusion features, the maximum training (100 %) and tenfold (98.43%) fault identification accuracy obtained from randomforest classifier. Considering ranked first six EEMD features, ANN gives least fault identification accuracy *i.e.* 92.18% and 78.12 % with both training and tenfold cross validation respectively, where as random forest gives 100% training accuracy and 93.75 % cross-validation accuracy. As observed from Table 7, WHT features gives bettertenfold cross-validation accuracy of 93.75 % with only two features whereas DWT features gives 93.75 % tenfold cross validation accuracy with nine ranked features.

Table 7 — Maximum training and tenfold cross validation accuracy.

Signal processing techniques	Random forest		SVM		ANN		IBK	
	Training	Tenfold	Training	Tenfold	Training	Tenfold	Training	Tenfold
Fusion of all three methods	100	98.43	98.43	90.62	96.87	89.06	100	96.87
Individual EEMD (6 features)	100	93.75	95.31	93.75	92.18	78.12	100	93.75
Individual WHT (2 features)	100	93.75	70.31	67.18	75.0	65.62	100	93.75
Individual DWT (9 features)	100	93.75	95.31	87.5	84.37	65.62	100	92.18

6 Conclusions

In the present work, the authors utilized three signal processing techniques and extracted features are used for fault identification. To choose the essential features, Mutual information is used as feature ranking methods. Four machine learning algorithms are used for comparison and decision making about the appropriate algorithm which correctly identifies the various bearing faults. Following observations are concluded:

(i) For identifying bearing faults accurately, the fusion of the features and ranking of features are most effective. Random forest gives a maximum 98 % ten-fold validation accuracy with only seven ranked features.

(ii) Maximum training and ten-fold validation accuracy obtained when the fusions of features extracted are used as compared to EEMD, WHT and DWT are used individually for fault identification.

(iii) After comparison of results, Random forest provides better bearing fault identification accuracy for all the fault cases considered.

(iv) Based on confusion matrix, classifiers are able to detect healthy bearings and defects in inner race effectively, as compared to other faults.

Acknowledgement

Authors would like to thank Professor Loparo and Case Western Reserve University, Cleveland, Ohio, the USA for providing bearing data set, which was used to carry out this study. Authors also acknowledge the support of PDPU Gandhinagar for providing necessary infrastructure to conduct the study.

References

1 Gangsar P & Tiwari R, *J Brazilian Soc Mech Sci Eng*, 41 (2019) 71.
2 Singh S & Kumar N, *IEEE Transac Industrial Informs*, 13(3) (2016) 1341.
3 Glowacz A & Glowacz Z, *Appl Acoust*, 117 (2017) 20.
4 Glowacz A, *Mech Sys Signal Process*, 133 (2019) 106226.
5 Glowacz A, Glowacz W, Kozik J, Piech K, Gutten M, Caesarendra W, Liu H, Brumeric F, Irfan M & Khan F Z, *Measure Sci Rev*, 19(6) (2019) 241.
6 Bhattacharyya A, Pachori R B, Upadhyay A & Acharya UA, *Appl Sci*, 7(4) (2017) 385.
7 Varanis M & Pederiva R, *J Brazilian Soc Mech Sci Eng*, 40(2) (2018) 98.
8 Banerjee A & Dutta A, *Procedia Technol*, 10 (2013) 623.
9 Gao Q, Tang H, Xiang J, Zhong Y, Ye S & Pang J, *Measure*, 134 (2019) 293.
10 Xiang X, Zhou J, Li C, Li Q & Luo Z, *Mech Syst Signal Process*, 23(4) (2009) 1313.
11 Dave V K, Vakharia V & Singh S, *Reliability, Safety and Hazard Assessment for Risk-Based Technologies: Proc ICRESH*, Springer, ISBN: 978-981-13-9007-4, (2020), p. 607.
12 Wang X, Liang X, Zheng J & Zhou H, *Signal Process: Image Comm*, 70 (2019) 47.
13 Tang B, Dong S & Song T, *Signal Process*, 92(1) (2012) 248.
14 Wu Z & Huang N E, *Adv Adaptive Data Analysis*, 1 (2009) 1.
15 Xiang X, Zhou J, An X, Peng B & Yang J, *Mech Sys Signal Process*, 22 (7) (2008) 1685.
16 Vakharia V, Gupta V & Kankar P, *Soft Computing*, 20(4) (2016) 1601.
17 Samanta B, Al-Balushi KR & Al-Araimi SA, *Soft Computing*, 10(3) (2006) 264.
18 Vakharia V & Gujar R, *Construc Build Mater*, 225 (2019) 292.
19 Safizadeh M & Latifi S, *Inform Fusion*, 18 (2014) 1.
20 Lu ZJ, Xiang Q & Xu L, *Procedia CIRP*, 17 (2014) 721.
21 Chen Z & Li W, *IEEE Transac Instrumen Measure*, 66(7) (2017) 1693.
22 Lei Y, He Z & Zi Y, *Mech Sys Signal Process*, 23(4) (2009) 1327.
23 Gujar R & Vakharia V, *Construc Build Mater*, 207 (2019) 519.
24 Breiman L, *Mach Learn*, 45(1) (2001) 5.
25 Loparo K,. *https://csegroups.case.edu/bearingdatacenter/pages/download-data-file*, (2020).
26 Vakharia V, Gupta V & Kankar P, *J Brazilian Soc Mech Sci Eng*, 39(8) (2017) 2969.