Research Article

Vibration Analysis of Shaft Misalignment Using Machine Learning Approach under Variable Load Conditions

A. M. Umbrajkaar [ID],[1] A. Krishnamoorthy,[1] and R. B. Dhumale[2]

[1]*Department of Mechanical Engineering, Sathyabama Institute of Science and Technology, Chennai, Tamilnadu, India*
[2]*Department of Electronics and Telecommunication, Sinhgad College of Engineering, Pune, Maharashtra, India*

Correspondence should be addressed to A. M. Umbrajkaar; amit.umbrajkar@dypiemr.ac.in

The Industry 4.0 revolution is insisting strongly for use of machine learning-based processes and condition monitoring. In this paper, emphasis is given on machine learning-based approach for condition monitoring of shaft misalignment. This work highlights combined approach of artificial neural network and support vector machine for identification and measure of shaft misalignment. The measure of misalignment requires more features to be extracted under variable load conditions. Hence, primary objective is to measure misalignment with a minimum number of extracted features. This is achieved through normalization of vibration signal. An experimental setup is prepared to collect the required vibration signals. The normalized time domain nonstationary signals are given to discrete wavelet transform for features extraction. The extracted features such as detailed coefficient is considered for feature selection viz. Skewness, Kurtosis, Max, Min, Root mean square, and Entropy. The ReliefF algorithm is used to decide best feature on rank basis. The ratio of maximum energy to Shannon entropy is used in wavelet selection. The best feature is used to train machine learning algorithm. The rank-based feature selection has improved classification accuracy of support vector machine. The result obtained with the combined approach are discussed for different misalignment conditions.

1. Introduction

All production and processing industries have been using rotary machines on a major scale. In order to ensure hassle-free operating conditions and fewer maintenance costs, it is essential to monitor machine health condition effectively. Eventually, machines get to adhere to the faulty conditions in due course of time due to various inherent causes. It is seen that out of many listed causes, misalignment is one of the prominent causes of fault set up. Hence, to avoid such faults, continuous monitoring is essential. Vishwakarma et al. [1] have discussed different modes of condition monitoring techniques. It emphasizes the importance of both time and frequency domain analysis for nonstationary signals. Tang et al. [2] have proposed an adaptive waveform decomposition method of the waveform to extract time-frequency features of nonstationary signals. The feature extraction for vibration signal of rolling bearing is carried out with the Adaptive Waveform Decomposition (AWD) algorithm and local frequency concept. The Adaptive Neuro Fuzzy Inference System (ANFIS) architecture has been implemented effectively for simulation of nonlinear components in online control systems. The effective implementation of ANFIS has reduced nondimensional error-index and minimized adjustable parameters very less than other methods such as cascade correlation neural network and backpropagation neural network [3]. Li et al. have explained [4] the importance of online condition monitoring and diagnosis of power equipment. A brief review of a transformer, gas-insulated switchgear, cable, generator, and capacitor are described with the help of big data, Internet of things, and cloud computing techniques. The wavelet gray moment vector approach is claimed as an effective tool in fault diagnosis of rotating machinery [5]. The detailed fault

classification method [6] is based on the wavelet packet energy ratio of resampled vibration signals. In comparison of wavelet transform with other methods such as variational mode decomposition and empirical mode decomposition, the evaluation of the upper as well as lower envelops is one of the main steps in calculations. For that reason, the error developed in envelop estimation will spread in the recursive decomposition results [7]. The analysis is carried out at a multilevel energy ratio to extract fault features of the vibration signal. Sohn and Farrar [8] have presented time series analysis for fault source in mechanical systems. A two-stage model combined with autoregression and exogenous input technique is used for damage location. Dhumale and Lokhande [9] have presented fault diagnosis of voltage source inverter. The author has used extracted features from normalized current signals effectively to train Artificial Neural Network (ANN).

The fault diagnosis system under variable load condition is developed for diagnosis of the voltage source inverter. Wilson Wang has proposed an extended Neuro Fuzzy system for real-time machinery condition monitoring. The developed monitoring system has been validated with experimental results and confirmed adaptability for different fault conditions [10]. Liu and Wang have explained fault diagnosis analysis for low-speed and heavy load slewing bearing. Different options viz. vibration, temperature, oil, stress, etc. are discussed and compared with innovative current analysis [11]. Shao et al. [12] have adopted deep wavelet autoencoder to handle unsupervised feature learning and developed an intelligent fault diagnosis system for rolling bearing. A multiple wavelet autoencoder is used to improve unsupervised feature learning ability. Tonks and Wang [13] have presented a combined approach of Artificial Neural Network (ANN) and Support Vector Machine (SVM) for fault identification in radial distribution systems. The principal component analysis technique is used for data analysis and faults are classified in combination with support vector classifiers. The change in thermometric condition in combination with SCADA system have been effectively used by Tonks and Wang [13] for detection of angular and offset misalignment in wind turbine shaft. The change in a condition of misalignment has been mapped with the changing temperature of the system. In this case, fault isolation is very essential to correlate misalignment with a change in temperature. The effect of torsional longitudinal vibration on aligned condition [14] has been studied through the simple lumped mass model and results are verified experimentally. A required coupling stiffness coefficient for reducing torsional vibration has been discussed. Acoustic emission technique [15] is used over conventional vibration analysis to detect angular misalignment of the shaft. The change in sound condition at support bearing is considered as a source of input. Especially at remote locations, the misalignment present is detected accurately with a combined approach of thermograph, i.e., thermal imaging and vibration analysis. This is claimed as an effective technique in which elevated positions are considered for measurement such as windmill gearbox [16]. A theoretical analysis of the combined effect of shaft misalignment and unbalance is presented in the first part of the paper. Experimental validation is carried out to support the claims in conclusion [17].

The fuzzy-based controlling of current to avoid nonlinear load drawbacks has been explained. The compensating currents are injected with the help of static current distribution compensator [18]. Singh et al. have explained helical gearbox fault diagnosis using wavelet theory and J48 algorithm. The maximum accuracy in feature extraction is claimed by using SYM8 wavelet [19]. Patra and Bruzzone [20] have explored combined advantage of self-organizing map neural network and support vector machine to select uncertain and diverse samples in image classification. The effective combination of feature selection and feature extraction technique with SVM [21] is used for the prediction of defective software modules. The correlation-based feature selection technique with SVM has been compared with other available techniques to prove accuracy claimed in results. The vibration signal obtained is normalized [22] for effective feature selection. The discrete wavelet transform is applied for suitable feature extraction. The discrete wavelet transform and fuzzy logic have been used in combination to predict shaft misalignment. In the review paper, Hsu and Lin [23] presented compared several methods of SVM and showed sample cases which are one-against-one (OAO) as a best suitable for practical use. The mathematical analysis part of binary and multiclass SVM with an example of disease classification is explained [24].

An integrated approach of data mining and machine learning method is proposed for classification of type of damage condition in wood poles [25]. The advantage of nonlinear mapping of SVM along with enhanced cat swarn optimization is used to predict compressive strength of high performance concrete [26]. The nonlinear behavior of magnetorheological elastomer base isolator is optimized based on artificial neural network and ant colony algorithm [27]. The artificial neural network is proposed for accurate estimation of modulus of elasticity by considering effect of Alkali-Silica reaction in concrete [28].

The statistical methods such as the fuzzy system are much better by formulating rules for handling ambiguity and defining the relationship between input and output. If there is no ambiguity in the information collected and since the data is labeled, there is no need to use fuzzy systems or unsupervised machine learning algorithm such as K-NN [29]. ANN is a nonlinear model that is easy to use and understand as a simple statistical method [30]. Most of the statistical methods are parametric models that require a high background of statistics; ANN is a nonparametric model. It cannot define the relationship between input and output and cannot deal with uncertainty. To overcome this, a number of approaches have been combined with ANN to select features, and so on [31]. Deep learning requires a large amount of data and needs to be trained in complex data models, which can be very expensive. It requires huge datasets to train.

In an overview of the literature study, the feature extraction-based condition monitoring technique is discussed for various faults other than misalignment [1–8]. The various fault analysis technique based on vibration, temperature,

stress analysis, and some of these in combination with ANN are considered for fault identification. In many cases of mechanical fault analysis, measurement of fault is essential to understand severity of fault generated. The major of faulty severity is not focused cases such as unbalance, misalignment, and crack analysis [9–16]. The SVM-based fault classification is used in various fields viz. static current distribution compensator, image processing, software defects, and decease classification. The SVM-based fault classification has been exercised in the domain of mechanical engineering [17–23, 32]. The on-field application and use of ANN-SVM approach for various engineering domain has been explained [24–27]. There are different artificial intelligence algorithms for fault classification. The selection depends on the type of problem and size of data available [28–31, 33].

The fault prediction and measure of fault severity, both are essential parts of run time condition monitoring. The present work focuses on the classification and measure of shaft misalignment under variable load conditions using combined approach of ANN and SVM. The normalized vibration signals are used for feature extraction. The suitable mother wavelet is selected on the basis of maximum Energy to Shannon Entropy (ESE) ratio. The ReliefF algorithm is used for best feature selection. The selected best features are used to train SVM for classification of misalignment as well as an input for ANN for measure of misalignment. The suitable structure for ANN is selected out of several trained structures based on the accuracy of prediction. The novelty of proposed **Classification and Prediction of Shaft Misalignment (CPSM)** is to classify the type of misalignment and to measure misalignment under variable speed conditions with minimum number of features and least data size for training. This is achieved by normalization of vibration signals before feature extraction. The results obtained show that accuracy of SVM and ANN classifier has been improved due to rank-based feature selection.

2. Methodology

The proposed CPSM is implemented to output signals obtained for healthy and faulty condition, and observations are recorded for all conditions. The outline of test rig used is shown in Figure 1. An accelerometer is placed at the casing of second bearing to sense vibration in all three directions viz. Longitudinal (V_g), Lateral (V_t), and Vertical (V_r). The misalignment is generated artificially in set up to visualize a proportional change in *Overall Vibration Level (OVL)*. The wide range of vibration levels is observed for a different range of misalignment and speed conditions. These output signals obtained are normalized in the range of 0 to 1. The normalization of the signal maintains distinctive values of extracted features under varying load conditions without loss of information. The normalized signals viz. V_{gNn}, V_{tNn}, and V_{rNn} are obtained from [22]

$$V_{jNn} = \frac{V_{jn}}{\max\left(V_j\right)}, \tag{1}$$

where j is a direction vector which represents g, r and t three direction. The output vibration signals are recorded in these directions.

The vibration signals are normalized and features such as Detailed coefficient (DC) and Average coefficient (AC) are extracted. Normalization reduces data size in the training of classifier and helps to improve accuracy. The selection of appropriate mother wavelet is carried out on the basis of DC. The DC and AC are obtained from

$$DC_{(BN)} = \sum_{p=0}^{q} V_{BN}[u] \times h(p - u). \tag{2}$$

$$AC_{(BN)} = \sum_{p=0}^{q} V_{BN}[u] \times g(p - u). \tag{3}$$

From equations (2) and (3), h and g are filtered coefficients, p is the number of samples, and u is shifting parameter. The extracted features are used for feature selection viz. Maximum (Max), Minimum (Min), Skewness, Kurtosis, Rms, and Entropy. The selected DC feature is revealed better with change in OVL for different conditions of misalignment. Hence, in **CPSM**, selection of correct mother wavelet is carried out on the basis of DC.

3. Experiment Facilities and Instrumentation

In the training of ANN and SVM, a large amount of real-time data with the actual misaligned condition is the foremost important part. It is obtained from the experimental setup. Figure 2 shows a pictorial view of the experimental setup. It comprises a motor, coupling, base plate, and two bearings. The vibration isolation pads at the base of a heavy foundation plate are used to isolate vibration from other sources. In the setup preparation, a major focus is projected on the actual induction of parallel and angular misalignment. The proper directional slot at the base of motor and base plate interface facilitates an easy induction of offset and parallel misalignment for experimental purpose on the artificial mode. It is very important to ensure zero misaligned states of an experimental setup in static conditions before carrying out the experiments. A fixture with special consideration has been prepared to verify zero misalignments. This fixture facilitates the use of the Face and Rim method to serve the purpose of checking alignment conditions. In this, the fixture is clamped on motor side coupling and the dial is simultaneously mounted on the face and rim part of rotor side coupling. The face dial calibrates angular alignment deviation and rim dial calibrates offset alignment deviation simultaneously in reference to motor side coupling. A variable frequency drive (VFD) is used to run setup at different operating speeds. The range of motor speed is closely considered with a standard rated speed of industrial motor selection.

In the implementation of CPSM, two sets of observations are recorded. One with varying speeds and constant misalignment and others with varying misalignment and constant speed. The few samples collected at 1200 rpm for
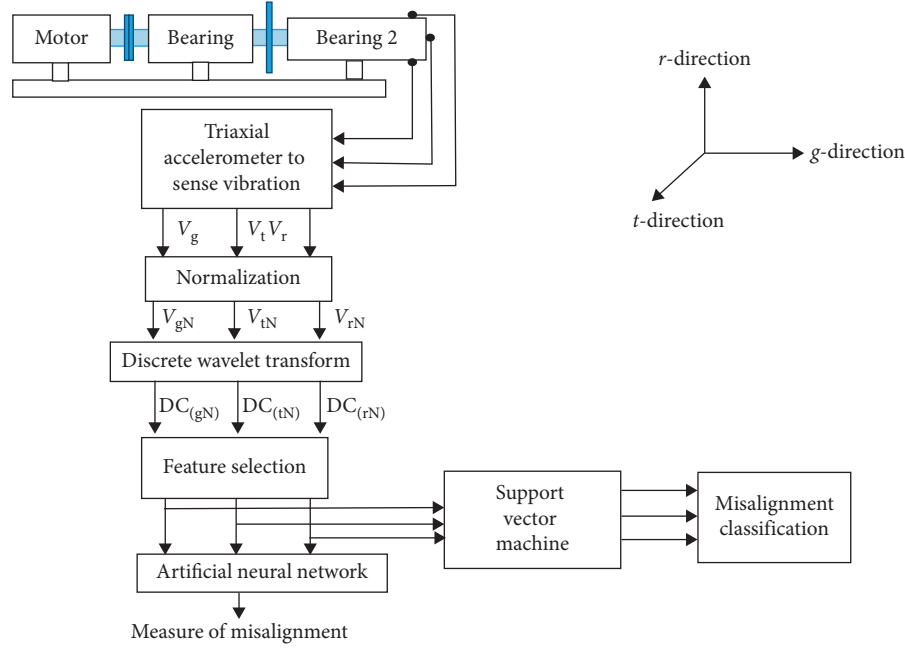
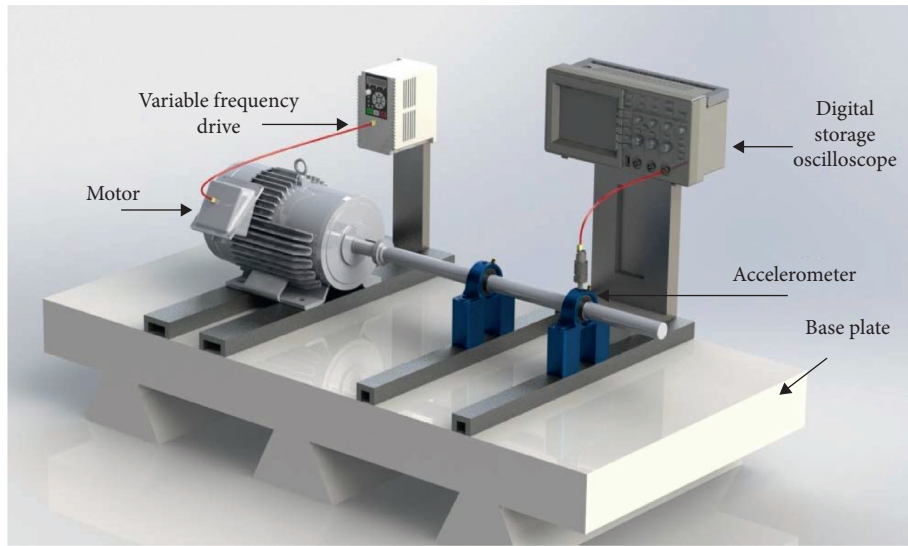Figure 1: SVM- and ANN-based misalignment prediction.

Figure 2: Experimental setup.

different misalignment conditions are presented in Figures 3–5. The variation of OVL with respect to change in misalignment and speed is shown in Figure 6. In order to prevent any unbalance or runout problem, a shaft along with a rotor was tested on a dynamic balancing machine before its assembly. The vibration signals are collected at bearing. An accelerometer (PCB make, Model: 352B70, measurement range: ±49000 m/s², frequency: 0.4 to 20 kHz) is used for sensing vibration signals in three directions. The directional slots are provided at the base plate, which enables to introduce offset and angular misalignment in the setup on artificial mode. The misalignment is introduced with a step of 0.02 mm, i.e., 7.88 mils for the entire range of experiments. The digital storage oscilloscope (DSO) (Tektronix

make TBS 1064, 60 MHz, 4 channels, measurement accuracy: vertical ±3%, from 10 mV/div to 5 V/div) is used to record and store vibration signals obtained from an accelerometer. These signals are analyzed with Discrete Wavelet Transform (DWT) and further considered as input for processing data with **SVM-ANN.** The input referred for experimentation is shown in Table 1.

4. Implementation

In the experimental setup, misalignment is introduced externally to obtain the vibration signal required for analysis. The output vibration signals are recorded in all three directions viz. Longitudinal (V_g), Lateral (V_t), and Vertical
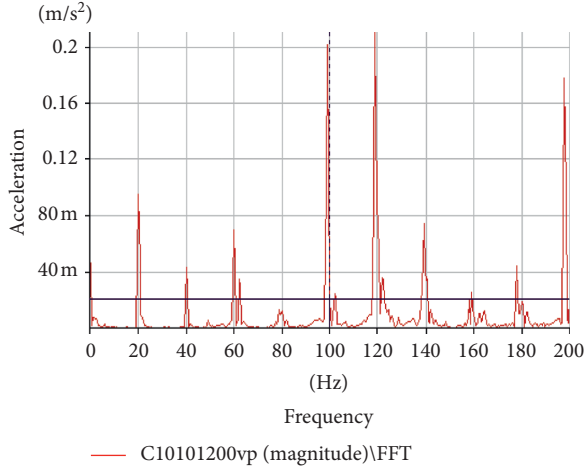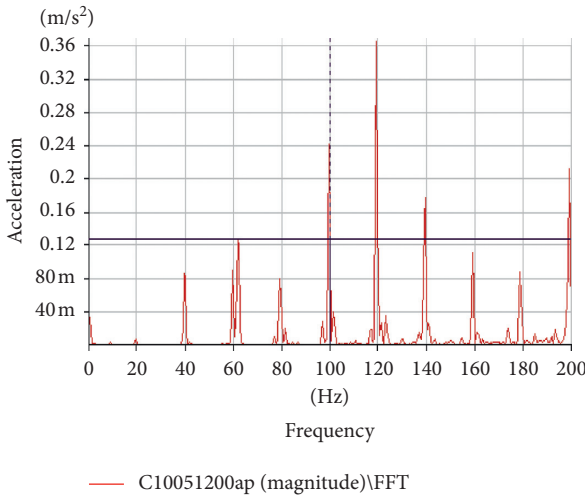
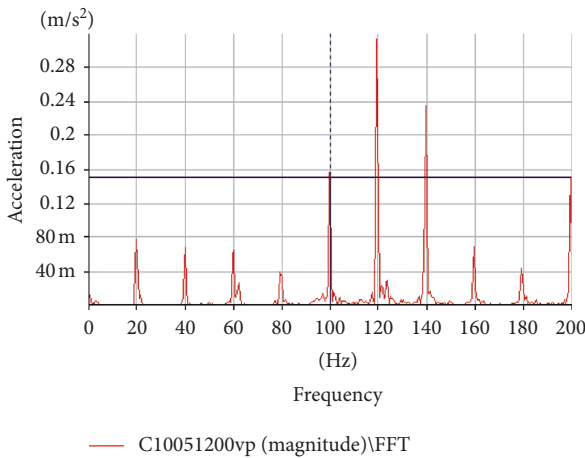Figure 3: 0.06 mm offset, 1200 rpm.

Figure 4: 0.08 mm offset, 1200 rpm.

Figure 5: 0.1 mm angular, 1200 rpm.

(V_r). For detailed discussion and comparison, a sample vibration signal at 1200 rpm and variation of misalignment in the range 0 to 0.2 mm is considered, as shown in Figure 7.
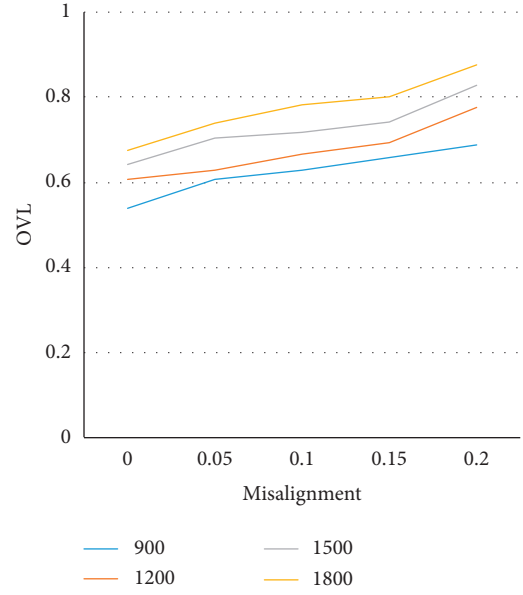
Figure 6: OVL v/s misalignment at different speeds.

Table 1: Input parameters for pilot experimentation.

Parameters	Conditions
Speed (rpm)	900, 1200, 1500, 1800
Parallel misalignment (mm)	0, 0.02, 0.04, 0.06, 0.08, 0.1, 0.12, 0.14, 0.16, 0.18, 0.2
Angular misalignment (mm/inch)	0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8.

It is observed that the overall vibration level (OVL) is increased with an increase in the value of misalignment, as compared in Figure 6.

The fault which may occur in the rotary machine is confirmed with the particular fault frequency. It clear that the misalignment of the shaft is observed at 1X and 2X frequency harmonics [16, 22]. For detailed discussion and comparison, a sample vibration signal at 1200 rpm and variation of misalignment in range 0 to 0.2 mm is considered, as shown in Figure 7. It is essential to understand fault information associated with nonstationary signals. This is possible with multilevel analysis using DWT. DWT is a widely used technique to obtain information in time and frequency domain.

The DC feature extracted is compared for various sample condition mentioned in Figure 7. The change in OVL for different misaligned condition has been reflected with corresponding change in DC values. It is compared and depicted in Figure 8. It is observed that few faults are registered at same common frequency as shown in the mechanical fault diagnostic chart [32]. In such cases, DWT assists to identify and isolate uncommon feature of fault.

The vibration signals measured with an accelerometer are shown in Figures 3–5, which are nonlinear in nature. In view of extracting useful information from nonlinear signals, it is essential to select proper signal processing technique and select proper features. The Mean, Max, Kurtosis, Skewness, RMS, and
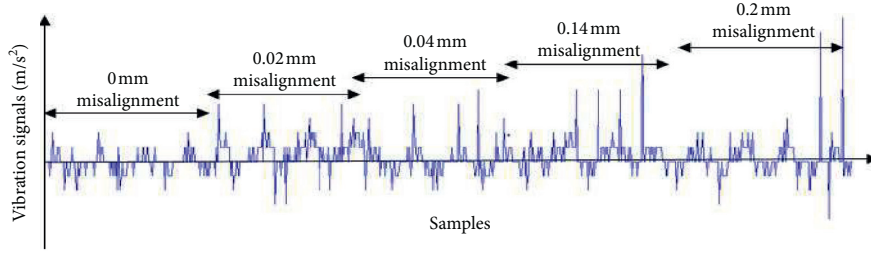
FIGURE 7: Vibration signals in vertical direction for parallel misalignment at 1200 rpm.
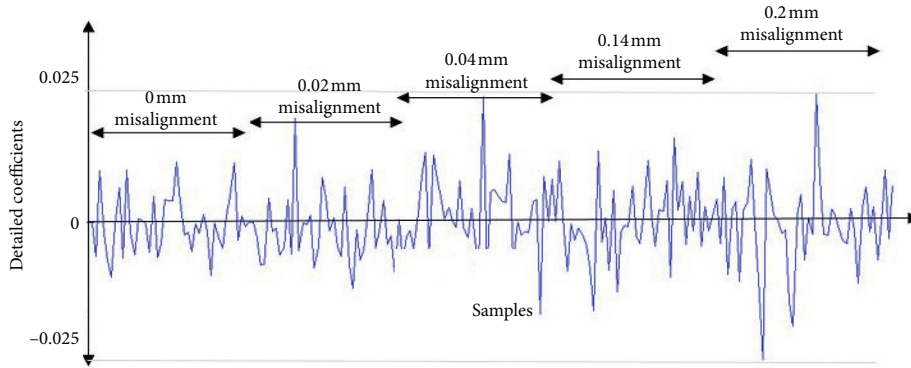

FIGURE 8: Detailed coefficients of vibration signals at the second level of decomposition.

Shannon entropy are commonly selected features that can be obtained from extracted feature of vibration signals. The required feature extraction of vibration signal is carried out using wavelet transform. Dwt is useful in time and frequency domain, which shows fault existing impulse effectively. Different mother wavelets are examined at different levels to select most suitable wavelet and suitable level based on maximum Energy to Shannon Entropy ratio (ESE), as shown in Table 2.

All mother wavelet considered are compared at different levels of decomposition based on ESE ratio, as shown in Figure 9. It is clearly observed for all considered mother wavelet that energy contained in signal is reduced as the level of signal decomposition increased. Therefore, level one is considered, as it contains maximum information for wavelet and feature selection. Different types of mother wavelet viz. Daubechies (DB), Coiflet, Symlet, HARR, DMEY, Biorthogonal, and Reverse Biorthogonal are considered during the analysis while using DWT. The correct level of mother wavelet is decided on the basis of maximum Energy to Shannon Entropy (ESE) ratio. This ratio is maximum at level 1 for all considered wavelet. The average of ESE for each class of fault and for each mother wavelet is plotted and compared in Figure 10. It is clear from Figure 10 that DB2 and SYM2 has higher ESE ratio. As mathematical function of these two wavelets are the same, so any one can be selected. Hence, DB2 mother wavelet at level 1 is selected for analysis.

5. Multiclass SVM Theory

The SVM is used as a classifier in the present work. The SVM refers to category of supervised learning algorithm in which set of output values are given to learning machine. Let us consider two classes of misalignment to be identified as Q_1 and Q_2. M is an unknown feature vector to be classified into classes.

The classification rule is applied, as shown in equation (4). In this equation, 'i' is the unknown input to be classified:

$$\left. \begin{array}{l} P(\mathbf{M_1}) = W^T \mathbf{M_i} + a > 0 \gg \mathbf{M_i} \epsilon Q_1 \\ P(\mathbf{M_2}) = W^T \mathbf{M_i} + a < 0 \gg \mathbf{M_i} \epsilon Q_2 \end{array} \right\}, \qquad (4)$$

where W^T is the orientation of hyperplane and 'a' is position of hyperplane.

The classifier is implemented to train the data by finding out the values of W^T and a. During training, ensure that the value of W^T and a is modified in such a way that $P(\mathbf{M_1})$ will come to the positive side of hyperplane. Similarly, modify the value of W^T and "a" in such way that $P(\mathbf{M_2})$ will come to the negative side of hyperplane. Support Vector Machine (SVM) finds the best position of line. SVM tries to keep the maximum distance between these classes and separating boundaries so that a small noise cannot misclassify the given feature of unknown input.

For every $\mathbf{M_i}$, assign class belonging to it. $\mathbf{M_i}$, belonging to class Y_i, where $Y_i = \pm 1$, can be written as

$$\left. \begin{array}{ll} \mathbf{M_i} \epsilon Q_1, & Y_i = +1, \\ \mathbf{M_i} \epsilon Q_2 & Y_i = -1 \end{array} \right\}. \qquad (5)$$

The generalized equation can be written as equation (6), irrespective of class:

TABLE 2: Comparison of mother wavelet at different levels on the basis of ESE.

| Level	DB1	DB2	BIOR 1.1	BIOR 1.3	COIF1	COIF2	COIF3	SYM2	RBIO1.3	DMEY
1	0.0389	0.0473	0.0389	0.0389	0.0287	0.0302	0.0366	0.0473	0.0394	0.0370
2	0.0200	0.0128	0.0200	0.0200	0.0296	0.0279	0.0276	0.0128	0.0196	0.0248
3	0.0107	0.0090	0.0107	0.0121	0.0124	0.0108	0.0105	0.0090	0.0111	0.0151
4	0.0065	0.0072	0.0065	0.0067	0.0065	0.0059	0.0066	0.0072	0.0059	0.0160
5	0.0051	0.0064	0.0051	0.0063	0.0068	0.0074	0.0079	0.0064	0.0059	0.0103
6	0.0025	0.0029	0.0025	0.0025	0.0023	0.0019	0.0016	0.0029	0.0029	0.0076
7	0.0024	0.0020	0.0024	0.0033	0.0021	0.0022	0.0023	0.0020	0.0026	0.0011
8	0.0028	0.0032	0.0028	0.0035	0.0041	0.0028	0.0020	0.0032	0.0031	0.0009
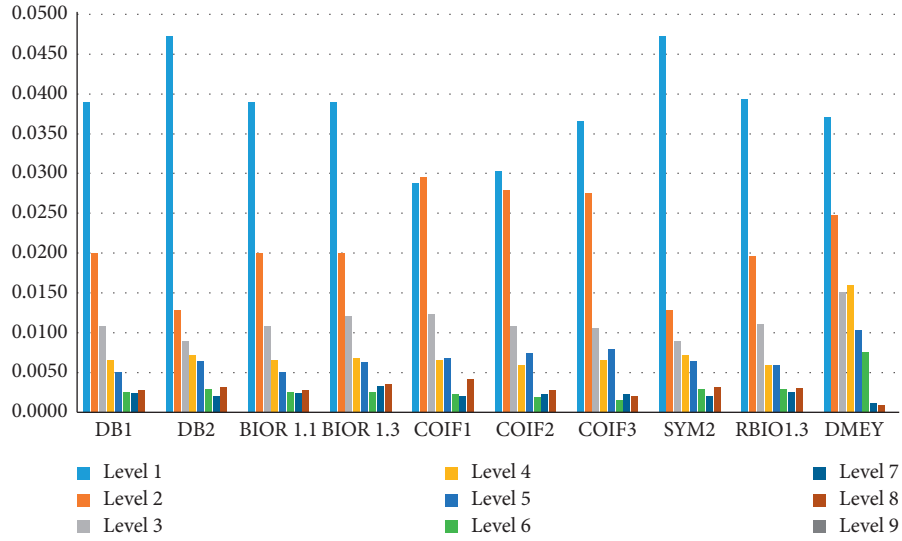
FIGURE 9: Comparison of mother wavelet at different levels.
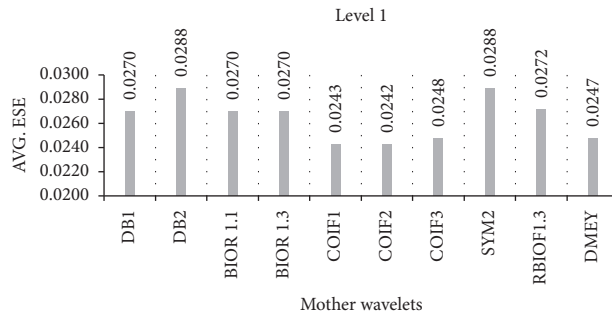
FIGURE 10: Comparison of DB mother wavelet at different levels.

$$Y_i \left(W_i \cdot \mathbf{M_i} + a \right) > 0. \qquad (6)$$

Once W and "a" are determined, the unknown vector P can be classified into two classes using equation (4). In SVM, representing maximum margin can be written as

$$P \left(\mathbf{M_1} \right) = W^T \mathbf{M} + a > \beta, \qquad (7)$$

where β is margin. The distance of $\mathbf{M}$ from the hyperplane is given by

$$\left(\frac{W \cdot \mathbf{M} + a}{\|W\|} \right) \geq \beta. \qquad (8)$$

By proper scaling, the β *parameter can be set to unity.* Hence, equation (8) can be written as

$$\begin{aligned} \mathbf{Y_i} \left(\mathbf{W_i} \cdot \mathbf{M_i} + \mathbf{a} \right) &= 1, \quad \text{if } \mathbf{M_i} \text{ is Support Vector,} \\ \mathbf{Y_i} \left(\mathbf{W_i} \cdot \mathbf{M_i} + \mathbf{a} \right) &> 1, \quad \text{if } \mathbf{M_i} \text{ is not Support Vector.} \end{aligned} \qquad (9)$$

From equation (8), it can be observed that the margin β can be maximized by minimizing $\|W\|$ and maximizing the bias 'a'. Therefore, a function to minimize the weight can be written as

$$\varnothing(W) = W^t W = \frac{1}{2} W \cdot W. \tag{10}$$

Similarly, for support vector, the constraint is obtained by the following equation to minimize the term $(1/2Wt \cdot nW)$.

The above constraint optimization problem can be converted into unconstraint optimization using the "Lagrangian Multiplier":

$$L(W, a) = \frac{1}{2}(W \cdot W) - \Sigma \alpha_i [Y_i (W \cdot \mathbf{M_i} + a)] - 1, \tag{11}$$

where α_i is the Lagrangian Multiplier Optimization of equation (11), which can be carried out by taking derivative with respect to 'a' and equating it to zero and equation (12) is obtained:

$$\sum_{i=1}^{n} \alpha_i Y_i = 0. \tag{12}$$

Similarly, Lagrangian Multiplier Optimization can be carried out by differentiating equation (11) w.r.t W, and equation (13) is obtained:

$$W = \Sigma \alpha_i y_i \mathbf{M_i}. \tag{13}$$

The above binary classification is applicable if class labels have only two values (k-class, $k < 2$). In some cases, it is required to deal with more than two classes in the actual fault diagnosis of the machine. In such a case (k-class, $k > 2$), Multiclass SVM classifier is used. It can be obtained by combining several binary classifiers. The various methods of obtaining Multiclass SVM are viz. One-against-all (OAA), One-against-one (OAO), and Direct acyclic graph (DAG). The OAO is observed most commonly as the considered method [24]. This method constructs $k * (k-1)/2$ classifier, where each one is trained on data from two classes. Let, ith and jth be two classes for training, which can be explained as follows:

$$\text{Minimize: } \frac{1}{2}\|w^{ij}\|^2 + c\sum_t \xi_t^{ij}(w^{ij})^T,$$

$$\text{for the condition } (w^{ij})^T \phi(x_t) + b^{ij} \geq 1 - \xi_t^{ij}, \quad \text{if } y_t = i,$$

$$(w^{ij})^T \phi(x_t) + b^{ij} \leq -1 + \xi_t^{ij}, \quad \text{if } y_t = j. \tag{14}$$

The decision is based on the following rule.

If $((w^{ij})^T \phi(x_t) + b^{ij})$ says that x is in ith class, then one vote will be added to ith class. If this is not true, then a vote will be added to the jth class. Accordingly largest vote count will decide class for x.

6. Training of ANN

The normalized Kurtosis features are used as an input to train ANN. The Kurtosis feature values are obtained for different misalignment conditions for three directions. The various combinations of ANN structures considered are as shown in Table 3. For training, the Levenberg–Marquardt algorithm method is used with tan sigmoid activation function. The selection of the best structure depends on the size training data, neurons in input, hidden, and output layer, and the initial weight is assigned to the input signals. In the range of 0 to 0.2 mm, 10 conditions of misalignment are considered for different operating speeds up to 2100 rpm, respectively. In the proposed work, for ANN training, 6000 samples are considered, which includes 4500 samples of the misaligned condition and 1500 samples of aligned (healthy) conditions. The samples are considered as 50% of the total for the training of ANN, 25% of the total sample for testing of ANN, and the remaining 25% for crossvalidation of ANN. The Machine Learning Models (MLMs) can guarantee the optimal performance, as it test several models, first with the fundamental ones. Crossvalidation is a method for assessing MLMs by training numerous MLMs on subsets of the accessible input data and assessing them on the complementary subset of the data. Use crossvalidation to detect overfitting, failing to generalize a MLM.

An input layer, hidden layer, and output layer are considered as the main part of the ANN structure. For the proposed ANN structure, it has three inputs. The attainment of accuracy controls the number of hidden layers. Hence, the number of hidden layers is considered a trial and error basis such as 5, 10, 15, 20, and 25. The learning rate during the training of ANN is varied in the range of 0.01 to 0.04. The number of epochs considered is 1800. The different structures considered for training and testing are shown in Table 3.

The ANN structure 3-20-11 with learning rate 0.03 is good for MoM in offset misalignment and angular misalignment. The performance of ANN during training is shown in Figure 11. The MSE, is 0.05 which is obtained at 1300 epochs. The MATLAB software is used for training, testing, and validation of ANN. The output error observed in the testing of trained ANN is calculated as below:

$$\text{Error} = \frac{(E_o - A_o) \times 100}{A_o}, \tag{15}$$

where E_o is expected output and A_o is actual output.

7. Results and Discussion

In the present study of shaft misalignment, the output vibration signals are obtained from the experimental setup. The output signals are normalized to achieve reduction in data size required for training as explained earlier.

The average ESE is more for DB2 and SYM2. The DB2 wavelet is selected as the suitable mother wavelet as explained earlier. The eighteen statistical features are obtained to analyze information in output signal. The ReliefF algorithm is used to optimize feature selection on rank basis. The sample vibration signal at 1200 rpm for misalignment is considered for presenting key points of analysis. It is clearly seen that all mother wavelet reflect good ESE ratio at level 1. The explanation in support with fact that disorder, i.e.,

TABLE 3: Comparison of ANN structure.

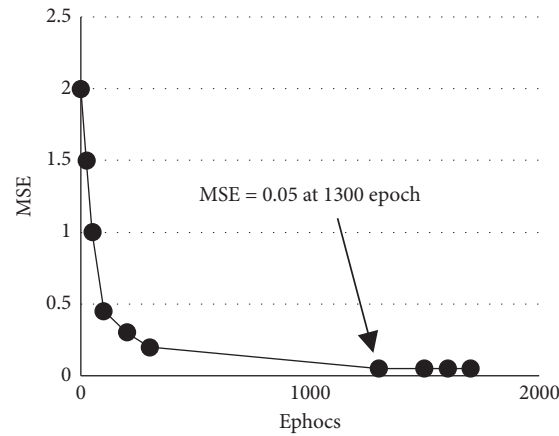| Structure	Learning rate	Accuracy
3-5-11	0.03	87.70
3-10-11	0.03	89.25
3-15-11	0.03	93.00
3-20-11	0.03	94.17
3-25-11	0.03	91.00
3-5-11	0.02	88.29
3-10-11	0.02	90.00
3-15-11	0.02	91.57
3-20-11	0.02	92.00
3-25-11	0.02	92.25

FIGURE 11: Variation of MSE in training of ANN.

TABLE 4: Ranking-based feature selection using ReliefF algorithm.

			Accuracy in percentage								
Position	Order/rank	Features	Complex tree	Medium tree	Simple tree	Linear SVM	Quadratic SVM	Cubic SVM	Fine Gaussian SVM	Medium Gaussian SVM	Coarse Gaussian SVM
1	6	Skew_1Z	33.1	23	16.4	12.0	20.5	23.4	19.8	15.2	8.4
1_2	5	Skew_1Y	58.5	28.7	20.3	25.9	39.3	53.7	40.1	24.7	10.7
1_3	4	Skew_1X	64.8	34.0	20.1	39.1	59.1	69.1	54.7	33.4	14.4
1_4	11	Min_1Y	76.2	49.9	20.0	54.8	70.3	76.8	74.0	53.1	28.2
1_5	15	Max_1Z	81.1	51.8	18.2	63.8	78.9	80.1	82.5	63.9	40.7
1_6	12	Min_1Z	83.0	53.7	18.2	70.7	80.3	80.2	85.2	67.1	44.7
1_7	14	Max_1Y	85.7	56.7	18.2	72.8	82.5	84.4	89.5	75.6	53.2
1_8	10	Min_1X	85.2	58.9	18.4	78.0	83.8	85.9	89.7	80.1	60.1
1_9	3	Kurtosis_1Z	87.3	55.0	18.4	78.9	84.9	86.4	89.3	80.1	63.4
1_10	2	Kurtosis_1Y	87.3	55.0	18.4	78.7	85.5	86.5	89.4	81.7	63.2
1_11	1	Kurtosis_1X	88.7	55.5	18.4	78.1	86.1	86.5	88.7	83.4	63.4
1_12	13	Max_1X	88.7	56.0	18.4	80.4	86.4	86.7	88.8	82.7	66.4
1_13	9	Rms_1Z	88.7	65.0	22.4	79.4	85.7	87.5	88.5	87.9	67.2
1_14	8	Rms_1Y	89.2	65.0	22.4	79.9	86.1	87	88.4	83.3	67.4
1_15	7	Rms_1X	88.4	65.0	22.4	79.7	85.9	86.5	88.1	83.2	67.2
1_16	18	EN_1Z	89.3	64.9	22.4	79.7	85.7	86.4	87.8	83.4	67.3
1_17	17	EN_1Y	89.2	64.7	22.4	79.3	85.7	86.2	87.6	83.1	67.2
1_18	16	EN_1X	89.1	64.7	22.4	79.2	85.9	86.5	87.4	83.7	66.9

entropy is minimum and information is maximum at the first level of signal decomposition. This is the basis for selection of level 1 for comparison.

The rank basis feature optimization is carried out by using ReliefF algorithm. The kurtosis feature is observed as rank 1 feature. Therefore, kurtosis features of all

TABLE 5: SVM multifault classifier output for selected samples at 1200 rpm.

Rotational speed (rpm)	Fault type	Number of samples	Classed by SVM			SVM: distance to H
			Healthy	Offset	Angular	
1200	Healthy	10	0	—	—	0.0139
	Misaligned	10	—	—	−1	−1.925
	Misaligned	10	—	1	—	1.239
	Misaligned	10	—	1	—	2.920
	Misaligned	10	—	—	−1	−1.110
	Misaligned	10	—	1	—	3.572

TABLE 6: Results obtained using classification and prediction of the shaft misalignment method.

Rotational speed (rpm)	Type of misalignment	Misalignment (mm)	SVM results		ANN result		% error
			SVM: distance to H	Classed by SVM	Expected outcome (E_o)	Actual outcome (A_o)	
500	Parallel	0	0.002	0	0.00	0.0030	0.30
		0.02	1.239	1	0.02	0.0204	2.00
		0.04	2.920	1	0.04	0.0414	3.50
		0.06	3.572	1	0.06	0.0622	2.75
	Angular	0	0.022	0	0.00	0.0019	0.19
		0.01	−1.925	−1	0.01	0.0102	2.00
		0.03	−1.110	−1	0.03	0.0312	4.02
		0.04	−1.592	−1	0.04	0.0410	2.69
1200	Parallel	0	0.014	0	0.00	0.0025	0.25
		0.02	1.239	1	0.02	0.0205	2.79
		0.04	2.920	1	0.04	0.0406	1.57
		0.06	3.572	1	0.06	0.0617	2.89
	Angular	0	0.022	0	0.00	0.0040	0.40
		0.01	−1.925	−1	0.01	0.0103	3.89
		0.03	−1.110	−1	0.03	0.0308	2.79
		0.04	−1.592	−1	0.04	0.0418	4.68
2100	Parallel	0	0.012	0	0.00	0.0010	0.10
		0.02	1.239	1	0.02	0.0251	2.57
		0.04	2.920	1	0.04	0.0408	2.15
		0.06	3.572	1	0.06	0.0616	2.67
	Angular	0	0.022	0	0.00	0.0030	0.30
		0.01	−1.925	−1	0.01	0.0104	4.21
		0.03	−1.110	−1	0.03	0.0311	3.92
		0.04	−1.592	−1	0.04	0.0408	2.21
Average % error:							2.28

signals are obtained using DB2 at level 1. The kurtosis feature of normalized vibration signals is considered as an input to train ANN and SVM. In the rank-based feature selection, the Kurtosis as a single feature shows efficiency 19.8 %. This efficiency has been improved upto 89.7 % by adopting combination of top eight ranked features, as depicted in Table 4. It is clear that eight top-ranked features are sufficient to obtain highest efficiency for fine Gaussian SVM classifier referred in this implementation.

In the CPSM approach, SVM and ANN are used for misalignment analysis. SVM classifier is applied to identify the class of misalignment. The misalignment condition which is to be classified has assigned a number as 0, 1, and −1 for healthy condition, offset misalignment, and angular misalignment, respectively. The few samples of output vibration signals are tested with SVM to confirm the output of fault classification, as shown in Table 5. It contains classification results obtained for three conditions related to healthy and misalignment. The sample cases of parallel and angular misalignment for varying load conditions are considered, as shown in Table 6.

The output of Multiclass SVM has obtained with good accuracy, as shown in Table 6. The selected kurtosis feature of DC of all normalized vibration signals is taken as input to SVM.

The points which has significant effect on improving the accuracy of CPSM are explained. The best part of the CPSM is the normalization of the vibration signal. The correct selection of mother wavelet is on the basis of the ESE ratio. The selection of the correct ANN structure has also contributed to minimizing MSE. The application of ReliefF algorithm for deciding rank of features and selecting top-ranked feature combination in training has also influenced in improving classification accuracy of misalignment.

8. Conclusion

A combined approach of Support Vector Machine and Artificial Neural Network is implemented successfully to obtain classification and prediction of shaft misalignment. The main contribution of study lies in implementing normalization of signals and ranking of features which is uncommon in problem of misalignment analysis. The selection of proper mother wavelet on the basis of maximum ESE has contributed well in feature selection. The implementation of the classification and prediction of the shaft misalignment method is seen with the least error in output results in misalignment fault classification and MoM. The accuracy of Support Vector Machine is seen to a good level for all conditions viz. healthy, parallel, and angular misalignment. In tested results of trained artificial neural network, a 3-20-11 structure is observed as the best artificial neural network structure for offset and angular measure of misalignment. The use of first eight ranked features has improved classification accuracy. It is observed that the selection of epochs and learning rate has also an effect on minimizing the error. The different samples tested for parallel and angular misalignment have classified successfully with the implementation of support vector machine. The average error observed in the output of the trained artificial neural network is 2.28%. It is concluded that the error in output of classification and prediction of shaft misalignment is within limit due to normalization, correct wavelet selection on the basis of maximum energy to Shannon entropy ratio, and rank-based feature selection using ReliefF algorithm. This approach is helpful in effective real-time condition monitoring of machines. The future scope of this work can be extended for fault prognosis of other faults related with bearing, gears, insufficient lubrication, and unbalance combined with misalignment to ensure effective condition-based maintenance.

Data Availability

Data used in this work can be made available on request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

References

[1] M. Vishwakarma, R. Purohit, V. Harshlata, and P. Rajput, "Vibration analysis & condition monitoring for rotating machines: a review," *Materials Today: Proceedings*, vol. 4, no. 2, pp. 2659–2664, 2017.

[2] Y. Tang, F. Lin, and Q. Zou, "Complexity analysis of time-frequency features for vibration signals of rolling bearing based on local frequency," *Shock and Vibration*, vol. 2019, Article ID 7190568, 13 pages, 2019.

[3] J.-S. R. Jang, "ANFIS: network-based fuzzy inference system," *IEEE Transactions on Systems, Man and Cybernetics*, vol. 23, no. 3, pp. 665–685, 1993.

[4] S. Li, Y. Nie, and J. Li, "Condition monitoring and diagnosis of power equipment: review and prospective," *Journal of Institution of Engineering and Technology*, vol. 2, no. 2, pp. 82–91, 2017.

[5] Z. YanPing, H. ShuHong, H. JingHong, S. Tao, and L. Wei, "Continuous wavelet grey moment approach for vibration analysis of rotating machinery," *Mechanical Systems and Signal Processing*, vol. 20, no. 5, pp. 1202–1220, 2006.

[6] X. Qi, "Diagnosis of misalignment faults by tacholess order tracking analysis and RBF networks," *Neuro Computing an International Journal*, vol. 169, pp. 439–448, 2014.

[7] P. Shi and W. Yang, "Precise feature extraction from wind turbine condition monitoring signals by using optimised variational mode decomposition," *IET Renewable Power Generation*, vol. 11, no. 3, pp. 245–252, 2017.

[8] H. Sohn and C. R. Farrar, "Damage diagnosis using time series analysis of vibration signals," *Smart Materials and Structures*, vol. 10, no. 3, pp. 446–451, 2001.

[9] R. B. Dhumale and S. D. Lokhande, "Neural network fault diagnosis of voltage source inverter under variable load conditions at different frequencies," *Measurement*, vol. 91, pp. 565–575, 2016.

[10] W. Wang, "An intelligent system for machinery condition monitoring," *IEEE Transactions on Fuzzy Systems*, vol. 16, no. 1, pp. 110–122, 2008.

[11] C. Liu and F. Wang, "A review of current condition monitoring and fault diagnosis methods for low speed and heavy loading slewing bearings," in *Proceeding of 9th International Conference on Modeling, Identification and Control*, Kunming, China, July 2017.

[12] H. Shao, J. Hongkai, L. Xingqiu, and W. Shuaipeng, "Intelligent fault diagnosis of rolling bearing using deep wavelet auto-encoder with extreme learning machine," *Knowledge-Based Systems*, vol. 140, pp. 1–14, 2018.

[13] O. Tonks and Q. Wang, "The detection of wind turbine shaft misalignment using temperature monitoring," *CIRP Journal of Manufacturing Science and Technology*, vol. 17, pp. 71–79, 2017.

[14] Q. Huang, X. Yan, Y. Wang, C. Zhang, and Z. Wang, "Numerical modeling and experimental analysis on coupled torsional-longitudinal vibrations of a ship's propeller shaft," *Ocean Engineering*, vol. 136, pp. 272–282, 2017.

[15] J. L. F. Chacon, E. A. Andicoberry, V. Kappatos, G. Asfis, T.-H. Gan, and W. Balachandran, "Shaft angular misalignment detection using acoustic emission," *Applied Acoustics*, vol. 85, pp. 12–22, 2014.

[16] A. R. Mohanty and S. Fatima, "Shaft misalignment detection by thermal imaging of support bearings," *IFAC-PapersOnLine*, vol. 48, no. 21, pp. 554–559, 2015.

[17] M. Xu and R. D. Marangoni, "Vibration analysis of A motor-flexible coupling-rotor system subject to misalignment and unbalance, part I: theoretical model and analysis," *Journal of Sound and Vibration*, vol. 176, no. 5, pp. 663–679, 1994.

[18] M. Amirrezai and H. Rezaie, "A new fuzzy based current control strategy with fixed switching frequency for improving

D-STATCOM performance," *Journal of Intelligent and Fuzzy Systems*, vol. 35, no. 3, pp. 32854–33294, 2018.

[19] A. Singh, V. Sugumaran, and M. Amarnath, "Fault diagnosis of Helical gearbox through vibration signals using J48 decision tree and wavelet," *Indian Journal of Science and Technology*, vol. 9, no. 33, 2016.

[20] S. Patra and L. Bruzzone, "A novel SOM-SVM based active learning technique for remote sensing image classification," *IEEE Transaction on Geoscience and Remote Sensing*, vol. 52, no. 11, 2014.

[21] R. Kumar and K. P. Singh, "SVM with feature selection and extraction techniques for defect-prone software module prediction," *Advances in Intelligent Systems and Computing*, pp. 279–289, Springer, Berlin, Germany, 2017.

[22] A. Umbrajkaar and A. Krishnamoorthy, "Vibration analysis using wavelet transform and fuzzy logic for shaft misalignment," *Journal of Vibroengineering*, vol. 20, no. 8, 2018.

[23] C.-W Hsu and C.-J. Lin, "A comparison of methods for Multiclass support vector machines," *IEEE Transactions on Neural Networks*, vol. 13, no. 2, pp. 415–425, 2002.

[24] R. B. Dhumale and N. D. Thombare, "Leaf decease classification using machine learning," *Journal of Emerging Technologies and Innovative Research*, vol. 6, no. 4, pp. 425–435, 2019.

[25] Y. Yang, U. Dakermann, J. Li, and E. Niederleithinger, "Wavelet packet energy based damage identification of wood utility poles using support vector machine multi-classifier and evidence theory," *Structural Health Monitoring*, vol. 18, no. 1, pp. 123–142, 2019.

[26] Y. Yu, W. Li, J. Li, and T. N. Nguyen, "A novel optimised self-learning method for compressive strength prediction of high performance concrete," *Construction and Building Materials*, vol. 184, pp. 229–247, 2018.

[27] Y. Yang, Y. Li, and J. Li, "Nonparametric modelling of magnetorheological elastomer base isolater based on artificial neural network optimized by ant colony algorithm," *Journal of Intelligent Material Systems and Structure*, vol. 26, no. 14, pp. 1789–1798, 2015.

[28] S. Golnaraghi and Z. Zangenehmadar, "Application of artificial neural network(s) in predicting formwork labour productivity," *Advances in Civil Engineering*, vol. 2019, Article ID 5972620, 11 pages, 2011.

[29] M. M. Najafabadi and F. Villanustre, "Deep learning applications and challenges in big data analytics," *Journal of Big Data*, vol. 1, 2015.

[30] O. I. Abiodun, A. Jantan, O. E. Omolara, and K. Dada, "State-of-the-art in artificial neural network applications: a survey," *Journal Heliyon*, vol. 4, no. 11, Article ID e00938, 2018.

[31] J. Maillo and J. Luengo, "Exact *K*-nearest neighbor classification for big datasets," in *Proceedings of the IEEE International Conference on Fuzzy Systems (Fuzz-IEEE)*, Naples, Italy, July 2017.

[32] ISO 10816: Mechanical Vibration Chart for Fault Diagnosis.

[33] T. N. Nguyen and Y. Yang, "Elastic modulus of ASR affected concrete: an evaluation using Artificial Neural Network," *Computers and Concrete*, vol. 24, no. 6, 2019.