Rotating blade faults classification of a rotor-disk-blade system using artificial neural network

Abdullahi Abubakar Mas’ud1, Ahmad Jamal2, Surajuddeen Adewusi3, Arunachalam Sundaram4
1,4Department of Electrical and Electronic Engineering, Jubail Industrial College, Kingdom of Saudi Arabia
1,2Prince Saud bin Thunayan Research Centre, Royal Commission for Jubail, Kingdom of Saudi Arabia
2,3Department of Mechanical Engineering, Jubail University College, Kingdom of Saudi Arabia

ABSTRACT

In this paper, the artificial neural network (ANN) has been utilized for rotating machinery faults detection and classification. First, experiments were performed to measure the lateral vibration signals of laboratory test rigs for rotor-disk-blade when the blades are defective. A rotor-disk-blade system with 6 regular blades and 5 blades with various defects was constructed. Second, the ANN was applied to classify the different x- and y-axis lateral vibrations due to different blade faults. The results based on training and testing with different data samples of the fault types indicate that the ANN is robust and can effectively identify and distinguish different blade faults caused by lateral vibrations in a rotor. As compared to the literature, the present paper presents a novel work of identifying and classifying various rotating blade faults commonly encountered in rotating machines using ANN. Experimental data of lateral vibrations of the rotor-disk-blade system in both x- and y-directions are used for the training and testing of the network.

Keywords: Artificial neural networks Blade faults Blade system Python Rotating machinery

This is an open access article under the CC BY-SA license.

Corresponding Author:
Abdullahi Abubakar Mas’ud
Department of Electrical and Electronic Engineering
Jubail Industrial College and Prince Saud bin Thunayan Research Centre, Jubail
P.O. Box, 10099, Kingdom of Saudi Arabia
Email: masud_a@jic.edu.sa

1. INTRODUCTION

For competitiveness and high quality operating requirements in industries, the need for effective rotating machines has been significant. This can be achieved in part through continuous monitoring of operating conditions to detect any fault before the rotating machines break down. Faults in rotating blades are very difficult to detect by direct blade vibration measurements and evaluation. This is because blades often operate in very harsh environments. In addition, no significant results have been obtained in the quest for a more efficient artificial intelligence system for real time monitoring of these faults yet.

The performance and reliability of spinning machines are maintained by constantly tracking their working conditions in order to detect any defects before causing a significant problem or failure of the system that may impact output targets. There are research articles on the rotor vibration response [1]-[3], that are usually monitored in the industry and analysed for identification of faults. These efforts have resulted in the development of knowledge base and standard procedures to identify machine faults such as shaft misalignment and unbalance, rotor eccentricity, blade crack, twisting, and rubbing. Many industrial rotating machines have rotors, disks, blades, and vanes. Hence, it is important to monitor the dynamic response of the machine parts such as rotor-disk-blade system that potentially can develop machine faults contributed by
blade defects. Despite the fact that vibration response of rotor can be easily measured with the help of proximity probe and accelerometer, direct measurement of vibration response of the blades sometimes becomes very difficult and expensive since blades sometimes operate at very extreme conditions of temperature and pressure encountered in gas and steam turbines. Also, strain gauges mounted on the blades to measure blade deflection have short operating life due to these extreme conditions, blade aerodynamics interference. As a result, non-contacting and non-intrusive techniques to measure the blade vibrations due to blade defects have been developed. One such technique is used by Lawson and Ivey [4], in which blade tip timing is measured using tip clearance capacitance probes.

Many researchers have extracted the blade dynamic response from coupled responses of rotor-disk-blade systems [5], [6]. For this, they established a relationship between rotor torsional vibration and blade dynamic response. However, one study shows blade fault detection from rotor torsional vibrations. There are few studies that exhibit relationship between the rotor-blade system and blade vibration [7]. Few studies exhibit relationship between the rotor-blade system and blade vibration [6], [8]. One of these studies measured [8], analytically and experimentally, the blade vibrations through the lateral vibration of the support of a rotor-blade system. Even though specific blade faults were not investigated in this study, they provided the basis for blade fault detection from rotor lateral vibration.

A Substantial amount of work on the detection of various rotating machine faults such as blade distortion, shaft misalignment, rotor unbalance, rotor crack, using machine learning algorithms has been undertaken. A number of review articles are found on fault diagnosis of rotating machinery using the artificial intelligence (AI). Liu et al. [9] presented a comprehensive review of previous research on such faults using naïve Bayes, support vector machine, deep learning algorithms and the k-nearest neighbor (KNN). The benefits, constraints, and practical implications of such AI algorithms were also discussed. Another review paper by Lei et al. [10] presented the machine learning algorithms utilized over the years to detect machine faults. They categorized the algorithms coined as intelligent fault diagnosis (IFD) into (a) traditional machine learning theories such as the probabilistic graphical method (PGM), artificial neural network (ANN), support vector machine (SVM), k-nearest neighbor (KNN), (b) convolutional neural network (CNN), ResNet and deep learning theory such as deep belief network (DBN), and (c) transfer learning theories such as transfer component analysis (TCA) and the generative adversarial network (GAN). They indicated that almost all the IFD could be utilized to diagnose rotating machinery faults. Both review papers focus on different machine learning algorithms and state-of-the-art techniques useful in diagnosing different rotating machine faults rather than on the faults or any specific machine fault. Sánchez et al. [11] used the random forest and KNN machine learning algorithms as classifiers of the gearbox and bearing faults. Through these, a methodological structure for detecting multiple faults in rotating machinery was discussed. From the vibration signal, they calculated thirty features in the time domain through function ranking techniques such as relief, information gain and the chi-square.

Considering some specific rotating machine faults, one of the faults is the shaft unbalance. Out of many papers on this type of fault, a representative paper by Gohari and Eyadi [12] presented a comparative study of two machine learning algorithms namely the decision tree algorithms and the KNN to identify the shaft unbalance parameters such as eccentric radius, disc number and the eccentric mass value. They concluded on the basis of their results that KNN estimates the faults better than the Decision Tree algorithm. Another fault is the blade rub-impact. Prosvirin et al. [13] identified this fault in turbines using a deep neural network (DNN) and the Auto encoder-based nonlinear function approximation. They used a two-step process for the diagnosis. First, they used a deep under-complete de-noising auto encoder for determining the nonlinear function of the system and then they used the resulting residual signals from the original signals as input to the DNN to find the current state of the blade-rotor system.

Classical ANN algorithms have been utilized by many researchers in the past to diagnose various rotating machine faults. Zhong et al. [14] utilized single back-propagation (BP), single ellipsoid backpropagation, and hierarchical diagnostic artificial neural network (HDANN) algorithms to classify different rotor system faults. They showed that HDANN required only known fault patterns instead of training the entire network. The fault data were collected experimentally through a motor and test rig rotor system and then fed into the three algorithms for fault identification. They concluded that HDANN classified the faults most accurately. Nahvi and Esfahanian [15], used multi-layer feed forward neural network with nonlinear neurons (sigmoidal activation function) to identify 40 rotating machine faults. They could train the neural network from selected data and then utilize it to identify the faults. Adewusi and Albedoor [16] used multi-layer feed-forward ANN to detect crack in a rotor. Two- and three-layer networks with different number of neurons in each layer were tested to determine an optimal ANN configuration. Vibration signals of a normal rotor and that with a propagating crack were used to train and test the ANN using back-propagation (BP) algorithm. The results showed that the performance of a three-layer ANN was better than that of two-layer ANN in detecting the severity of the propagating crack in the rotor. Nguyen et al. [17], on
the other hand, worked on the gearbox system and developed a deep neural network based diagnosis model to diagnose and classify multi-level tooth cut gear faults operating under variable shaft speeds. They first implemented the adaptive noise control approach to remove any unwanted noise in the original vibration signal and then employed the stacked sparse autoencoder based deep neural network to diagnose and classify the gearbox faults.

In addition to the application of ANN to diagnose specific rotating machine faults, some previous works are reported here identifying rotor-bearing system faults using multi-layer BP neural network [18], rolling element bearing faults using time-domain features in ANN [19], minor faults such as scratch and hole in the bearings of induction motors using five machine learning algorithms and convolutional neural network [20], and compound bearing defects such as inner race, outer race and roller defect for embedded systems under varying rotational speeds using MobileNet-v2; a light state-of-the-art convolutional neural network model [21]. In addition, two articles related to rotating shafts addressed the detection of ball bearing faults producing vibrations in Rotating shafts using multi-layer feed forward neural network [22] and induction motor shaft misalignments using multi-scale entropy (MSE) coupled with back-propagation neural network [23].

Related to the present study of utilizing ANN for blade faults identification, some previous work can be found. Rao and Srinivas [24] used single layer neural network to locate the optimal position of rotational constraint along the length of a pre-twisted blade on a rotor system in order to increase the lowest natural frequency undergoing torsional vibration. Two papers are found that address the blade faults of variable rotational speed wind turbines. First is by Chen et al. [25] in which they used the deep learning approach by embedding attention mechanism into long-short-term memory neural network to identify the wind turbine blade imbalance fault due to the accumulation of ice on the blades. Second paper by Joshuva et al. [26] presented a study investigating various wind turbine blade faults using first variational mode decomposition (VMD) for experimental data signal pre-processing and then multi-layer perceptron (MLP) for blade fault classification. The experimental data of lateral vibrations of the shaft in only vertical axis, i.e., y-axis was acquired specific to wind turbines.

2. EXPERIMENT AND ESTABLISHMENT OF THE SIMULATION MODEL

2.1. Experimental set-up and data measurement

Figure 1 shows the experimental test rig, consisting of a variable speed electric motor, a rotor, a 6-blade disk and 2 ball bearing supports. The disk and the 6 normal blades were designed and fabricated in a workshop. In addition to the 6 normal blades, more blades were designed and fabricated and the following defects were introduced on the blades: twisted, bent, bulged, cracked, slotted bulged, cracked and slotted blades. Some of the blades with defects are shown in Figure 2 (a). Four accelerometers are attached to the two supports along the x- and y-axes to measure the lateral vibration of the rotor-disk-blade system. The signals of the accelerometers are connected to a six-channel B&K data acquisition system and then to a laptop with Pulse software (Figure 2 (b)). The Pulse software is capable of measuring, analysing and storing vibration signals from the accelerometers and the signals can be exported to excel file for further processing.

![Experimental test-rig for rotor-disk-blade responses at a constant running speed](image)

First, the test-rig was operated by running the electric motor and rotor-disk-blade system at a speed of 1000 rpm and the vibration signals in time domain for the set-up with 6 normal blades were obtained from the four accelerometers and saved in the laptop. One of the 6 normal blades was then replaced with one blade with a defect and the set-up was operated again at 100 rpm to obtain the vibration signal of the system with a defective blade. The defective blade was later removed and subsequently replaced with other defective blades one after the other to obtain the vibration signals for all the defective blades. All the time domain vibration...
signals of the rotor-disk-blade system were exported to excel file for analysis. The vibration signals from the x- and y-axis accelerometers attached to the support that is close to the disk-blade system are considered for further analysis.

![Image](image1.png)

**Figure 2.** (a) three defective blades (bent, twisted and broken blades); (b) PC and B&K pulse system and vibration data measurement and analysis accessories.

### 2.2. Data collection procedure
Experiments are conducted on the different blades of a wind turbine by varying the frequency (F) from 0 Hz to 2000 Hz. The x- and y-coordinates for different sets of blades ‘namely’ the good blades without any defects, twisted blades, bent blades, bulged blades, cracked blades, and slotted blades; all of which are recorded in an excel file. The dataset so prepared has 13 attributes and 2134 observations. The attributes are the frequency, x- and y-axis vibration measurements of the good blades without any defects, twisted blades, bent blades, bulged blades, cracked blades, and slotted blades.

### 2.3. Simulation model
Python is a high-level programming language. Python is a programming language that can be utilized to create desktop graphical user interface (GUI) programs [28]. Furthermore, since Python is a high-level programming language, it helps you to concentrate on the application's core features while it handles the mundane programming tasks. Python was used in this study to develop the ANN for classifying various rotating blade faults commonly encountered in rotating machines. Compared to other software’s, python is a general-purpose programming language. Python gives more control over one's code organization and better management of namespace. There are three main steps to develop the ANN in python [29], i.e. 1) using the provided data as input 2) predicting the output 3) comparison of the prediction with the actual output and 4) changing its internal state in order to correctly predict the next time.

### 3. ARTIFICIAL NEURAL NETWORK
ANN refers to a computational model mimicking the human brain and its information processing capabilities [30], [31]. An ANN is an ensemble of highly interconnected processing units known as the neurons [32]. There are a number of numerical values between the neurons that shows the strength of the connections known as the weights [23], [33]. The ANN is capable of self-organization and acquisition of information, i.e., learning [34].

Figure 3 shows a typical three feed forward neural network, comprising of the input layer, hidden layer and output layer. In this network, the middle layer nodes only receive information from the nodes of the previous layer and forward their outputs to consecutive layer nodes only. The hidden layers are not connected with the outside world directly. The computational elements or nodes used in neural network models are normally nonlinear and generally analog. As shown in Figure 3, the simplest node sums the weighted inputs and transfers the result through nonlinearity. The major part of the ANN is its learning algorithm. The BP, which is a supervised learning algorithm, is the most widely used for various applications [35], [36].
3.1. Back-propagation algorithm

A type of supervised learning is the BP algorithm [37], [38]. It takes place in two stages, ‘namely’ forward propagation of information and backward propagation of errors [14]. It learns by presenting the input and output data patterns repeatedly and propagating the error backward each time and adjusting the weights until the error minimizes to a reasonable low value. Some of the BP algorithm’s fundamental weaknesses include longer convergence period and susceptibility to failure of training. One of the modifications introduced to resolve this issue include adding a momentum term for faster training, thereby compromising additional memory space. The following equation gives the output of the neuron [37]:

$$X_i^{[k]} = f(t_i^{[k]}) = f(\sum_{j=0}^{N} w_{ij}^{[k]} x_j^{[k-1]})$$  \hspace{1cm} (1)

Where $w_{ij}$ and $x_j$ are the associated weights and the inputs of the ANN, respectively. Figure 4 show how the BP algorithm is implemented together with the processing elements. Inputs ($X_{p1}, \ldots, X_{pNo}$) are initially propagated and the response is computed via the network. Then the output error ($T_{p1}, \ldots, T_{pNM}$) is propagated backward through the network and the weights are modified according to the algorithm based on the gradient descent [30]. This method continues until the average mean square error at the output reaches an acceptable minimum value. For this work, the BP algorithm will be implemented in the ANN to classify the different blade faults.
4.1. Exploratory data analysis

The next important step after data collection is the exploratory data analysis. The data is analysed by summarizing the basic statistical description of the data for finding anomalies, missing values, duplicate values, and patterns or relationship between the attributes. Statistical description of the data reveals that all attributes are floating point numbers and there are no missing values in the dataset. The range of the attribute frequency is from 0 Hz to 2000 Hz. For all other attributes, the minimum value is a very small continuous floating point value that has a power of 1e-9. Further analysis of the data using duplicated in pandas, a data analysis software, reveal there are no duplicate rows in the dataset. To further analyse the dataset for outliers, skewness and kurtosis functions in pandas have been used and the results reveal high positive value for skew and kurtosis for all attributes except the frequency. High positive value for skew indicates all attributes except frequency are right skewed which indicates the data distribution is not symmetrical and highly skewed with longer tails to the right side. High positive values for kurtosis reveal numerous outliers in the dataset for all attributes except frequency. The results shown in Table 1 indicate significant reduction in skewness and kurtosis values which indicates the transformed dataset is now closer to normal distribution with reduced outliers. Exploratory data analysis reveals that the dataset has numerous outliers and is positively highly skewed. Therefore, the training of the ANN would not be ideal and with the help of a suitable transformation algorithm, the dataset has to be transformed to a new dataset devoid of outliers and skewness.

![](https://example.com/Table_1.png)

Table 1. Analysis of the attributes in the dataset

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Skewness Before processing</th>
<th>Kurtosis Before processing</th>
<th>Skewness After processing</th>
<th>Kurtosis After processing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frequency</td>
<td>0</td>
<td>1.2</td>
<td>-1.97483</td>
<td>5.53236</td>
</tr>
<tr>
<td>Good X</td>
<td>14.94039</td>
<td>254.9368</td>
<td>-0.90639</td>
<td>0.649114</td>
</tr>
<tr>
<td>Good Y</td>
<td>45.25426</td>
<td>2074.963</td>
<td>-1.42593</td>
<td>4.291290</td>
</tr>
<tr>
<td>Twisted X</td>
<td>14.87656</td>
<td>324.2824</td>
<td>-0.74694</td>
<td>-0.13571</td>
</tr>
<tr>
<td>Twisted Y</td>
<td>43.77001</td>
<td>1981.886</td>
<td>-1.12677</td>
<td>2.473182</td>
</tr>
<tr>
<td>Bent X</td>
<td>11.33756</td>
<td>162.3006</td>
<td>-0.84875</td>
<td>0.286339</td>
</tr>
<tr>
<td>Bent Y</td>
<td>44.99905</td>
<td>2058.742</td>
<td>-1.36612</td>
<td>4.188487</td>
</tr>
<tr>
<td>Bulged X</td>
<td>24.84025</td>
<td>759.1271</td>
<td>-0.90449</td>
<td>0.610648</td>
</tr>
<tr>
<td>Bulged Y</td>
<td>46.12201</td>
<td>2129.461</td>
<td>-1.55131</td>
<td>5.372109</td>
</tr>
<tr>
<td>Cracked X</td>
<td>20.21336</td>
<td>546.9635</td>
<td>-0.88551</td>
<td>0.480553</td>
</tr>
<tr>
<td>Cracked Y</td>
<td>45.89052</td>
<td>2115.043</td>
<td>-1.4902</td>
<td>4.656801</td>
</tr>
<tr>
<td>Slotted X</td>
<td>18.88787</td>
<td>463.5696</td>
<td>-0.82665</td>
<td>0.072927</td>
</tr>
<tr>
<td>Slotted Y</td>
<td>46.16197</td>
<td>2131.942</td>
<td>-1.47583</td>
<td>5.158623</td>
</tr>
</tbody>
</table>

4.2. Data transformation

Even many transformations are available, log transformation is applied to all attributes of this dataset to reduce the skewness to make the data as normal as possible. For the attribute frequency, the log transformation is applied to all values except the first value which is zero. For all other values, the application of the Log transformation results in a negative value and hence, after applying Log transformation, the sign of the data is changed from negative to positive. The basic statistical details of the transformed dataset are summarized in Table 2.

![](https://example.com/Table_2.png)

Table 2. Statistical description of the data after transformation

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Count</th>
<th>Mean</th>
<th>Standard deviation</th>
<th>Minimum</th>
<th>25%</th>
<th>50%</th>
<th>75%</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frequency</td>
<td>2134</td>
<td>2.86629</td>
<td>0.43307</td>
<td>-0.02803</td>
<td>2.69890</td>
<td>2.99993</td>
<td>3.17602</td>
<td>3.30996</td>
</tr>
<tr>
<td>Good X</td>
<td>2134</td>
<td>5.69455</td>
<td>0.78286</td>
<td>2.48956</td>
<td>5.15335</td>
<td>5.89308</td>
<td>6.29755</td>
<td>8.73845</td>
</tr>
<tr>
<td>Good Y</td>
<td>2134</td>
<td>5.72442</td>
<td>0.57929</td>
<td>1.59341</td>
<td>5.44511</td>
<td>5.85112</td>
<td>6.10766</td>
<td>8.52216</td>
</tr>
<tr>
<td>Twisted X</td>
<td>2134</td>
<td>5.60984</td>
<td>0.86286</td>
<td>2.40997</td>
<td>4.92536</td>
<td>5.90164</td>
<td>6.30942</td>
<td>8.67312</td>
</tr>
<tr>
<td>Twisted Y</td>
<td>2134</td>
<td>5.67138</td>
<td>0.59989</td>
<td>1.98165</td>
<td>5.33377</td>
<td>5.80116</td>
<td>6.08955</td>
<td>8.81861</td>
</tr>
<tr>
<td>Bent X</td>
<td>2134</td>
<td>5.67871</td>
<td>0.79395</td>
<td>2.75327</td>
<td>5.16924</td>
<td>5.89219</td>
<td>6.29509</td>
<td>8.68163</td>
</tr>
<tr>
<td>Bent Y</td>
<td>2134</td>
<td>5.71686</td>
<td>0.60327</td>
<td>1.47977</td>
<td>5.42954</td>
<td>5.84326</td>
<td>6.12088</td>
<td>8.95687</td>
</tr>
<tr>
<td>Bulged X</td>
<td>2134</td>
<td>5.72402</td>
<td>0.77280</td>
<td>2.25688</td>
<td>5.20475</td>
<td>5.93149</td>
<td>6.31493</td>
<td>8.66393</td>
</tr>
<tr>
<td>Bulged Y</td>
<td>2134</td>
<td>5.76086</td>
<td>0.60550</td>
<td>0.89438</td>
<td>5.47030</td>
<td>5.89358</td>
<td>6.15173</td>
<td>8.96419</td>
</tr>
<tr>
<td>Cracked X</td>
<td>2134</td>
<td>5.71243</td>
<td>0.78809</td>
<td>2.34252</td>
<td>5.20453</td>
<td>5.93768</td>
<td>6.31626</td>
<td>8.39567</td>
</tr>
<tr>
<td>Cracked Y</td>
<td>2134</td>
<td>5.74478</td>
<td>0.60007</td>
<td>1.29346</td>
<td>5.46098</td>
<td>5.87850</td>
<td>6.13801</td>
<td>8.76337</td>
</tr>
<tr>
<td>Slotted X</td>
<td>2134</td>
<td>5.67220</td>
<td>0.81912</td>
<td>2.46975</td>
<td>5.08905</td>
<td>5.92917</td>
<td>6.30308</td>
<td>8.38482</td>
</tr>
<tr>
<td>Slotted Y</td>
<td>2134</td>
<td>5.69869</td>
<td>0.60130</td>
<td>0.73089</td>
<td>5.39308</td>
<td>5.83621</td>
<td>6.10179</td>
<td>8.76579</td>
</tr>
</tbody>
</table>

Rotating blade faults classification of a rotor-disk-blade system using ... (Abdullahi Abubakar Mas’ud)
5. DESIGN OF THE ANN AND THE PREDICTED RESULTS

5.1. Design of the ANN

This section describes the architecture of the ANN designed to predict the 10 faults (5 defects in each x- and y-directions) in the blades of the rotating machines. The input layer of ANN consists of three neurons ‘namely’ the frequency, good x and good y attributes. The output layer consists of 10 neurons namely, x- and y-axis vibration measurements of twisted blades, bent blades, bulged blades, cracked blades and slotted blades. After trial and error with a number of neurons, a single hidden layer with 80 neurons with linear activation function is considered between the input and output layer as shown in Figure 5. Before choosing the linear activation, correlation between the attributes were considered and a positive correlation between the input and output attributes was found as shown in the heat map (see Figure 6).

![Figure 5. The architectures of the ANN to find the faults in the rotating machine](image)

Figure 5. The architectures of the ANN to find the faults in the rotating machine

![Figure 6. Heat map of the correlation between the attributes after data transformation](image)

Figure 6. Heat map of the correlation between the attributes after data transformation

5.2. Selection of the training and testing dataset

Scikit-learn machine learning library available for machine learning for python programming [21] is used to split the transformed dataset into training and testing dataset. Ninety percent of the dataset is used for training and ten percent of the dataset is used for testing. 1920 observations are randomly split as training data and 214 observations are split as testing data.

5.3. Training of the ANN

TensorFlow is a machine learning platform used for training neural networks and it is an open source library with plenty of fast numerical computing algorithms available to create ANN models and to train them. Keras is an API that runs on top of TensorFlow and enables fast experimentation to fine tune and train ANN models. Training of ANN in this research is carried out using Keras and TensorFlow as it helps to sequentially build and test neural network very quickly with minimal lines of code.
Adam optimizer (adaptive moment estimator) is used for upgrading the weights of the dense layers of the ANN during the training. Adam uses adaptive learning rate during training phase and changes the individual learning rates for each parameter based on squared gradients and momentum to enable faster convergence. Adam is superior to stochastic gradient descent optimizer which uses a learning rate that does not vary during the iterations and hence preferred for training the ANN.

During training the mean squared error (MSE) loss function is minimized. The maximum epoch is specified as 80. Very common issue with training neural network model is the problem of over fitting. To overcome this problem EarlyStopping feature available in keras to stop training when a monitored metric has stopped improving has been used. The patience argument available in EarlyStopping function is set to five which stops the training when the loss function has not significantly improved over 5 iterations. The 20 per cent of the training data was used for validation to test the model during training. The convergence of the model is provided in Table 3. In Table 3, MSE, MAE, VAL_MSE and VAL_MAE refer to mean squared error during training, mean absolute error during training, mean squared error during validation, and mean absolute error during validation respectively. As seen from Table 3 the model converged in 17 iterations when there is no significant improvement in mse during training. The early stopping feature used in this work ensures the training continues until the error significantly decreases and error for training and validation data is almost similar to avoid over fitting. The mse during training and validation is almost the same at the end of the 17th iterations which indicates the model will not over fit for unseen data. The plot the convergence characteristics of the model are shown in Figure 7 (a) and Figure 7 (b). From Figure 7 (a) and Figure 7 (b), it is evident that during training of ANN, the MSE and MAE reduces to the lowest possible value.

<table>
<thead>
<tr>
<th>Epoch</th>
<th>MSE</th>
<th>MAE</th>
<th>VAL_MSE</th>
<th>VAL_MAE</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>18.711683</td>
<td>2.649993</td>
<td>0.358549</td>
<td>0.439097</td>
</tr>
<tr>
<td>1</td>
<td>0.231891</td>
<td>0.369927</td>
<td>0.171996</td>
<td>0.322538</td>
</tr>
<tr>
<td>2</td>
<td>0.143537</td>
<td>0.290199</td>
<td>0.129434</td>
<td>0.275002</td>
</tr>
<tr>
<td>3</td>
<td>0.129264</td>
<td>0.274902</td>
<td>0.12706</td>
<td>0.270735</td>
</tr>
<tr>
<td>4</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>5</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>6</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>7</td>
<td>0.064647</td>
<td>0.189005</td>
<td>0.054906</td>
<td>0.17281</td>
</tr>
<tr>
<td>8</td>
<td>0.060739</td>
<td>0.182184</td>
<td>0.048822</td>
<td>0.163773</td>
</tr>
<tr>
<td>9</td>
<td>0.066706</td>
<td>0.193218</td>
<td>0.06819</td>
<td>0.208151</td>
</tr>
<tr>
<td>10</td>
<td>0.069894</td>
<td>0.182181</td>
<td>0.05758</td>
<td>0.182925</td>
</tr>
<tr>
<td>11</td>
<td>0.064329</td>
<td>0.189406</td>
<td>0.061712</td>
<td>0.195263</td>
</tr>
</tbody>
</table>

Figure 7. Plot of (a) MSE, and (b) MAE during training of ANN

5.4. ANN prediction results

In this section, the plot of the actual value and the predicted value for all 10 faults are shown. In Figures 8 (a)-(j), the trained models are evaluated on the test data and the plots for the 10 attributes are shown. The predicted values for each fault are shown in the y-axis, while the actual values are shown in x-axis. The result shows that ANN can predict the x- and y-axes lateral vibrations of most of the faults very closely. A 45° line indicates a prediction with minimal errors and the data points away from the 45° line indicates the error in prediction by the ANN. Generally, the results imply that the actual and predicted values for most of the faults are closely similar with wider variations in few fault examples.
6. **CONCLUSION**

In this paper, the ANN has been applied to classify different faults of rotating machines. First, the lateral vibrations of a rotor-disk-blade device consisting of 6 healthy blades were analyzed experimentally.

Figure 8. Plot of actual and predicted values for 10 faults, (a) twistedx, (b) twistedy, (c) bentx, (d) benty, (e) bulgedx, (f) bulgedy, (g) crackedx, (h) crackedy, (i) slottedx, and (j) slottedy.
and through finite element modal analysis. The system responses in two directions were obtained at 1000 rpm. Five faulty blades (twisted, bent, bulged, cracked, and slotted) were replaced in turn by one of the 6 safe blades. Comparisons of a twisted blade's response magnitudes with those of healthy blades indicate that the vibration amplitudes in the acceleration response range of the first two resonant frequencies decreased, while the amplitude nearer to the running speed increased.

After obtaining the data on the aforementioned rotational machine faults, the ANN was utilized to be able to classify and discriminate the x- and y-axis lateral vibrations of the rotor. The ANN was coded in python software. Initially, correlation between the attributes was considered and it was shown that there is a positive correlation between the input and output attributes. For training the ANN, 90% of the dataset is used and 10% of the dataset is used for testing. The result shows that ANN can very accurately predict the two-dimensional (x- and y-axes) lateral vibrations of most of the faults. The results imply that the ANN can effectively recognize and discriminate different lateral vibration in a rotor depending on training and testing with different data samples of the faults types. Further research concentrates on the application of the ANN for practical fault detection and classification.

ACKNOWLEDGEMENTS

Acknowledgements are due to Jubail Industrial College and Jubail University College, Royal Commission for Jubail.

REFERENCES


**BIOGRAPHIES OF AUTHORS**

**Abdullahi A. Masud (PhD, CEng)** received his BEng. Degree in Electrical Engineering from Ahmadu Bello University (ABU) in 1999. Subsequently, he pursued his Master’s degree in the same discipline at ABU and graduated in 2003. He also obtained a PhD in Electrical Engineering in 2013 from the Glasgow Caledonian University, Scotland, UK, under the supervision of Prof. Brian Stewart and Prof. Scott McMeekin. In 2002 he became an assistant lecturer in the Faculty of Engineering at the Ahmadu Bello University, Zaria, Nigeria. In 2013 he joined Juhail Industrial College, Saudi Arabia, and is currently an Associate Professor in the Department of Electrical and Electronic Engineering. He has several publications in high impact factor journals and conference proceedings in high voltage partial discharge and renewable energy. He is a Chartered Engineer, a Member of the IET, and a registered Engineer (COREN) Nigeria.
Ahmad Jamal obtained his B.S. in Mechanical Engineering degree in 1998 University of Engineering and Technology, Lahore, Pakistan. His M.S. degree in Mechanical Engineering was from King Fahd University of Petroleum and Minerals, Dhahran, Saudi Arabia in 2002 with specialization in Thermo-Fluid. He secured his Ph.D. degree from McGill University, Montreal, Canada in 2014 in the area of Fluid-Structure Interaction (FSI). He is the author of more than 20 refereed journals and conferences in the fields of Thermo-Fluid, FSI, Reliability, and Machine Learning. He undertook many research projects related to Thermo-Fluid and FSI. He has also supervised many undergraduate design projects related to Thermo-Fluid and Applied Mechanics. His new research avenue is in the field of Control Theory.

Dr. Surajudeen Adewusi, P. Eng., MASME is an Assistant Professor and Chairman of Mechanical Engineering Department at Jubail University College, Saudi Arabia. He obtained Ph.D. in Mechanical Engineering from Concordia University, Canada in 2009 and Masters of Science in Mechanical Engineering from King Fahd University of Petroleum & Minerals, Saudi Arabia in 2000. He was a recipient of the Natural Sciences and Engineering Research Council of Canada (NSERC) and Quebec Province Research Fund on Nature and Technology (FQRNT) Postdoctoral Research Fellowships in 2011. Surajudeen has presented research/technical papers at international conferences and co-authored 30 research papers in international journals and conference proceedings. His area of expertise include Experimental and Finite Element Modeling, Analysis and Control of Systems (Dynamic Characterization of Systems), Rotating Machinery Vibration Condition Monitoring and Faults Diagnosis (Rotordynamics), Human Biodynamic Responses to Occupational Vibrations (Biomechanics) and Renewable Energy. He has 25 years of combined academic, industrial and research work experiences. Dr. Adewusi is a member of the Professional Engineers of Ontario (PEO), Canada and American Society of Mechanical Engineers (ASME).

Dr. Arunachalam Sundaram received the B.E. degree in Electrical and Electronics Engineering from Annamalai University, Chidambaram, India, in 2001, and the M.E. degree in Power Systems Engineering and the Ph.D. degree in Electrical Engineering from Anna University, Chennai, India, in 2004 and 2014, respectively. He received the Gold Medal from Anna University for securing first rank in M.E power systems engineering. He has also completed Post Graduation Program in Artificial Intelligence and Machine Learning from University of Texas at Austin offered by Great Learning. He is a senior member in IEEE. He is currently working as Associate Professor in Jubail Industrial College, Kingdom of Saudi Arabia, Al-Jubail. He has over 17 years of academic experience where he was involved in developing short course and long courses for various industries. His research interest includes artificial intelligence and machine learning, optimization applied to power system engineering, and smart grids.