IJE TRANSACTIONS A: Basics Vol. 34, No. 01, (January 2021) 212-223

International Journal of Engineering

Journal Homepage: www.ije.ir

Gear Fault Detection using Machine Learning Techniques- A Simulation-driven Approach

V. C. Handikherkar*, V. M. Phalle

Department of Mechanical Engineering, Veermata Jijabai Technological Institute (VJTI), Mumbai, India

PAPER INFO

Paper history:
Received 12 August 2020
Received in revised form 18 September 2020
Accepted 30 October 2020

Keywords:
Machine Learning
Simulated Data
Vibration Analysis
Gear Fault Diagnosis
Condition Monitoring

ABSTRACT

Machine Learning (ML) based condition monitoring and fault detection of industrial equipment is the current scenario for maintenance in the era of Industry-4.0. The application of ML techniques for automatic fault detection minimizes the unexpected breakdown of the system. However, these techniques heavily rely on the historical data of equipment for its training which limits its widespread application in industry. As the historical data is not available for each industrial machine and generating the data experimentally for each fault condition is not viable. Therefore, this challenge is addressed for gear application with tooth defect. In this paper, ML algorithms are trained using simulated vibration data of the gearbox and tested with the experimental data. Simulated data is generated for the gearbox with different operating and fault conditions. A gearbox dynamic model is utilized to generate simulated vibration data for normal and faulty gear condition. A pink noise is added to simulated data to improve the exactness to the actual field data. Further, these simulated-data are processed using Empirical Mode Decomposition and Discrete Wavelet Transform, and features are extracted. These features are then fed to the training of different well-established ML techniques such as Support Vector Machine, Random Forest and Multi-Layer Perceptron. To validate this approach, trained ML algorithms are tested using experimental data. The results show more than 87% accuracy with all three algorithms. The performance of the trained model is evaluated using precision, recall and ROC curve. These metric show the affirmative results for the applicability of this approach in gear fault detection.

doi: *10.5829/ije.2021.34.01a.24*

NOMENCLATURE

$I_m / I_b / I_1 / I_2$	Mass moment of inertia of rotor/load/pinion/gear	c_p / c_g	Torsional damping of flexible coupling Input/Output
M_1 / M_2	Input/Output torque from Motor/Load	k_1 / k_2	Vertical Radial stiffness of bearing Input/Output
m_1 / m_2	Mass of pinion/gear	c_1 / c_2	Vertical Radial viscous damping coefficient of bearing Input/Output
R_{b1} / R_{b2}	Base circle of pinion/gear	y_1 / y_2	Linear displacement of Pinion/Gear in the y-direction
k_p / k_g	Torsional stiffness of flexible coupling Input/Output	$\theta_m / \theta_b / \theta_1 / \theta_2$	Angular displacement of motor/load/pinion/gear

1. INTRODUCTION

Rotating machinery are the most essential systems of the industrial machinery. Gearboxes are the most widely used sub-systems of the rotating machinery that are vulnerable to failure and system breakdown. As they operate under harsh operating conditions, which may develop fault on gears. Also, continuous operation under these conditions causes gear to degrade and leads to the failure. Failure of gear causes the transmission system breakdown, production and economic loss.

Different maintenance strategies such as breakdown or unplanned, preventive or scheduled and Condition Based Maintenance (CBM) are employed to ensure the satisfactory operation of rotating machinery over its useful life. Earlier was the breakdown or unplanned

*Corresponding Author Email: *vishwadeepch@gmail.com* (V. C. Handikherkar)

Please cite this article as: V. C. Handikherkar, V. M. Phalle, Gear Fault Detection using Machine Learning Techniques- A Simulation-driven Approach, International Journal of Engineering, Transactions A: Basics Vol. 34, No. 01, (2021) 212-223

maintenance in which maintenance is carried out only at the breakdown; preventive or scheduled maintenance was carried out at predefined intervals, and CBM was carried out based on the information on the condition of machine [1]. Out of this CBM strategy gained popularity in the industry as it avoids unnecessary maintenance. In current Fourth Industrial Revolution (i.e. Industry 4.0) for industry equipment maintenance, machine learning based condition monitoring system are being developed for automatic fault diagnosis [2]. Machine learning has been applied not only in industry equipment maintenance but also in different fields such as roadways maintenance [3], predicting student grades [4] etc.

Vibration analysis is a most widely used condition monitoring technique for gear fault daignosis. In literature two approaches have been used for the gear fault diagnosis, one is data driven approach and other is physical model based approach. The data-driven approach purely rely on the historical or in-service data of gearbox to predict the faults in gear, and physical model based approach makes use of physics based models to create a virtual system to mimic the vibration characteristics of gearbox under different operating conditions [5]. Subsequent section discusses the literature on these two approaches.

Several researchers have used Machine Learning (ML) techniques for developing automatic fault detection of industrial machinery based on the data driven approach. Recently, Lei et al. [6] presented a review of different ML techniques employed for machine fault diagnosis. To develop a fault diagnosis technique based on the data driven approach using the ML techniques, require a historical data of in-service equipment or experimental data to train the ML algorithms. ML techniques like Support Vector Machine (SVM), k-Nearest Neighbour (kNN), Artificial Neural Network (ANN) Ensemble techniques etc. [6] have been employed for the bearing and gear fault diagnosis. Samanta [7] used this approach for the binary classification (i.e. healthy and faulty) of gear using SVM and ANN. In this input features were selected and optimized using the genetic algorithm, SVM resulted in better classifier over the ANN. Similarly, Samanta et al. [8] used three different ANN classifiers such as Multi-Layer Perceptron (MLP), Radial Basis Function Network and Probabilistic Neural Network for the bearing fault classification. Using genetic algorithm and Probabilistic Neural Network a test accuracy of 100% was obtained. Further, the effectiveness of pre-processing of data using Discrete Wavelet Transform (DWT) on the classification by SVM and ANN was studied by Tyagi and Panigrahi [9], and results show that pre-processing improves the performance of both the classifiers and that SVM outperforms ANN. Discrete wavelet transform and multi-layer perceptron was used by Sanz et al. [10] to determine the gear condition status and the model is able to predict 1% decrease in the mesh stiffness. Shen et al. [11] used a transductive SVM for gear fault classification for data having more numbers of unlabeled data than labelled data; in this features were extracted using Empirical Mode Decomposition (EMD). Shao et al. [12] also utilized an EMD technique with higher-order cumulant method for gear fault classification and developed a virtual system for gear damage detection. Li et al. [13] proposed a bearing fault detection method using Improved Iterative Windowed Interpolation Discrete Fourier Transform technique. For the combined gear and bearing fault detection Dhamande and Chaudhari [14] proposed that, features extracted using continuous and discrete wavelet transform have been more prominent in detecting the combined fault than time and frequency domain features. A highest accuracy of 90% and 97% for training and testing respectively was obtained using the SVM. Attaran et al. [15] developed bearing fault detection technique based on kurtogram in time-frequency domain using ANN. A 100% training accuracy was noted for ANN. Bajric et al. [16] used features extracted using the discrete wavelet transform and time synchronous averaging method for a wind turbine gearbox fault detection. Researchers have also utilized ensemble techniques such as Random Forest (RF) for the fault classification. Han and Jiang [17] used RF classifier for the bearing fault classification; in this the variational mode decomposition and autoregressive model parameters have been employed for the fault feature extraction. Cerrada et al. [18] utilized RF classifier for the gear fault classification and used a genetic algorithm to select the best features and a best precision value of 0.9781 was obtained. Patil and Phalle [19] have used Random Forest, Gradient Boosting Classifier and Extra Tree classifier ensemble techniques for the bearing fault classification, in this features were ranked using decision tree and randomized lasso feature ranking technique and fed to these classifiers. Results showed that the features, ranked using DT technique, when fed to the classifier provided better accuracy compared to randomized lasso with fewer features and execution time. A highest accuracy of 98.21% was recored using DT ranking technique. In literature cited above fault diagnosis system was developed based on the data driven approach and utilized an experimental test rig to generate the training dataset for the training of ML algorithms.

In physical model based approach, dynamic model is used to mimic the actual operating conditions of the gearbox, and vibration response of gearbox under different conditions can be studied theoretically. Numerous dynamic models of the gearbox have been developed by researchers to study the vibration characteristics of the gearbox under healthy and faulty gear conditions. Liang et al. [5] presented a review of different gearbox fault dynamic models developed. The vibrations in gears are caused due to fluctuation in applied load, speed and Time-Varying Mesh Stiffness (TVMS), transmission errors etc. When faulty tooth engages the TVMS changes, due to this change in vibration response is observed. Therefore, the calculation of the TVMS for normal and faulty gear condition is essential. Researchers have developed different analytical methods such as potential energy, square waveform and finite element method for calculating the TVMS [5]. For obtaining the vibration response different models have been developed. Bartelmus [20] developed a dynamic model having 8-Degrees of Freedom (DoF) incorporating torsional and lateral motion and friction. Howard et al. [21] developed a 16-DoF model to study the effect of crack on gear tooth and friction between the tooth in contact on the vibration response. Abouel-seoud et al. [22] developed a model for wind turbine gearbox having twelve DoF to study the vibration response of gearbox under three faults like crack, spall and tooth breakage. A single-stage spur gearbox model incorporating the gyroscopic effect was developed by Mohammed et al. [23]. Literature cited above discusses the use of physical model based approach in gear fault diagnosis and the study is limited to calculating TVMS, obtaining the vibration response under different fault conditions and identifing the most sensitive condition indicators. The vibration response obtained using dynamic model is not having any noise, but in actual practice vibration response is masked with the environmental noise and determining the fault in such noisy data using the condition indicators is not possible.

It is clear that for the application of the data driven approach historical data is required for training of ML algorithm and in physical model based method study is limited to calculating TVMS, obtaining vibration response under different fault conditions and determining sensitive condition indicators. But these condition indicators does not perform well in case of actual vibration data.

Most of the studies reported in the literature for fault diagnosis of mechanical components using ML techniques have been utilizing a data-driven approach. The dependence of this approach on historical data from in-service equipment or data from the experimental test setup to train the ML algorithm restricted its full spread implementation in industrial machinery fault diagnosis. As in-service data for the equipment is not available and generating data using experimental test rig is not viable. Also, the cost associated with creating the data for each fault and at different operating conditions is very high; and model trained using these data are often valid for the condition and machine for which the data is collected. Also, training of ML algorithms are dependent on the diversity of data i.e. data at different operating conditions, more diverse data better is training. But generating this kind of diverse data experimentally is not feasible. Therefore, an alternative approach is required to overcome this limitation of the data driven approach and generate the diverse data with minimum cost.

The present paper addresses the limitation of the data-driven approach by employing a dynamic model of gearbox to generate the training dataset that includes the extensive variety of operating and fault conditions, for the training of ML algorithm. The data is generated by simulating the actual conditions of the gearbox therfore this approach is called as a simulation driven approach. In this a single-stage spur gearbox dynamic model is employed to generate simulated vibration acceleration data for different gear conditions (normal and faulty) at different loads and speeds. ODE15s solver function in Matlab is used to solve the equations of motion to obtain the simulated vibration response. Pink noise is added to simulated data to improve its exactness to the actual field data. These data are then processed using the DWT and EMD signal processing techniques, and features are extracted to create a training dataset. Various extensively used ML algorithms like SVM, MLP and RF are trained using this simulated training data set and tested using the experimental data.

2. METHODOLOGY

Figure 1 shows the schematic of the methodology of the simulation-driven approach adopted for gear fault detection. The simulation-driven gear fault diagnosis approach proposed in this work involves obtaining vibration acceleration data for normal and faulty gear conditions at different loads and speeds using the six DoF gear dynamic model. Pink noise is added to obtained simulated vibration data to improve the exactness

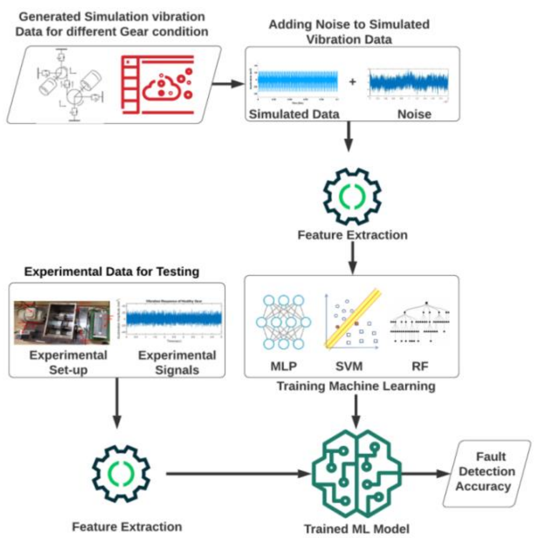


Figure 1. Schematic for Simulation Driven Fault Detection Methodology

towards the actual vibration data. Discrete Wavelet Transform (DWT) and Empirical Mode Decomposition (EMD) are used to process these vibration signals, and then statistical features are extracted. Using these features training data set is prepared and fed for the training of ML algorithms such as SVM, MLP and RF. To test this simulation-driven approach experimental data is collected from the experimental test rig. This experimental data is processed similarly as simulated data and features are extracted to prepare the testing data. Also, the theoretical background of signal processing and ML techniques used in this study are presented in this section.

2. 1. Discrete Wavelet Transform (DWT) DWT is an effective tool for signal and image processing in a wide range of research as well as in industrial applications [14,16]. The wavelet transform gives both frequency and time domain information about the signal. The continuous wavelet transform of signal *x(t)* is

$$W_{\psi}(\tau,s)=\int_{-\infty}^{+\infty} x(t)\psi^{*}_{\tau,s}(t)dt \tag{1}$$

where $\psi^{*}_{\tau,s}(t)$ is a conjugate of $\psi_{\tau,s}(t)$, that is the scaled and shifted version of the transforming function, called a mother wavelet which is defined as:

$$\psi_{\tau,s}(t)=\frac{1}{\sqrt{s}}\psi\left(\frac{t-\tau}{s}\right) \tag{2}$$

The transformed signal is a function of translation (τ) and scale (*s*) parameters. Other wavelet functions can be derived using the mother wavelet. Scale and translation correspond to frequency band and time information respectively in the transform domain. The DWT is derived from the discretization of $W_{\psi}(\tau,s)$ given by

$$DWT(j,k)=\frac{1}{\sqrt{2^{j}}}\int_{-\infty}^{\infty} x(t)\psi\left(\frac{t-2^{j}k}{2^{j}}\right)dt \tag{3}$$

Vector A and D are obtained by passing the signal x through low and high pass filters. approximate and detailed coefficients are obtained by downsampling these vectors. By repeating the process of decomposition using the approximate coefficients, different levels of DWT coefficients can be obtained.

Researchers have used different mother wavelet functions and levels of decomposition in bearing and gear fault diagnosis. In this study, db5 is used as a mother wavelet function and 3rd level decomposition for gear fault diagnosis.

2. 2. Empirical Mode Decomposion (EMD) EMD is an effective adaptive signal processing technique. Finite numbers of intrinsic mode functions (IMF) can be obtained by decomposing complicated data [24, 25]. A number of extrema and zero crossings are same for linear or non-linear mode. The procedure for obtaining these IMF's from a given signal is as follows.

Firstly, all the local extrema are determined, and a cubic spline curve is employed to connect all the local maxima to obtain the upper envelop and all local minima to obtain the lower envelop. The entire signal should be enclosed by these upper and lower envelops. A mean of upper and lower envelop is m_1 and the difference between the *x(t)* and m_1 gives the first component h_1

$$x(t)-m_1=h_1 \tag{4}$$

If h_1 is an IMF, then h_1 is the first component of *x(t)*. If h_1 is not an IMF, the above procedure is repeated considering h_1 as original signal

$$h_1-m_{11}=h_{11} \tag{5}$$

After sifting for *k* times, h_{1k} becomes an IMF, that is

$$h_{1(k-1)}-m_{1k}=h_{1k}=c_1 \tag{6}$$

c_1 *is the* first IMF component obtained from the original. Separating c_1 from *x(t)*, We get

$$r_1=x(t)-c_1 \tag{7}$$

Now, r_1 is considered as the original signal and the process described above is repeated *n* time to obtain *n*-IMFs of signal *x(t)*.

$$x(t)=\sum_{j=1}^{n} c_j+r_n \tag{8}$$

The above procedure is repeated till r_n becomes a monotonic function from which no more IMFs can be drawn out.

Thus, decomposition of the signal results into *n*-empirical modes and a residue r_n. Where r_n is the mean trend of *x(t)*. The IMFs contain different frequency bands and the components of frequency included in each frequency band are different, and they change with the variation of signal *x(t)*, r_n represents the central tendency of signal *x(t)*.

2. 3. Machine Learning Techniques In the present work Support Vector Machine, Multi-Layer Perceptron and Random Forest, widely used and popular ML techniques are utilized for the gear fault classification.

2. 3. 1. Support Vector Machine (SVM) SVM is a supervised ML technique employed for the classification and regression in the small sample dataset. In SVM classifier, data is separated by a decision boundary known as the hyperplane such that the margin of separation between two classes is maximized and points that decide the margin are called as support vectors as shown in Figure 2.

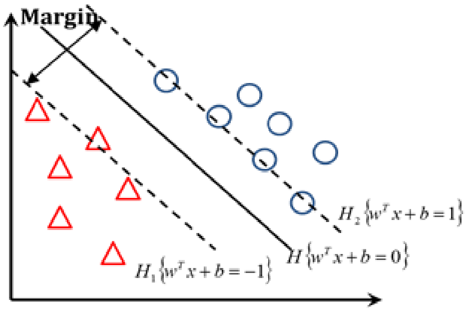


Figure 2. Hyperplane classifying two classes

This is done by minimizing the quadratic function under linear inequality constraints. Consider a training sample set $\{(x_i, y_i)\}$; $i=1$ to N, where N is the total number of samples. It is to determine the separation plane with the smallest generalization error. The labels associated with the two classes, i.e. triangle and circle class are $y_i=-1$ and $y_i=+1$ respectively. Slack variables are considered $\xi_i \geq 0$ for non-separable data. The hyperplane $f(x)=0$ that separates the given data can be obtained as a solution to the following optimization problem [4].

$$\text{Minimize } \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{N}\xi_i \text{ Subject to} \quad \begin{cases} y_i\left(w^T x_i+b\right)\geq 1-\xi_i \\ \xi_i \geq 0, i=1,2,\dots N \end{cases} \tag{9}$$

where, C is a constant representing error penalty

2. 3. 2. Multi-Layer Perceptron (MLP) **MLP** is a type of Back Propagation Artificial Neural Network which includes an input layer, output layer and two or more weighted perceptron (hidden layers) [8]. The targets of each layer are characterized by a function, and subsequently, the targets of preceding layers are served as inputs to the succeeding layer. Each layer consists of a specific number of neurons which carry information to neurons of subsequent layers.

For example, consider a training set (X_i^j, Y_i^j) with N features and M samples where $i=1, 2, 3, \dots M$ and $j= 1, 2, 3, \dots N$. The first layer will consist of N neurons with each j^{th} neuron containing M samples of j^{th} feature. Neurons in subsequent layers are also fixed accordingly (with an additional node neuron), but the last layer will have only one neuron for regression model and for the classification problem number of neurons will be equal to a number of classes. The parameter which controls the mapping of information from one layer to another, say Ø_j is multiplied with training samples X_i^j and the resultant matrix multiplication is fed as a parameter to the function. For example, a neuron a_1 of the second layer is represented as:

$$a_1 = f\left(\phi_1^T, X_j^1\right) \tag{10}$$

where, f is the sigmoid function.
Hence, the prediction is given by a_L. To minimize the loss function, this is followed by back-propagation method wherein another parameter say ∂ is computed for each layer with the help of gradient descent. The values of ∂ are computed from the last layer to the first layer and are updated by means of gradient descent. Also, for the last layer:

$$\partial_L = \left(a_L - y\right) \tag{11}$$

where, y is the target.
The following layers are given by:

$$\partial_{L-1} = \left[\phi_{L-1}\right]^T * \partial_L * \left[\left(a_{L-1}\right)*\left(1-a_{L-1}\right)\right] \tag{12}$$

where, $\left[\left(a_{L-1}\right)*\left(1-a_{L-1}\right)\right]$ is the differentiated term of the sigmoid function. Hereafter, the product of two ∂ and a are computed to add over each iteration to get the gradient which will minimize the loss function. This process is done with regularization.

Hence the 'accumulator' i.e. $\frac{d\left(L\left(a_L, y\right)\right)}{dy}$ (i.e. the derivative of loss function) is given by:

$$D_{ij} = \frac{1}{M}\left(\Delta_{ij}^{(L)}\right) + \alpha * \phi_{ij}^{(L)} \tag{13}$$

where, α is the regularization parameter and M is the number of samples.
Also, initially $\Delta_{ij}^{(L)}=0$ and it is updated as follows:

$$\Delta^{(L)} = \Delta^{(L)} + \partial_{L-1} * \left[a_L\right]^T \tag{14}$$

The accumulator i.e. D is added over each step to minimize the loss function and get better value of prediction i.e. a_L.

2. 3. 3. Random Forest (RF) is a polpular ensemble machine learning technique, in which many decision trees are built on bootstrapped sample same as bagging from training dataset and to get the output prediction, prediction from each tree is averaged in case of regression and majority voting in case of classification. RF is better than a single decision tree and reduces the over-fitting by averaging. High variance and low bias of decision trees make it unstable. RF consists of bootstrap aggregating (bagging) with a randomized selection of features at each split. A bootstrap aggregating algorithm improves the stability and accuracy of an individual predictive model [18,19].

Consider A, number of trees generated by the random forest. Using bootstrap sampling, select the subset from the training dataset. For each subset grow tree $\{t_1(x), t_2(x),\ldots,t_A(x)\}$ where $x \in \{x_1, x_2, \ldots, x_n\}$. x is an n-dimensional feature vector, and this feature vector is prepared by randomly selecting n features from a total of N features. Using the Gini impurity (I_g) pick a node e out of those n variables for the best split point.

$$I_{G(p)} = 1 - \sum_{a=1}^{b} R_a^2 \tag{15}$$

where, a is the number of classes (a=1, 2, 3, …., b) and R_a is the fraction of instances tagged with class a. Then split the node into two daughter nodes and repeat the above procedure for growing A number of trees. $Y_1, Y_2,\ldots Y_k$ is the output of these trees, where $k \in \{1, 2, \ldots, A\}$ is the prediction for a classified object by the k^{th} tree, and a collection of all individuals make a final classification decision.

3. SIMULATION DRIVEN APPROACH

In the present work, a model developed by Bartelmus [20] is used to obtain the vibration response. The model includes both torsional and lateral motions. In this gearbox, the casing is assumed as rigid so as vibration propagates linearly along the casing. Figure 3 depicts the model used in this work.

In this model the system is rotated by a motor and a load torque is applied at the output. M_1 and M_2 are the motor and load torque, respectively. Flexible couplings are used to connect the motor shaft and input shaft on which pinion is mounted, and load shaft and shaft on which gear is mounted. Shafts containing pinion and gear are mounted on bearings, and these bearings are mounted on the rigid casing. A model consists of two parameters stiffness and damping and includes both linear and rotational (lateral and torsional) equations of motion. Equations (16)-(23) represent the equations of motion

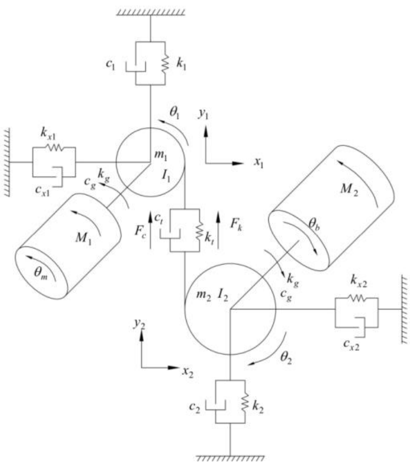


Figure 3. Single-stage spur gearbox model [20]

for the system shown in Figure 3. x - direction vibration response is a free response (Equations (16) and (17)). When the system is stable response in this direction will disappear; therefore, vibration in y - direction is considered here to get the response of healthy and faulty gears. Single fault on a single pinion tooth is considered for obtaining vibration response for faulty gear condition. Vibration response of faulty gears includes responses of gearbox having faults such as spalled, cracked and chipped gear tooth. Firstly, TVMS of normal and faulty conditions is calculated using the potential energy method presented in [26-28]. The value of calculated TVMS which act as an internal excitation, is considered while obtaining the vibration response.

$$m_1\ddot{x}_1 = -k_{x1}x_1 - c_{x1}\dot{x}_1 \tag{16}$$

$$m_2\ddot{x}_2 = -k_{x2}x_2 - c_{x2}\dot{x}_2 \tag{17}$$

$$m_1\ddot{y}_1 + c_1\dot{y}_1 + k_1y_1 = -F_k - F_c \tag{18}$$

$$m_2\ddot{y}_2 + c_2\dot{y}_2 + k_2y_2 = F_k + F_c \tag{19}$$

$$I_1\ddot{\theta}_1 = k_p(\theta_m - \theta_1) + c_p(\dot{\theta}_m - \dot{\theta}_1) - R_{b1}(F_k + F_c) \tag{20}$$

$$I_2\ddot{\theta}_2 = R_{b2}(F_k + F_c) - k_g(\theta_2 - \theta_b) - c_g(\dot{\theta}_2 - \dot{\theta}_b) \tag{21}$$

$$I_m\ddot{\theta}_m = M_1 - k_p(\theta_m - \theta_1) - c_p(\dot{\theta}_m - \dot{\theta}_1) \tag{22}$$

$$I_b\ddot{\theta}_b = -M_2 + k_g(\theta_2 - \theta_b) + c_g(\dot{\theta}_2 - \dot{\theta}_b) \tag{23}$$

$$F_k = k_t(R_{b1}\theta_1 - R_{b2}\theta_2 + y_1 - y_2) \tag{24}$$

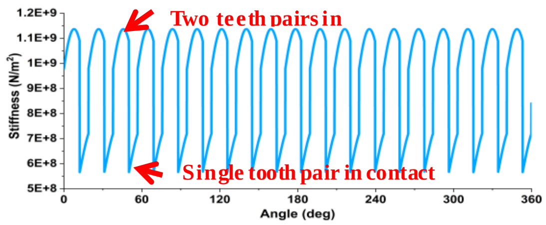
$$F_c = c_t(R_{b1}\dot{\theta}_1 - R_{b2}\dot{\theta}_2 + \dot{y}_1 - \dot{y}_2) \tag{25}$$

The gear parameters utilized in this model to obtain the simulated vibration data are given in Table 1. Figure 4 (a-b) shows the sample of TVMS obtained for normal and faulty gear conditions respectively for one rotation of pinion. The equations of motion (Equations (18)-(23)) are solved simultaneously using ODE15s solver function in Matlab to obtain the vibration response. The vibration response is obtained for different load and speed conditions. Figure 5 (a-b) shows the vibration response obtained at 1800rpm for normal and faulty gear condition, respectively for 0.1 s.

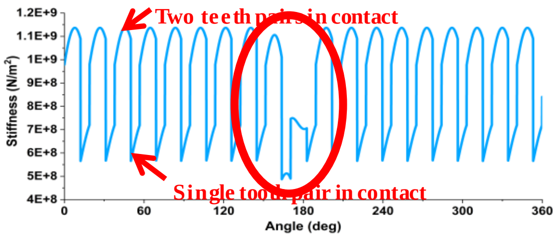
To increase the exactness of simulated vibration signals with the actual field data noise is added to the simulated response. A pink noise (1/f power spectrum) is added, as pink noise is present everywhere in nature, electronics, machinery, and numerous other fields [29]. Chen et al. [30] used this pink noise power spectrum to generate the bearing signals. Randomly this noise is

TABLE 1. Gear parameters for simulated data

Gear Parameter	Value
Number of teeth on Pinion/Gear	$N_p = 19,\ N_g = 48$
Pressure angle	20^0
Diametral Pitch	$P = 0.2032\ m^{-1}$
Width of Teeth	$L = 0.16\ m$
Contact ratio	$C_r = 1.6456$
Young's Modulus	$E = 2.068 \times 10^{11}\ Pa$
Poisson's Ratio	$\upsilon = 0.3$
Mass of the Pinion/Gear	$m_p = 0.96\ kg,\ m_g = 2.88\ kg$
Mass moment of inertia of the Pinion/Gear	$I_p = 4.3659 \times 10^{-4}\ kgm^2$ $I_g = 8.3602 \times 10^{-4}\ kgm^2$
Mass moment of inertia of the motor/Load	$I_m = 0.0021\ kgm^2$ $I_b = 0.0105\ kgm^2$
Tortional stiffness of the coupling	$k_p = k_g = 4.4 \times 10^4\ Nm/rad$
Damping coefficient of the coupling	$c_p = c_g = 5.0 \times 10^5\ Nm/rad$
Radial stiffness of the bearing	$k_1 = k_2 = 6.56 \times 10^7\ N/m$
Damping coefficient of the bearing	$c_1 = c_2 = 1.8 \times 10^5\ Ns/m$
Base circle radius of Pinion/Gear	$R_{b1} = 0.02834\ m,$ $R_{b2} = 0.0716\ m$



(a) Gear with no fault



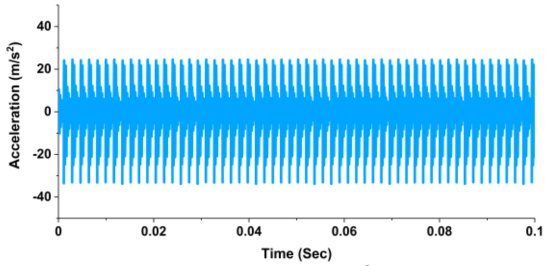
(b) Gear with fault

Figure 4. Time Varying Mesh Stiffness (TVMS)

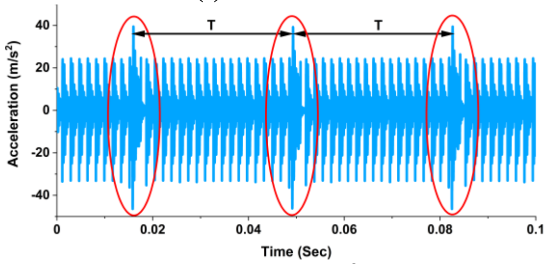
added to simulated vibration signals so that S/N ratio can range from 1 to 30.

4. EXPERIMENTAL TEST RIG

To validate the simulation-driven approach ML algorithms trained using the simulated training data is to be tested using the experimental data. For the broader



(a) Gear with no fault



(b) Gear with fault

Figure 5. Simulated Vibration Response at 1800rpm

applicability of this approach in gear fault detection, gearbox test rig having different gear geometric parameters than the one used to obtain the simulated data is to be employed for generating the testing/validation data. Figure 6 shows the gearbox experimental setup used to collect the validation data. Validation data for normal and faulty conditions are collected at different load and speed conditions. The gear test rig has spur gear having full depth involute profile and has 25 and 61 number of teeth on pinion and gear respectively and 20^0 pressure angle, 4 mm module, 20 mm width. An electric motor drives the pinion, and a load is applied at the driven end using the magnetic brake. Tri-axial accelerometer mounted on the gearbox is connected to the OROS data acquisition system to collect the vibration data. Validation or testing data contains an equal number of samples of normal and faulty gear condition.

5. FEATURE ETRACTION AND SELECTION

Numerous features have been proposed in the literature for gear fault diagnosis using time domain, frequency

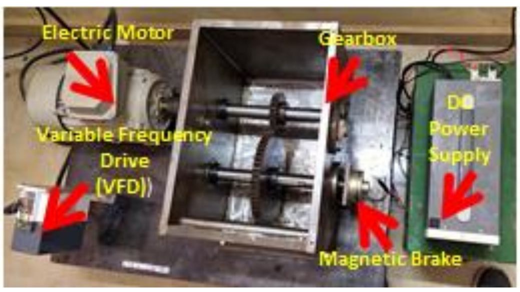


Figure 6. Gearbox experimental test rig

domain, wavelet decomposition and Empirical Mode Decomposition etc. In the present work, wavelet decomposition and EMD are used to process vibration signals. A signal is decomposed in 3 levels using the db5 mother wavelet function, and also the signal is processed using EMD to obtain the IMFs. Statistical features Skewness, Kurtosis, Root Mean Square (RMS) Crest Factor, Shape Factor, Impulse Factor and Clearance Factor are extracted from all three decomposition levels of wavelet and IMF's of EMD. Equations (26)-(32) presents the mathematical expression for calculating these features.

In literature different methods of feature reduction are available such as Principal Component Analysis (PCA), Factor Analysis (FA), Independent component analysis (ICA), High correlation between two columns. These methods were utilized for feature reduction but application of first three methods i.e. PCA, FA and ICA didn't yield satisfactory results. As PCA is a widely used feature reduction technique, results obtained using this is presented here.

5. 1. Principal Component Aanalysis (PCA) PCA is a feature dimensionality reduction technique which is used for the compression and classification of the data. The dimensionality of data set is reduced by finding the new variables that are smaller than the original data set and retain most of the information. These new variables are called as the principal components which are uncorrelated.

In this paper for feature reduction high correlation between two columns method is adopted. In which correlation between the extracted features is determined. Features having high correlation coefficient are linearly dependent and have the same effect as the dependent variable. Therefore, one of the two features having high correlation coefficients is removed. The remaining features are then fed one by one to ML algorithms and a trial and error approach is adopted to select the features giving best results. Table 2 presents a list of selected features. The extracted features from both simulated and experimental data are normalized to the zero mean and to a range of ±1 to prepare training and testing data set.

$$\boldsymbol{Skewness(SK_a)} = \frac{\sum_{n=1}^{N}(a(n)-\mu_a)^3}{(N-1)\sigma_a{}^3} \quad (26)$$

$$\boldsymbol{Kurtosis(K_a)} = \frac{\sum_{n=1}^{N}(a(n)-\mu_a)^4}{(N-1)\sigma_a{}^4} \quad (27)$$

$$\boldsymbol{Root\ Mean\ Square(a_{rms})} = \sqrt{\frac{\sum_{n=1}^{N}(a(n))^2}{N-1}} \quad (28)$$

$$\boldsymbol{Crest\ Factor} = \frac{a_p}{a_{rms}} \quad (29)$$

$$\boldsymbol{Shape\ Factor\ (SF_a)} = \frac{a_{rms}}{\frac{1}{N}\sum_{n=1}^{N}|a(n)|} \quad (30)$$

TABLE 2. Selected Extracted Features

Feature Extraction Method	Features
Wavelet Transform	1st level Detailed Coefficients - Crest Factor 3rd Level Detailed Coefficients – Skewness 3rd Level Approximation Coefficients - Crest Factor
Empirical Mode Decomposition	IMF2 – RMS IMF4 – RMS, Shape Factor, Kurtosis IMF6 – Clearance Factor, Impulse Factor, Kurtosis IMF7 – Kurtosis IMF8 – Shape Factor

$$\boldsymbol{Impulse\ Factor(IF_a)} = \frac{a_p}{\frac{1}{N}\sum_{n=1}^{N}|a(n)|} \quad (31)$$

$$\boldsymbol{Clearance\ Factor(CLF_a)} = \frac{a_p}{\left(\frac{1}{N}\sum_{n=1}^{N}\sqrt{|a(n)|}\right)^2} \quad (32)$$

6. RESULTS AND DISCUSSION

A simulated vibration data is obtained for different load and speed conditions, as described in section 3. While getting the vibration data single fault is considered on a single pinion tooth. Figure 4(a-b) shows the TVMS for the healthy and faulty gear, respectively, for the one pinion revolution. The contact ratio of gear is around 1.6, i.e. for 60% of duration two pairs of teeth are in contact and hence the increase in TVMS in this region. Figure 4(a) shows the TVMS for the healthy gear, which shows a similar pattern for all the tooth engagements. However, for the faulty case (Figure 4(b)), when faulty tooth engages decrease in the TVMS is observed for the duration of engagement of faulty tooth. The faulty tooth contact region is highlighted in Figure 4(b). This change in the TVMS affects the vibration response of the gear. This variation in vibration response of normal and faulty gear is clearly shown in Figure 5(a-b), respectively. In Figure 5(b) the part of the vibration response whenever faulty tooth engages is highlighted. As the fault is considered on the pinion, for every pinion rotation this faulty tooth will come in contact and hence the change in vibration response. *T* in Figure 5(b) represents the time period for one pinion rotation. To improve the exactness of the simulated data towards the actual data noise is added. These signals are then processed using the DWT and EMD, and total 84 statistical features are extracted from all levels of decomposition of DWT and IMF's of EMD. To reduce the dimensionality of the dataset

Principal Component Analysis (PCA) was employed, and ML algorithms were tested for a different number of principal components, and its classification accuracies were obtained. Figure 7 shows the testing accuracy obtained for all three ML techniques for a different number of principal components. This figure indicates that accuracy obtained using PCA for a different number of components was between 65 and 78% for all three classifiers. These accuracy values obtained using PCA were not satisfactory. Therefore, to reduce the dimensionality of the dataset, a different method is utilized. In this, a correlation among the extracted features is determined. One of the two features having high correlation coefficients is removed as features having high correlation coefficient are linearly dependent and have the same effect as the dependent variable. Remaining features are then fed one by one to train the ML algorithm and features giving best results are selected. Table 2 presents a list of selected features.

The ML algorithms are trained using these selected features, and ten cross-validation accuracy is obtained. The ML algorithm's parameter tuning is done based on simulated data and a grid search algorithm. For the wider applicability of the simulation-driven approach, ML algorithms trained using simulated data is validated by experimental data generated using the gearbox test rig having different gear geometric parameters than the one used to create the simulated data. The classifier accuracy is a metric utilized in this study for the evaluation of the classification algorithm. Figure 8 shows training and testing accuracies obtained for SVM, MLP and RF machine learning algorithms.

Training accuracy of more than 90% and testing accuracy of more than 87% is obtained for all the three classifiers. The highest training and testing accuracy of 95 and 90% is obtained using the RF classifier. The values of accuracies for all three classifiers imply the applicability of the simulation-driven approach for gear fault detection. The performance of the proposed simulation-driven approach is compared with the existing data-driven approach. ML algorithms are trained and tested using experimental data in the data-driven

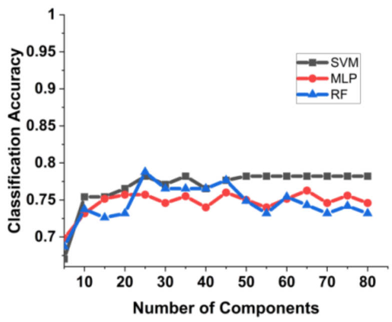


Figure 7. Testing Accuracy of ML algorithms for different principal components

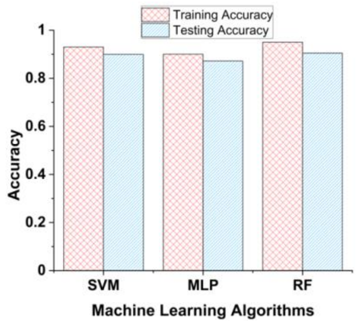


Figure 8. Accuracies by using Simulation Driven Approach

approach. Figure 9 shows the accuracies obtained using the data-driven approach. The accuracies obtained for all the three classifiers using data-driven and simulation-driven approach are comparable, and there is very little difference in their accuracies. Therefore, the simulation-driven approach can be applied for the gear fault classification when actual historical data is not available.

Classifier accuracy is considered as a metric for the evaluation of the ML algorithm. However, for the complete performance evaluation of the classification algorithm precision and recall are also evaluated. Precision is the ratio of correctly classified positive cases to the total number of positively classified cases and recall is the ratio of correctly classified positive cases and the total number of actual positive cases. Figure 10 shows the values of precision and recall for each classifier. For all the three classifiers precision and recall values are more than 87%. The higher values of precision mean the less number of instances are classified as faulty when it is normal, that is fewer chances of false alarms for the maintenance and avoid undue maintenance task. The higher value of recall increases the chances of positive alarms as lesser instances are classified as normal when it is faulty, to carry out the maintenance task when it is actually required. All the three classifiers show the considerable values of precision and recall (Figure 10), which increases the confidence in the simulation-driven approach presented in this paper for gear fault detection. Precision and recall values for the RF classifier are slightly higher than the other classifiers.

Another metric used for the evaluation of the binary classification is the Area under the receiver operating characteristic curve (AUC). The curve of true positive rate Vs false positive rate at different classification threshold values is known as the Receiver Operating Characteristic curve abbreviated as ROC. Figure 11 depicts the ROC curve for the three classifiers. The AUC values are 0.97, 0.92 and 0.98 for SVM, MLP and RF, respectively. A higher value of AUC means that classifier is confident that randomly selected positive instance is positive than the randomly selected negative instances as positive. The ROC curve (Figure 11) and the AUC values obtained for three classifiers shows the agreement with

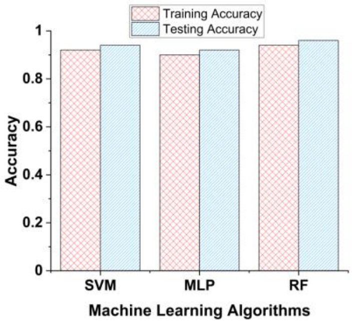


Figure 9. Accuracies by using Data Driven Approach

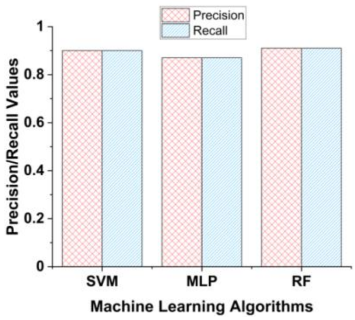


Figure 10. Precision and Recall for SVM, MLP and RF

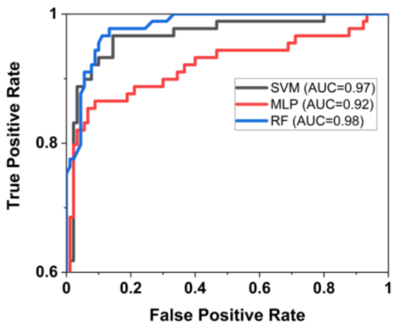


Figure 11. ROC curve for SVM, MLP and RF classifiers

the applicability of the simulation-driven approach for the gear fault classification.

7. CONCLUSION

Condition monitoring and fault diagnosis of the gearbox using the machine learning techniques utilize the data-driven approach, but the requirement of in-service/experimental data for the training of ML algorithm has prevented its widespread application in the industry. Also, for the better training of ML algorithm high diverse data is required. The present paper implements a simulation-driven approach for the spur gear fault detection and validates this approach using the experimental data. Simulated vibration response is generated using a gearbox dynamic model to produce the extensive variety of data of at different operating and fault condition. Firstly, TVMS is calculated for normal and faulty gear condition. TVMS is reduced when faulty gear tooth engages, and this results in the change in vibration response. ODE15s solver function in Matlab is employed to solve the equations of motion to obtain the simulated vibration response. This simulated vibration data is not having any noise, but in actual practical application the signal is masked with environmental noise. Therefore, a pink noise is added to improve the exactness of the simulated vibration signals to the actual vibration signals. Pink noise is added because it is present everywhere in nature, electronics, machinery, and numerous other fields. These signals are then processed using the EMD and DWT signal processing techniques. Features are extracted from simulated data and fed for the training of ML algorithm. For the wider applicability of the simulation-driven approach, ML algorithm trained using the simulated data is validated by using experimental data collected from a test rig having different gear parameters. The results show that the simulation-driven approach gives the considerable training and testing accuracy for all the three classifiers. This approach is then compared with the existing data driven approach and the results obtained are comparable. However, the performance of the this approach can be improved by considering the dynamic model with high degrees of freedom and considering more effect such as friction, gyroscopic effect etc. In the present era of Industry 4.0 the proposed method has the potential to improve the machine learning based condition monitoring of the gearbox using simulated data. However, the simulated data cannot replace the experimental data, but it can serve as a starting point for the gear fault diagnosis when no historical data is available. The implementation of this approach will help in monitoring the condition of the gearbox from the day of installation.

8. REFERENCES

1. Salameh, J. P., Cauet, S., Etien, E., Sakout, A., & Rambault, L., "Gearbox condition monitoring in wind turbines: A review", ***Mechanical Systems and Signal Processing***, Vol. 111, (2018), 251-264. DOI: 10.1016/j.ymssp.2018.03.052
2. Wang, K., "Intelligent predictive maintenance (IPdM) system–Industry 4.0 scenario", ***WIT Transactions on Engineering Sciences***, Vol. 113, (2016), 259-268. DOI: 10.2495/IWAMA150301
3. Karimzadeh, A., & Shoghli, O., "Predictive Analytics for Roadway Maintenance: A Review of Current Models, Challenges, and Opportunities", ***Civil Engineering Journal***, Vol. 6, No. 3, (2020), 602-625. DOI: 10.28991/cej-2020-03091495
4. Damuluri, S., Islam, K., Ahmadi, P., & Qureshi, N. S., "Analyzing Navigational Data and Predicting Student Grades Using Support Vector Machine", ***Emerging Science Journal***, Vol. 4, No. 4, (2020), 243-252. DOI: 10.28991/esj-2020-01227

5. Liang X, Zuo MJ, Feng Z., "Dynamic modeling of gearbox faults: A review", ***Mechanical Systems and Signal Processing***, Vol. 98, (2018), 852-876. DOI: 10.1016/j.ymssp.2017.05.024

6. Lei Y, Yang B, Jiang X, Jia F, Li N, Nandi AK., "Applications of machine learning to machine fault diagnosis: A review and roadmap", ***Mechanical Systems and Signal Processing***, Vol. 138, (2020), 106587. DOI: 10.1016/j.ymssp.2019.106587

7. Samanta, B., "Gear fault detection using artificial neural networks and support vector machines with genetic algorithms", *Mechanical systems and signal processing*, Vol. 18, No. 3, (2004), 625-644. DOI: 10.1016/S0888-3270(03)00020-7

8. Samanta B, Al-Balushi KR, Al-Araimi SA., "Artificial neural networks and genetic algorithm for bearing fault detection", ***Soft Computing***, Vol. 10, No. 3, (2006), 264-271. DOI: 10.1007/s00500-005-0481-0

9. Tyagi, S., Panigrahi, S. K., "A DWT and SVM based method for rolling element bearing fault diagnosis and its comparison with Artificial Neural Networks", ***Journal of Applied and Computational Mechanics***, Vol. 3, No. 1, (2017), 80-91. DOI: 10.22055/JACM.2017.21576.1108

10. Sanz J, Perera R, Huerta C., "Gear dynamics monitoring using discrete wavelet transformation and multi-layer perceptron neural networks", ***Applied Soft Computing***, Vol. 12, No. 9, (2012), 2867-2878. DOI: 10.1016/j.asoc.2012.04.003

11. Shen Z, Chen X, Zhang X, He Z., "A novel intelligent gear fault diagnosis model based on EMD and multi-class TSVM", ***Measurement***, Vol. 45, No. 1, (2012), 30-40. DOI: 10.1016/j.measurement.2011.10.008

12. Shao R, Hu W, Huan X, Chen L., "Multi-damage feature extraction and diagnosis of a gear system based on higher order cumulant and empirical mode decomposition", ***Journal of Vibration and Control***, Vol. 21, No. 4, (2015), 736-754. DOI: 10.1177/1077546313482342

13. Li, X., Han, L., Xu, H., Yang, Y. and Xiao, H., "Rolling Bearing Fault Analysis by Interpolating Windowed DFT Algorithm", ***International Journal of Engineering, Transactions A: Basics***, Vol. 32, No. 1, (2019), 121-126. DOI: 10.5829/ije.2019.32.01a.16

14. Dhamande LS, Chaudhari MB., "Compound gear-bearing fault feature extraction using statistical features based on time-frequency method", ***Measurement***, Vol. 125, (2018), 63-77. DOI: 10.1016/j.measurement.2018.04.059

15. Attaran, B., Ghanbarzadeh, A. and Moradi, S., "A Novel Intelligent Fault Diagnosis Approach for Critical Rotating Machinery in the Time-frequency Domain", ***International Journal of Engineering, Transactions A: Basics***, Vol. 33, No. 4, (2020), 668-675. DOI: 10.5829/IJE.2020.33.04A.18

16. Bajric R, Zuber N, Skrimpas GA, Mijatovic N., "Feature extraction using discrete wavelet transform for gear fault diagnosis of wind turbine gearbox", Shock and Vibration, (2016), DOI: 10.1155/2016/6748469

17. Han T, Jiang D., "Rolling bearing fault diagnostic method based on VMD-AR model and random forest classifier", ***Shock and Vibration***, (2016), DOI: 10.1155/2016/5132046

18. Cerrada M, Zurita G, Cabrera D, Sánchez RV, Artés M, Li C., "Fault diagnosis in spur gears based on genetic algorithm and random forest", ***Mechanical Systems and Signal Processing***, Vol. 70-71, (2016), 87-103. DOI: 10.1016/j.ymssp.2015.08.030

19. Patil S, Phalle V., "Fault Detection of Anti-friction Bearing using Ensemble Machine Learning Methods", ***International Journal of Engineering, Transactions B: Applications***, Vol. 31, No. 11, (2018), 1972-1981. DOI: 10.5829/ije.2018.31.11b.22

20. Bartelmus W., "Mathematical modelling and computer simulations as an aid to gearbox diagnostics", ***Mechanical Systems and Signal Processing***, Vol. 15, No. 5, (2001), 855-871. DOI: 10.1006/mssp.2001.1411

21. Howard I, Jia S, Wang J., "The dynamic modelling of a spur gear in mesh including friction and a crack", ***Mechanical Systems and Signal Processing***, Vol. 15, No. 5, (2001), 831-853. DOI: 10.1006/mssp.2001.1414

22. Abouel-seoud SA, Dyab ES, Elmorsy MS., "Influence of tooth pitting and cracking on gear meshing stiffness and dynamic response of wind turbine gearbox", ***International Journal of. Science and Advanced. Technology***, Vol. 2, No. 3, (2012), 151-165. DOI:

23. Mohammed, O. D., Rantatalo, M., & Aidanpää, J. O., "Dynamic modelling of a one-stage spur gear system and vibration-based tooth crack detection analysis", ***Mechanical Systems and Signal Processing***, Vol. 54-55, (2015), 293-305. DOI: 10.1016/j.ymssp.2014.09.001

24. Parey A, El Badaoui M, Guillet F, Tandon N., "Dynamic modelling of spur gear pair and application of empirical mode decomposition-based statistical analysis for early detection of localized tooth defect", ***Journal of Sound and vibration***, Vol. 294, No. 3, (2006), 547-561. DOI: 10.1016/j.jsv.2005.11.021

25. He Q, Li P, Kong F., "Rolling bearing localized defect evaluation by multiscale signature via empirical mode decomposition", ***Journal of Vibration and Acoustics,*** Vol. 134. No. 6, (2012). DOI: 10.1115/1.4006754

26. Handikherkar V.C., Phalle V.M., "Vibration Analysis Based Spalling Defect Severity Assessment of Spur Gearbox Using a Dynamic Model", ***Developments and Novel Approaches in Nonlinear Solid Body Mechanics***, Vol. 130, (2020), 363-375. DOI: 10.1007/978-3-030-50460-1_20

27. Wu S, Zuo MJ, Parey A., "Simulation of spur gear dynamics and estimation of fault growth", ***Journal of Sound and Vibration***, Vol. 317, No. (3-5), (2008), 608-624. DOI: 10.1016/j.jsv.2008.03.038

28. Chaari F, Baccar W, Abbes MS, Haddar M., "Effect of spalling or tooth breakage on gearmesh stiffness and dynamic response of a one-stage spur gear transmission", ***European Journal of Mechanics-A/Solids***, Vol. 27, No. 4, (2008), 691-705. DOI: 10.1016/j.euromechsol.2007.11.005

29. Bak, P., Tang, C., & Wiesenfeld, K., "Self-organized criticality: An explanation of the 1/f noise", ***Physical Review Letters***, Vol. 59, No. 4, (1987), 381. DOI: 10.1103/PhysRevLett.59.381

30. Chen, B., Yan, Z., & Chen, W., "Defect detection for wheel-bearings with time-spectral kurtosis and entropy", ***Entropy***, Vol. 16, No. 1, (2014), 607-626. DOI: 10.3390/e16010607

Persian Abstract

چکیده

نظارت بر شرایط مبتنی بر یادگیری ماشین (ML) و تشخیص عیب تجهیزات صنعتی سناریوی فعلی برای نگهداری در دوره Industry-4.0 است. استفاده از تکنیک های ML برای تشخیص خودکار عیب ، خرابی غیرمنتظره سیستم را به حداقل می رساند. با این حال ، این تکنیک ها به شدت به داده های تاریخی تجهیزات برای آموزش آن متکی هستند که کاربرد گسترده آن را در صنعت محدود می کند. از آنجا که داده های تاریخی برای هر ماشین صنعتی در دسترس نیست و تولید داده ها به صورت آزمایشی برای هر شرایط خطا قابل استفاده نیست. بنابراین ، این چالش برای استفاده از چرخ دنده با نقص دندان حل می شود. در این مقاله ، الگوریتم های ML با استفاده از داده های ارتعاش شبیه سازی شده گیربکس آموزش داده شده و با داده های تجربی آزمایش می شوند. داده های شبیه سازی شده برای گیربکس با شرایط کارکرد و خطای مختلف تولید می شود. از یک مدل دینامیکی گیربکس برای تولید داده های ارتعاش شبیه سازی شده برای شرایط دنده نرمال و معیوب استفاده شده است. یک نویز صورتی به داده های شبیه سازی شده اضافه می شود تا دقت داده های درست را بهبود بخشد. علاوه بر این ، این داده های شبیه سازی شده با استفاده از تجزیه حالت تجربی و تغییر شکل موجک گسسته پردازش می شوند و ویژگی ها استخراج می شوند. این ویژگی ها سپس به آموزش تکنیک های مختلف ML کاملاً ثابت مانند ماشین بردار پشتیبان ، جنگل تصادفی و پرسپترون چندلایه منتقل می شود. برای تأیید این روش ، الگوریتم های ML آموزش دیده با استفاده از داده های تجربی آزمایش می شوند. نتایج بیش از ۸۷٪ دقت با هر سه الگوریتم را نشان می دهد. عملکرد مدل آموزش دیده با استفاده از منحنی دقت ، فراخوان و ROC ارزیابی می شود. این معیارها نتایج مثبتی را برای کاربرد این روش در تشخیص عیب دنده نشان می دهد.