



# Bearing Damage Analysis with Artificial Intelligence Algorithms

## A New Feature Extraction Approach for Current-Based Signals

André da Silva Barcelos<sup>1</sup> · Fábio Muniz Mazzoni<sup>2</sup> · Antonio J. Marques Cardoso<sup>1</sup>

Received: 23 February 2021 / Revised: 21 May 2021 / Accepted: 7 June 2021 / Published online: 13 September 2021  
© Brazilian Society for Automatics--SBA 2021

### Abstract

Three-phase induction motors are widely used in industrial facilities, where the maintenance of these machines is preponderant for industrial processes. Recent research reports that vibration-based data acquisition is the most common approach to perform bearing condition monitoring because it can extract more relevant information. However, the acquisition of vibration-based signals is expensive, requiring accelerometers and other external devices to transmit and process the signal information. Otherwise, current-based signals are directly measured by the supply system or inverters, enabling the current-based data acquisition in most industrial cases. In this context, this work introduces a new current-based method to identify bearing damages, applying artificial intelligence algorithms. Experimental and on-site tests present promising results, validating this approach for bearing damage diagnosis.

**Keywords** Three-phase induction motor · Bearing · Diagnosis · Artificial intelligence

## 1 Introduction

The conversion of electrical energy into mechanics is present in most industrial processes. Three-phase induction motors (TIMs) are the principal equipment responsible for driving pumps, compressors, valves, conveyors, propellers, elevators, etc. Recent research estimates that 70% of the European Union's industrial energy consumption is directly related to three-phase electric motor applications due to low-cost man-

ufacturing and versatility for high-performance applications (Cardoso, 2018; Merizalde et al., 2017).

These motors present electric faults and mechanical damages owing to problems from construction and operating conditions. These problems may result in reduced performance or industrial process interruptions. Studies published by the Institute of Electrical and Electronics Engineers and by the Electric Power Research Institute report that 40% of TIM failures and damages are related to bearings (Cerrada et al., 2018).

Bearing damages can be separated into two categories. The first category is the punctual damages, which appear on a delimited bearing surface. The characteristics of these damages are holes, scratches, particles, corrosion, electrical discharge, material removal, impact points, and others. These damages produce impulsive vibration frequencies (Barcelos & Cardoso, 2021).

The second category is the distributed damages, exemplified by flushing, encrustations, loss of material, corrosion, wear, generalized roughness, or other forms that propagate throughout the entire length of the bearings' raceways. These types of damages produce continuous mechanical vibrations with low magnitude harmonics. Recent research allows predicting distributed damages with models (Cardoso, 2018; Irfan, 2019; Randall, 2011).

---

An early version of paper was presented at XXIII Congresso Brasileiro de Automática (CBA 2020).

---

André da Silva Barcelos passed away on May 31, 2021.

---

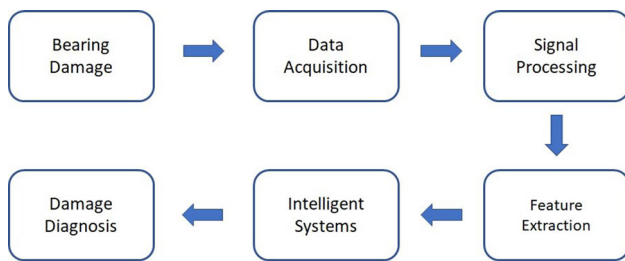
✉ Antonio J. Marques Cardoso  
ajmcardoso@ieee.org

André da Silva Barcelos  
andre.s.barcelos@ubi.com

Fábio Muniz Mazzoni  
fabiomuma@hotmail.com

<sup>1</sup> CISE-Electromechatronic Systems Research Centre, University of Beira Interior, Covilhã, Portugal

<sup>2</sup> Universidade Federal Fluminense, Instituto de Ciência e Tecnologia – ICT, Covilhã, Portugal



**Fig. 1** Bearings diagnosis overview (Color figure online)

Various approaches acquire signals from electrical, mechanical, thermal, or other sources to monitor the bearings condition. Then, signal processing techniques interpret the bearings behavior in the time, frequency, and time-frequency domain. The first three blocks of Fig. 1 introduce the data acquisition and signal processing overview.

Consequently, after the signal processing task, the feature extraction methods use statistical tools and signal measurements to extract more relevant information, constructing a database, as described in Fig. 1. Then, intelligent systems algorithms may classify this database, performing bearing damage diagnosis (Bazan et al., 2017).

Several industrial processes acquire vibration-based signals to extract features and monitoring the bearings condition. Also, the current-based bearings analysis presents satisfactory results in recent years (Bessous et al., 2018; Leite et al., 2014). Some advantages of current-based techniques include real-time, remote, and non-invasive monitoring without the need for new sensors and devices (Cardoso, 2018; Cerrada et al., 2018).

The time-frequency analysis with the wavelet transform (WT) becomes a recurrent signal processing technique because of the advantage of multiresolution decomposition (Bessous et al., 2018; Yu et al., 2019). The kurtogram (Leite et al., 2014), enhanced kurtogram (Chen et al., 2016), and autogram (Moshrefzadeh & Fasana, 2018) are time-frequency representations based on spectral kurtosis (Wang et al., 2019; Zhang et al., 2019). Also, the empirical mode decomposition (EMD) calculates intrinsic mode functions (IMFs) to represent the current-based signal behavior with Hilbert transform in the time-frequency domain (Dragomiretskiy & Zosso, 2013). Both time-frequency methods are signal processing techniques that allow the feature extraction from current-based signals.

Furthermore, data-driven methods as stacked autoencoders (Jiang et al., 2018), variational autoencoders (San Martin et al., 2019), deep convolutional neural networks (Chen et al., 2020; Zhang et al., 2018), extreme learning (Zhao et al., 2020), generative models (Cao et al., 2018), among others, can extract automatic features from raw signals (Xia et al., 2018).

These signal processing techniques and data-driven algorithms have satisfactory performance in vibration-based

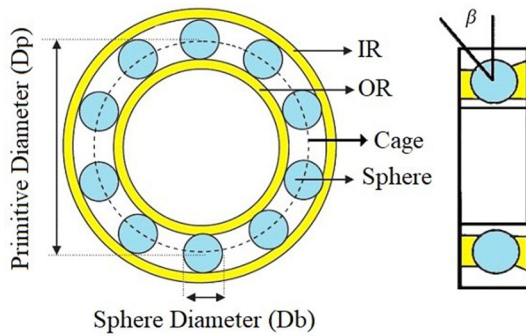
approaches. However, although current-based signals allow monitoring the bearing behavior, it is difficult to identify the category, location (inner ring  $\times$  outer ring), and severity of distributed and earlier damages (Hoang & Kang, 2019; Leite et al., 2014).

Current-based signals have a low signal-to-noise ratio that buries the information and increases the interference in most cases. Thus, it remains challenging to exclude harmonic interference when the current-based signal has accidental impulses, non-stationary behavior, quasi-overlapping impulses, and non-Gaussian noise (Barcelos & Cardoso, 2021; Wang et al., 2019). Another relevant issue is that in most industrial facilities, the induction motors are prevented from performing under damaged conditions because of industrial safety reasons. In this condition, it is impracticable to acquire a labeled database from damaged bearings (Barcelos et al., 2021).

Based on these premises, this work introduces a novel feature extraction method for current-based signals. This method calculates the covering dimension (CD) from wavelet decomposition and IMFs to construct a database. The CD is an index based on the fractal theory that measures the asymptotic or periodic local behavior of signals. Indeed, bearing damages from current-based signals produce low magnitude harmonics that increase the CD. Experimental and on-site tests are performed with a support vector machine (SVM) and an artificial neural network (ANN) to identify bearing damages from a labeled database. Subsequently, this work performs on-site tests with the support vector data description (SVDD) algorithm to identify bearing damages from an unlabeled database.

In summary, this paper introduces a new feature extraction method that employs the CD from orbits to detect bearing damage in current-based signals. The wavelet transforms and the EMD are the signal processing techniques used in this work, while the ANN and SVM are the supervised learning classifiers. In addition, the SVDD is used in the novelty detection framework to conduct positive unlabeled learning when the labeled database is unavailable. The KLD and FDR measure the centers and boundary behavior of distributions, improving the SVDD accuracy. After the SVDD detects the novelty, which in this case is the damaged bearing signal, the ANN and SVM may perform classification, enabling a supervised learning classifier to solve a complex PUL problem.

The following sequence introduces the bearing's models in Sect. 2. The wavelet transform and EMD are presented in Sect. 3. The feature extraction and the covering dimension are reported in Sect. 4. Section 5 presents the ANN, SVM, and SVDD formulation. Section 6 summarizes the methodology, introduces the databases, test procedure, and the results from experimental and on-site tests. Section 7 contains the main conclusions.



**Fig. 2** Geometric details from bearings with 10 spheres (Color figure online)

## 2 Data Acquisition

Bearing damages cause dysfunctions in the induction motor magnetic field resulting in harmonics in the stator currents. These harmonics depend on the supply frequency, motor speed, and geometric characteristics of bearings, such as inner raceway (IR) radius, sphere diameter ( $D_b$ ), cage assembly, number of spheres, and outer raceway (OR) radius. Figure 2 presents the principal geometric details of bearings with ten spheres (Rao et al., 2019).

The characteristic frequency ( $f_c$ ) for the outer ring ( $f_o$ ), inner ring ( $f_i$ ), and spheres ( $f_b$ ) is defined in (1), (2), and (3), as follows:

$$f_c = f_o = \frac{N_b}{2} f_r \left( 1 - \frac{D_b}{D_p} \cos \beta \right) \tag{1}$$

$$f_c = f_i = \frac{N_b}{2} f_r \left( 1 + \frac{D_b}{D_p} \cos \beta \right) \tag{2}$$

$$f_c = f_b = \frac{D_p}{D_b} f_r \left( 1 - \frac{D_b^2}{D_p^2} \cos^2 \beta \right) \tag{3}$$

where  $f_r$  is the rotor frequency,  $D_p$  is the primitive diameter,  $D_b$  is the sphere diameter,  $N_b$  is the number of spheres, and  $\beta$  is the contact angle (Bessous et al., 2018). The punctual damage frequencies ( $f_p$ ) and harmonics are calculated as follows:

$$f_p = |f_r \pm k f_c| \quad \text{where } k \in \mathbb{N} \tag{4}$$

The inner ring ( $f_{di}$ ) and outer ring ( $f_{do}$ ) frequencies and harmonics from distributed damages are calculated as follows:

$$f_{di} = n f_r \left[ \left( N_b - \frac{D_{ir}}{D_b} \right) \left( \frac{D_p - D_b}{2 D_p} \right) + \frac{D_{ir}}{D_b} - 1 \right] \tag{5}$$

$$f_{do} = n f_r \left[ \left( N_b - \frac{D_{or}}{D_b} \right) \left( \frac{D_p - D_b}{2 D_p} \right) \right] \tag{6}$$

where  $n \in \mathbb{Z}$ , the inner diameter is  $D_{ir}$ , and the outer diameter is  $D_{or}$ . With these models, it is possible to search for

the predicted-model harmonics in the time-frequency domain (Irfan, 2019).

## 3 Signal Processing

### 3.1 Wavelet Transform

The Fourier transform (FT) is a recurrent approach to acquire stationary information from signals in the frequency domain. The short-time Fourier transform (STFT) uses a constant resolution window to acquire non-stationary information in the time-frequency domain (Aimer et al., 2019).

The STFT becomes inadequate in several applications with non-stationary signals that require a variable resolution window. The multiresolution analysis overcomes the constant window drawbacks introducing windows with different resolutions (Bessous et al., 2019; Kamiel & Howard, 2019).

In this context, the wavelet transform is a multiresolution signal processing technique that uses different resolution windows to represent information in the time-frequency domain. The inner product between wavelets and functions is the wavelet transform, as follows:

$$\langle f(t), \psi_{a,b}(t) \rangle = \frac{1}{\sqrt{a}} \int f(t) \psi \left( \frac{t-b}{a} \right) dt \tag{7}$$

where  $\psi_{a,b}(t)$  is the wavelet function,  $a$  and  $b$  are scaling and shifting parameters. Also, the wavelet function must belong to the Lebesgue space  $L^2$ , with regularity, and finite energy, as follows:

$$\int_{-\infty}^{+\infty} \psi(t) dt = 0 \tag{8}$$

$$\int_{-\infty}^{+\infty} |\psi(t)|^2 dt < \infty \tag{9}$$

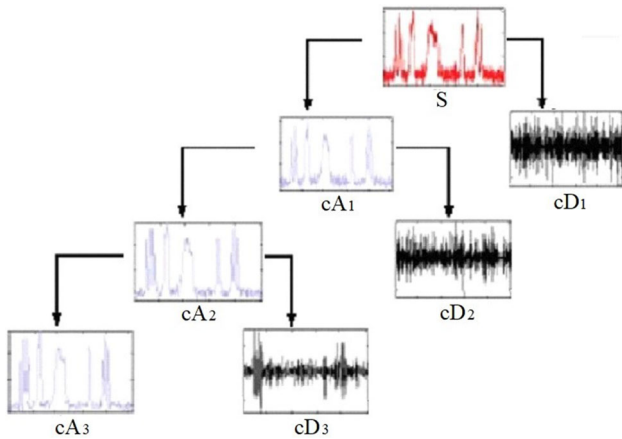
An admissible wavelet  $\psi(t)$  implies that  $\hat{\psi}(0) = 0$ , where hat is the Fourier transform operator.

### 3.2 Discrete Wavelet Transform

The discrete wavelet transforms (DWTs) define the parameters  $a$  and  $b$  as integers, where  $b$  depends on  $a$ . This redefinition allows obtaining the discrete wavelet  $\psi_{m,n}(t)$  with integer parameters for scaling ( $m$ ) and shifting ( $n$ ) according to (10). The inner product of (11) performs the discrete wavelet transform as follows:

$$\psi_{m,n}(t) = a_0^{-m/2} \psi(a_0^{-m} t - n b_0) \tag{10}$$

$$\langle f(t), \psi_{m,n}(t) \rangle = \frac{1}{\sqrt{a^m}} \sum_{k=-\infty}^{+\infty} f(k) \psi \left( \frac{t - n b_0 a_0^m}{a_0^m} \right) \tag{11}$$



**Fig. 3** The multiresolution decomposition of successive wavelets transforms into approximation and details coefficients (Color figure online)

Therefore, the DWT decomposes the function  $f(t)$  throughout low-pass (Lp) and high-pass (Hp) filters. This decomposition generates two new signals with coefficients  $l[k]$  and  $h[k]$  from Lp and Hp filters (BayroCorrochano, 2019).

The construction of Lp and Hp filters demands that a scaling function  $\phi(t)$  depends on  $l[k]$  coefficients, and the wavelet  $\psi_{m,n}(t)$  depends on  $\phi(t)$  and  $h[k]$  as follows:

$$\phi(t) = \sqrt{2} \sum_{k \in \mathbb{Z}} l[k] \phi(2t - k) \tag{12}$$

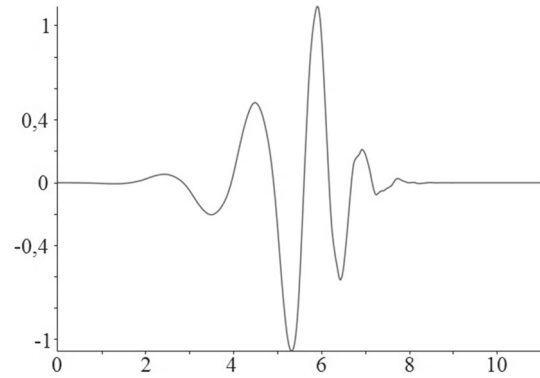
$$\psi(t) = \sqrt{2} \sum_{k \in \mathbb{Z}} h[k] \phi(2t - k) \tag{13}$$

The  $l[k]$  coefficients from Lp filters are the approximation coefficients (cA). Also, the  $g[k]$  coefficients from Hp filters are the details coefficients (cD) (Gupta et al., 2019). The multiresolution decomposition of Fig. 3 consists of successive discrete wavelet transforms of the signal  $S$  to extract several levels of details coefficients.

The parameter  $m$  of DWT rescales every successive decomposition, narrowing the frequency resolution. In the bearing diagnosis context, the last level of successive decomposition with non-redundant information is defined by (14) as follows:

$$L(cD) = \text{int} \left[ \frac{\log\left(\frac{f_a}{f_r}\right)}{\log(2)} \right] + 1 \tag{14}$$

where  $L(cD)$  is the last level of cD,  $f_a$  is the sampling frequency, and  $f_r$  is the rotor frequency (Bessous et al., 2019; Ghods & Lee, 2016).



**Fig. 4** Daubechies wavelet with order  $N = 12$

### 3.3 Daubechies wavelets

The Daubechies wavelets are functions  $\psi_{a,b}(t)$  in orthonormal bases with compact support, regularity, and the maximum number of vanishing moments (Daubechies, 1988). Fig. 4 presents the Daubechies wavelet with order  $N = 12$ , compact support  $N - 1 = 11$ , and  $N/2$  vanishing moments.

This wavelet has maximum and minimum, finite energy, and zero average, according to (8) and (9). The existence of  $k$  vanishing moments from wavelets is defined by (15) as follows:

$$\int x^k \psi(x) dx = 0 \tag{15}$$

The number of vanishing moments is related to the smoothness for time-frequency representation and the capacity to approximate polynomials (Narendiranath et al., 2017).

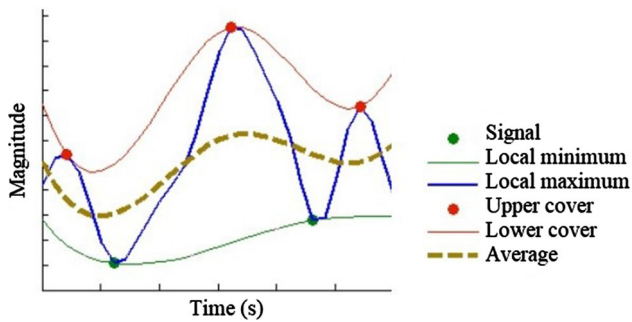
### 3.4 Hilbert-Huang Transform

The Hilbert-Huang transform (HHT) decomposes a function into a complex plane, keeping the instantaneous frequency and localizing phenomena in time (Bessous et al., 2019). The HHT is calculated with (16) to represent the non-stationary signals in time-frequency domain.

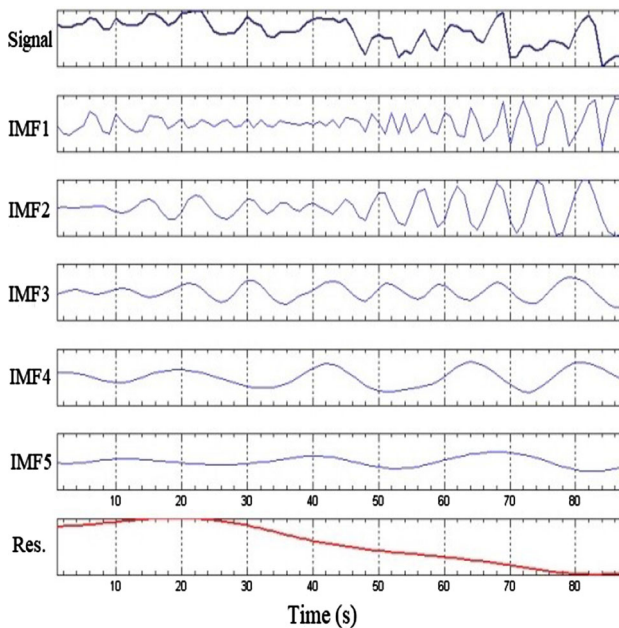
$$z(t) = s(t) + \frac{i}{\pi} P \int_{-\infty}^{\infty} \frac{s(\tau)}{t - \tau} d\tau = a(t) e^{i\theta(t)} \tag{16}$$

where  $i$  is the imaginary number,  $P$  is the Cauchy principal value, and  $s(t)$  is the signal. The function  $a(t)$  contains the variable magnitude, and  $\theta(t)$  represents the angular variations in the complex plane.

The empirical mode decomposition (EMD) is a recurrent technique that breaking down a signal to extract intrinsic mode functions (IMFs), allowing the HHT for sampled signals (Bessous et al., 2019). A function is defined as IMF when the number of extrema and zero-crossings differ at most by one.



**Fig. 5** The enveloping procedure to obtain IMFs with sifting algorithm (Color figure online)



**Fig. 6** Subsequent IMFs in blue and sifting residual in red (Color figure online)

The proceedings to construct IMFs are named sifting. The maxima and minima cubic spline interpolation connects the extrema points of the sampled signal. Figure 5 shows the enveloping and the average.

The first proto-IMF is the difference between the signal and the envelope average. The sifting is repeated with the first proto-IMF that becomes data and generates the second proto-IMF. Figure 6 presents the IMF's and the summation of the residuals from sifting process.

A recurrent stop criterion for EMD compares the energy of each IMF. Different values for subsequent IMFs mean orthogonality loss. Furthermore, each IMF contains a part of the information from the signal behavior allowing feature extraction without HHT calculation.

## 4 Feature Extraction

A time series  $x(t)$  can be transformed into circular trajectories to form an orbit representation in a multidimensional space (Alligood et al., 1998). With this premise,  $x(t)$  is sampled into a discrete series  $x(n)$  to build a  $D$ -dimensional vector  $z(n)$  with time delay  $\tau$ , as follows:

$$z(n, \tau) = [x(n), x(n - \tau), \dots, x(n - (D - 1)\tau)] \quad (17)$$

The parameter  $\tau$  can be calculated by mutual information minimization between each vector pairs (Alligood et al., 1998). This procedure unfolds  $z(n)$  projections in orbital representation, allowing new insights from geometric shapes, distributions, periodicity, and trajectories.

The Minkowski–Bouligand dimension is an intrinsic measure, which consists of covering orbits and surfaces, with an overlapping set of open disks with area  $A(\epsilon)$  and radius that lies on surface. Specifically, the covering dimension ( $CD_M$ ) is defined as:

$$D_M = \lim_{\rightarrow 0} \left( \frac{\log[A(\epsilon)]}{\log(\epsilon)} \right) \quad (18)$$

Replacing the disks by boxes with cover area  $\epsilon^2 N(\epsilon)$  to perform well-behavior computation, the covering dimension is redefined as follows:

$$D_M \approx \lim_{f \rightarrow 0} \left( \frac{\log[N(f)]}{\log(1/f)} \right) \leq \lim_{f \rightarrow 0} \left( \frac{\log[A(f)]}{\log(f)} \right) \quad (19)$$

where  $N(\epsilon)$  is the number of boxes with maximum size  $\epsilon$  that cover the surface. The  $k$ -grid zoom method  $P(k\epsilon)$  rescales the size  $\epsilon$  to estimate  $N(k\epsilon)$  at each  $k$ -scale as follows:

$$P(k, f) = \sum_{i=1}^{n/k} \max\{x_{k(i-1)+1}, x_{k(i-1)+2}, \dots, x_{k(i-1)+k+1}\} - \min\{x_{k(i-1)+1}, x_{k(i-1)+2}, \dots, x_{k(i-1)+k+1}\} \quad (20)$$

$$N(k) = \frac{P(k)}{k} + 1 \quad (20)$$

where  $k \in \mathbb{Z}^+$  is the grid scale in the discrete space and  $n$  is the number of samples. The geometric interpretation of  $\log(k, f) - \log(N(k, f))$  finds the section with better linearity at end points  $k_1$  and  $k_2$  as follows:

$$\log(N(k, f)) = a \log(k, f) + b, \quad k_1 \leq k \leq k_2 \quad (21)$$



The least square method fit the data to obtain the slope  $\hat{a} \approx a = CD_M$  as follows:

$$\hat{a} = - \frac{g(k) \sum_{k=1}^M \{f(k, N(k)) - \sum_{k=1}^M f(k, N(k))\}}{g(k) \sum_{k=1}^M \log^2(k) - \{\sum_{k=1}^M \log(k)\}^2} \tag{22}$$

where the  $g(k)$  and  $f(k, N(k))$  functions are described as follows:

$$g(k) = (k_2 - k_1 + 1) \tag{23}$$

$$f(k, N(k)) = \log(k) \log(N(k)) \tag{24}$$

During the k-scale reducing proceeding, nonlinear signals with asymptotic or periodic behavior change the  $CD_M$  iteratively. Therefore, the  $CD_M$  measures local persistent behaviors as brown noises, pink noises ( $|f|^{-n}$ ), harmonics, or others asymptotically persistent structures over rescaled shorter periods (Maragos & Sun, 1993; Shen et al., 2020).

Thus, this work calculates the average, harmonic average, kurtosis, skewness, energy, and entropy from six wavelet details coefficients (cDs) and six IMFs. Also, it transforms the cDs and IMFs into orbits to calculate the  $CD_M$  and construct two databases with 42 features each.

## 5 Intelligent Artificial Algorithms

### 5.1 Artificial Neural Network

In this work, after the feature extraction step, the artificial neural network (ANN) performs supervised learning from the databases. Indeed, the features input the ANN, which is responsible for output an accurate classification, separating the healthy and damaged bearing signals, considering different speed and loading conditions.

Moreover, the ANN output is a measure for the separation possibility, reflecting the accuracy for performing bearing damage diagnosis in similar conditions. This algorithm uses neurons in multiple layers architectures to perform classification. This algorithm uses neurons in multiple layers architectures to perform classification. A cost function  $J(\mathbf{w})$  relates the predicted output  $g(\mathbf{w}^T \mathbf{x}_i + \mathbf{b}_i)$  of ANN with a labeled output  $\mathbf{y}_j$  in a set with  $n$  inputs as follows:

$$\min J(\mathbf{w}) = \frac{1}{n} \sum_{i=1}^n f(g(\mathbf{w}_i^T \mathbf{x}_i + \mathbf{b}_i), \mathbf{y}_j) \tag{25}$$

where  $\mathbf{w}_i$  is a weight vector,  $\mathbf{x}_i$  is the random variable from the  $i$ th feature, and  $\mathbf{b}_i$  is the bias. The function  $f(\cdot, \cdot)$  search

for convex distances between predicted and labeled output, while the activation function  $g(\cdot)$  inserts nonlinear behavior on the weighted features (Witten et al., 2016).

The backpropagation algorithm adjusts the weights with a learning rate based on misclassification. The regularization can provide sparsity and penalize the overfitting. Other internal procedures can improve the learning rate, avoid underfitting and overfitting, and improve accuracy (Haykin et al., 2009; Witten et al., 2016).

### 5.2 Support vector machine

The support vector machine (SVM) is a machine learning algorithm that builds separation hyperplanes based on an optimum weight vector  $w_o$  (Haykin et al., 2009). In this paper, the SVM is used to classify the bearing damages, contrasting accuracy with the ANN algorithm. It is possible to establish an optimization problem with Lagrange multipliers ( $\lambda_i$ ) limited by a constant  $V$ , as follows:

$$\begin{aligned} \max \sum_{i=1}^n \lambda_i - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \lambda_i \lambda_j \mathbf{d}_i \mathbf{d}_j \mathbf{x}_i^T \mathbf{x}_j \\ \text{s.t. } \sum_{i=1}^N \lambda_i \mathbf{d}_i = 0 \text{ and } 0 \leq \lambda_i \leq V \end{aligned} \tag{26}$$

where  $\mathbf{x}_i$  is a vector with label  $\mathbf{d}_i$ . The optimal weight vectors  $\mathbf{w}_o$  and the bias  $\mathbf{b}_o$  are according to Eqs. 27 and 28:

$$\mathbf{w}_o = \sum_{i=1}^N \lambda_i \mathbf{d}_i \mathbf{x}_i \tag{27}$$

$$\mathbf{b}_o = 1 - \mathbf{w}_o^T \mathbf{x}_i \text{ iff } \mathbf{d}_i = 1 \tag{28}$$

The kernel  $\mathbf{x}_i^T \mathbf{x}_j$  from (26) can be replaced by symmetrical functions  $k(x, x^0)$ . Also,  $k(x, x^0)$  must be continuous, have eigenfunctions  $\varphi(x)$  and  $\varphi(x^0)$  with positive eigenvalues, and satisfy the Mercer’s condition (Witten et al., 2016).

### 5.3 Support Vector Data Description

Support vector data description (SVDD) is a one-class algorithm from positive unlabeled learning (PUL) models. The SVDD search a minimal volume hypersphere containing most of the positive data in a trade-off between volume and outlier rejection (Benkedjough et al., 2012; Noumir et al., 2012). The hypersphere can be described by a center  $c$  and radius  $r$  as follows:

$$\begin{aligned} \min J(r) = r^2 + \frac{1}{vn} \sum_{i=1}^n h_i \\ \text{s.t. } \|f(x_i) - c\| \leq r^2 + h_i \text{ and } h_i \geq 0. \end{aligned} \tag{29}$$

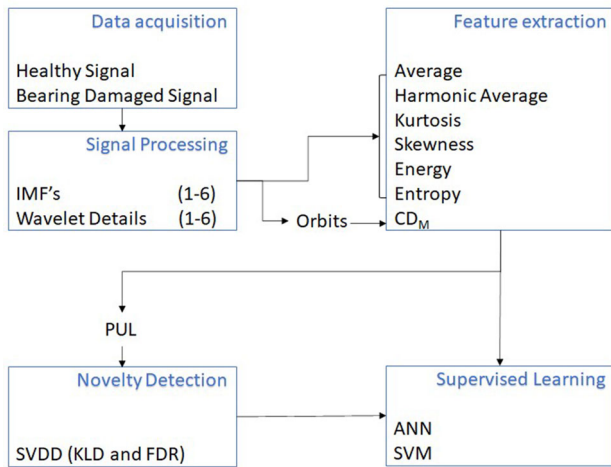


Fig. 7 The methodology for this work (Color figure online)

where  $h_i$  are slack variables,  $v$  is the trade-off between hypersphere volume and the number of data points rejected in a set with  $n$  elements. With Karush–Kuhn–Tucker conditions, Lagrange multipliers  $\lambda_i$ , and kernel trick  $K(\cdot, \cdot)$ , the dual optimization problem is obtained as follows:

$$\begin{aligned} \max L(\mathbf{x}) &= \sum_{i=1}^n \lambda_i K(\mathbf{x}_i, \mathbf{x}_i) - \sum_{i=1}^n \sum_{j=1}^n \lambda_i \lambda_j K(\mathbf{x}_i, \mathbf{x}_j) \\ \text{s.t. } 0 \leq \lambda_i &\leq \frac{1}{vn}, \sum_{i=1}^n \lambda_i = 1, c = \sum_{i=1}^n \lambda_i f(x_i) \end{aligned} \tag{30}$$

Thus, a new sample  $x_i$  is typified as an outlier when  $\|f(x_i) - c\| > r$ . In this work, two successive SVDDs are calculated with a Gaussian kernel to detect changes in healthy bearings signals. The first SVDD uses PU data (healthy) from  $(-\tau, t_0)$  interval, while the second SVDD uses the current data from  $(t_0, \tau)$  time interval.

The comparison between these two successive SVDDs with Kullback–Liebler divergence (KLD) and Fisher dissimilarity ratio (FDR) allows monitoring the global behavior and the relative center movement of data distribution (Desobry et al., 2005; Noumir et al., 2012).

## 6 Tests and Results

### 6.1 Methodology

Since the previous sections explain each algorithm separately, Fig. 7 summarizes the methodology of this work. The data acquisition consists of obtaining healthy and bearing damaged signals, while the signal processing techniques output the IMFs and the wavelets CD, as described in the first two blocks.

Table 1 Type of bearing damages

Serie	Local	Damage	Severity	Type
KA01	OR	EDM	1	Earlier
KA07	OR	Drilling	1	Punctual
KA08	OR	Drilling	2	Punctual
KA05	OR	EE	1	Earlier
KA04	OR	Pitting	1	Earlier
KA16	IR and OR	Pitting	2	Punctual
KA22	OR	Pitting	1	Punctual
KI16	IR	Pitting	3	Punctual
KI17	IR	Pitting	1	Punctual
KI18	IR	Pitting	2	Punctual
KA30	OR	ID	1	Distrib
KB27	IR and OR	ID	1	Distrib
KB24	IR and OR	ID	3	Distrib
KI04	IR and OR	ID	1	Earlier
UA16	IR and OR	ED	1	Earlier
K002	NA	NA	NA	Healthy
U002	NA	NA	NA	Healthy

The feature extraction block consists of measure the IMFs and CD information, while the orbits of these functions and coefficients are used to calculate the  $CD_M$ . In PUL context, the SVDD performs novelty detection with KLD and FDR measures, while the supervised learning is performed by ANN and SVM algorithms, as described in the last block

### 6.2 Paderborn University Dataset

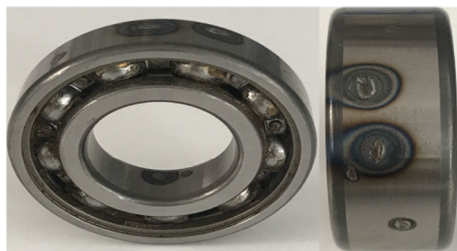
The tests performed in this work use the time series developed by the Chair of Design and Drive Technology from the University of Paderborn in Germany, which contains the current-based signals from an induction motor with healthy and damaged bearings (Lessmeier et al., 2016). Table 1 presents the damaged bearings time series, with a description of the damage location (inner raceway–IR and outer raceway–OR).

The third column presents the damages, where an electric drilling machine (EDM) provides artificial damages, a drilling machine produces the holes, and an electric engraver (EE) makes the scratches. The other three types are pitting, indentations by plastic deformations (ID), and electrical discharge (ED). The severity is an index that considers the extension and magnitude of each damage, and finally, the fifth column is the characteristic of the damages.

These time series are available at speeds of 900 rpm and 1500 rpm (N09 and N15), with loading conditions 0.1 Nm and 0.7 Nm (M01 and M07). KA01, KA04, KI04, and KA05 are time series with earlier bearing damages, which produces harmonics with low SNR.



**Fig. 8** Test rig with the current probe, acquisition board, loading system, and induction motor (Color figure online)



**Fig. 9** Bearing with earlier punctual damage caused by electrical discharge. (Color figure online)

### 6.3 CISE Dataset

The ED is a recurrent punctual damage on industrial motors. This work uses the test rig available at CISE—Electromechatronic Systems Research Centre at the University of Beira Interior in Portugal to acquire this type of data.

The test rig (Fig. 8) consists of a three-phase squirrel cage induction motor with 2.2 hp, four poles, and a supply frequency of 50 Hz. The electric motor couples a load system that provides a stable load of 0.1 and 0.7 Nm with speed and torque control. Current probes send the stator currents to an acquisition board with a sampling frequency of 44 kHz.

The healthy bearing data acquisition from this test rig produces a healthy time series named U002 in Table 1. The damaged bearing of Fig. 9, which is named UA16, has earlier punctual damages in both rings caused by electrical discharges.

The damage diameter of the inner ring is 1.5 mm. Spheres and cage remain healthy. The outer ring has two damages with 2.0 mm and another damage with 1.5 mm of diameter.

### 6.4 Test Procedure

The discrete wavelets transform and the EMD is applied to each signal of Table 1 to generate six levels of details coefficients (cD) and six successive IMFs. These signal processing techniques create two distinct data set with seven features (42 variables each) calculated on intervals of 0.05 seconds.

**Table 2** Intelligent algorithms accuracy with wavelets for signal processing

Serie	ANN	SVM $p_1$	SVM $p_2$	SVMg
Earlier	0.941	0.952	0.929	0.911
Punctual	0.991	0.986	0.944	0.920
Dist	0.966	0.974	0.968	0.931

**Table 3** Intelligent algorithms accuracy with EMD signal processing

Serie	ANN	SVM $p_1$	SVM $p_2$	SVMg
Earlier	0.930	0.963	0.913	0.904
Punctual	0.988	0.990	0.961	0.936
Dist	0.958	0.943	0.963	0.932

These data sets are separated by 70% for training, 15% for validation, and 15% for tests.

The ANN algorithm contains one hidden layer that changes within a range from eight to sixteen neurons to allow accuracy improvement. The training is limited to 1000 epochs with an adaptive learning rate from 0.32 to 0.08 and momentum from 0.2 to 0.1. The backpropagation algorithm has stopping criteria at 300 seconds, and the Lasso regularization controls the overfitting.

The soft margin SVM algorithm is performed with Gaussian kernel (SVMg) and polynomial  $[1, xy, x,y]^p$  (SVM $p$ ) kernel, where  $p_1 = 1$  and  $p_2 = 2$ . The parameter  $V$  starts at  $n = 1$  with  $n \in \mathbb{N}$  and changes  $2^n$  per epoch with stop criteria of 0.01 %. The SVDD algorithm starts with Gaussian kernel at  $\tau = 0.25$  seconds and  $t_0 = 0$ . The restriction is the Euclidean distance  $\|f(x_i - c)\|$ , while  $h_i$  regularizes the cost function  $J(r)$ .

The approach to performing a supervised test in SVDD is to insert a batch of healthy signals into the algorithm to construct the hyperplanes. Then, a batch of mixed signals (healthy and damaged) is provided to algorithm perform outlier rejection and change detection.

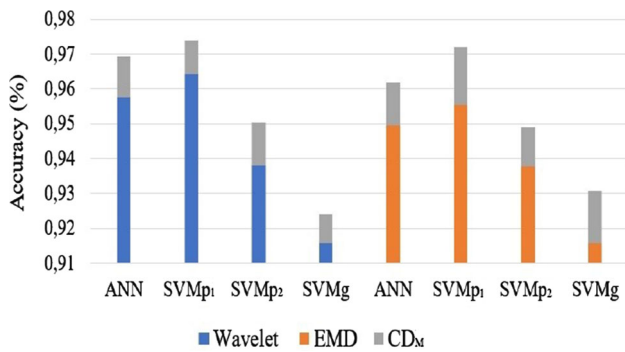
### 6.5 Results from Experimental Tests

The average accuracy classification with wavelets can be seen in Table 2. Although the results are similar, SVM $p_1$  has superior performance for earlier and distributed damage identification, while ANN has the best performance for punctual damage identification.

Next, Table 3 presents the bearing damage accuracy classification with EMD. The SVM $p_1$  algorithm has superior performance for earlier and punctual damages, while SVM $p_2$  reaches the best accuracy for distributed damages identification.

The wavelet transform and EMD present similar results with different algorithms, attesting that these methods can be applied in bearing diagnosis from current-based signals.





**Fig. 10** Intelligent algorithms accuracy improvement with  $CD_M$  (Color figure online)

**Table 4** On-site test accuracy with EMD and wavelet signal processing

	ANN	SVM $p_1$	SVM $p_2$	SVM $g$
EMD	0.915	0.939	0.887	0.873
Wavelet	0.948	0.951	0.923	0.899

Figure 10 compares the accuracy of wavelets and EMD associating the contribution of the  $CD_M$  as a feature.

With this feature, the average accuracy rises 1.1% for the wavelet signal processing and 1.45% for the EMD.

## 6.6 On-Site Tests Results

On-site tests were performed in two industrial motors of a gas processing facility. These motors have current-based condition monitoring without vibration data acquisition. One of these motors is kept in operation, while a spare remains on hot standby due to uninterrupted processing demand.

Initially, only the positive unlabeled database is available to collect data (healthy). Therefore, the stator current of these motors was monitored by the SVDD algorithm to identify changes in the data behavior. During testing, the SVDD algorithm identifies changes in the data distribution caused by wear.

This type of damage increases the bearings temperature and motor vibration until a pre-establish limit for maintenance scheduling. Therefore, this motor is kept operating, while the features acquired after wear damage identification are labeled as damaged.

In this context, the ANN and SVM algorithms have sufficient information to perform off-line training on healthy versus damage dichotomy with a labeled database. Table 4 presents the accuracy of the intelligent algorithms from wavelet and EMD.

The wavelet signal processing produces the superior result, and the SVM $p_1$  is the best algorithm to perform classification. When the bearing temperature surpasses the pre-established limit, the damaged bearing was replaced by

a new one. Then, SVDD, ANN, and SVM monitoring the new bearing behavior in real-time. These algorithms identify a new distributed bearing damage caused by overload misalignment a week later.

## 7 Conclusion

This paper proposes a new feature extraction approach for current-based signals using the Daubechies wavelet transform and empirical mode decomposition as signal processing techniques. The covering dimension is introduced as a feature that can improve the accuracy of the ANN and SVM for detecting bearing damage. The experimental results confirm that the detection of incipient damage is a difficult task, reaching 96.3% of accuracy for the EMD-SVM $p_1$  configuration, although other setups yield similar results.

The detection of punctual damages yields promising results, reaching 99.0% of accuracy for EMD-SVM $p_1$  and wavelet-ANN setup. Furthermore, distributed damages present better results for wavelet-SVM $p_1$  configuration, reaching 97.4% of accuracy, and therefore, the wavelet or EMD are appropriate for bearing damage detection with current-based signals and ANN or SVM as classifiers. Another relevant aspect to note is that the EMD produces quasi-orthogonal IMFs, which provide cross-information and therefore can reduce the efficiency of classifiers. Indeed, most of the experiments in this study demonstrate that Daubechies wavelets, which are always orthogonal by definition, perform better than EMD.

The on-site tests are performed in the SVDD framework to attain novelty detection when a labeled database is unavailable. Since the classifiers in this case are based on SVDD performance, these on-site experiments are predicted to perform slightly worse than supervised learning. As a result, the wavelet-SVM $p_1$  and EMDSVM $p_1$  configurations achieve 95.1% and 93.9% of accuracy, respectively, which are remarkable results in a positive unlabeled learning context. The main advantage of these methods is the capability to perform current-based bearing condition monitoring and earlier damage detection for different bearing damage types, damage location, and severity within a supervised learning or positive unlabeled learning context.

**Funding** This work was supported by the European Regional Development Fund (ERDF) through the Operational Programme for Competitiveness and Internationalization (COMPETE 2020), under Project POCI-01-0145FEDER-029494, and by National Funds through the FCT—Portuguese Foundation for Science and Technology, under Projects PTDC/EEI-EEE/29494/2017, UIDB/04131/2020, and UIDP/04131/2020.

## Declarations

**Conflict of interest** The authors declare that they have no conflict of interest.

## References

- Aimer, A. F., Boudinar, A. H., Benouzza, N., Bendiabdellah, A., & Mohammed-El-Amine, K. (2019). Bearing fault diagnosis of a pwm inverter fed-induction motor using an improved short time fourier transform. *Journal of Electrical Engineering and Technology*, 14(3), 1201–1210.
- Alligood, K. T., Sauer, T. D., Yorke, J. A., & Chillingworth, D. (1998). *Chaos: an introduction to dynamical systems* (Vol. 1). Society for Industrial and Applied Mathematics.
- Barcelos, A. S., & Cardoso, A. J. M. (2021). Current-based bearing fault diagnosis using deep learning algorithms. *Energies*, 14(9), 2509.
- Barcelos, A. S., Mazzoni, F. M., & Cardoso, A. J. M. (2021). Análise de avarias em rolamentos, utilizando algoritmos de inteligência artificial. *Brazilian Journal of Development*, 7(3), 29080–29093.
- Bayro-Corrochano E (2019) Applications of lie filters, quaternion fourier, and wavelet transforms. In: *Geometric Algebra Applications*, Springer, pp 489–517
- Bazan, G. H., Scalassara, P. R., Endo, W., Goedel, A., Godoy, W. F., & Palácios, R. H. C. (2017). Stator fault analysis of three-phase induction motors using information measures and artificial neural networks. *Electric Power Systems Research*, 143, 347–356.
- Benkedjough T, Medjaher K, Zerhouni N, Rechak S (2012) Fault prognostic of bearings by using support vector data description. In: IEEE Conference on Prognostics and Health Management
- Bessous, N., Zouzou, S., Bentrach, W., Sbaa, S., & Sahraoui, M. (2018). Diagnosis of bearing defects in induction motors using discrete wavelet transform. *International Journal of System Assurance Engineering and Management*, 9(2), 335–343.
- Bessous N, Sbaa S, Megherbi A (2019) Mechanical fault detection in rotating electrical machines using mcsa fft and mcsa-dwt techniques. *Bulletin of the Polish Academy of Sciences Technical Sciences* 67(3)
- Cao S, Wen L, Li X, Gao L (2018) Application of generative adversarial networks for intelligent fault diagnosis. In: IEEE 14th International Conference on Automation Science and Engineering, pp. 711–715
- Cardoso, A. J. M. (2018). *Diagnosis and fault tolerance of electrical machines, power electronics and drives*. IET.
- Cerrada, M., Sanchez, R. V., Li, C., Pacheco, F., Cabrera, D., de Oliveira, J. V., & Vásquez, R. E. (2018). A review on data driven fault severity assessment in rolling bearings. *Mechanical Systems and Signal Processing*, 99, 169–196.
- Chen, S., Meng, Y., Tang, H., Tian, Y., He, N., & Shao, C. (2020). Robust deep learning-based diagnosis of mixed faults in rotating machinery. *IEEE/ASME Transactions on Mechatronics*, 25(5), 2167–2176. <https://doi.org/10.1109/TMECH.2020.3007441>
- Chen, X., Feng, F., & Zhang, B. (2016). Weak fault feature extraction of rolling bearings based on an improved kurtogram. *Sensors*, 16(9), 1482.
- Daubechies, I. (1988). Orthonormal bases of compactly supported wavelets. *Communications on Pure and Applied Mathematics*, 41(7), 909–996.
- Desobry, F., Davy, M., & Doncarli, C. (2005). An online kernel change detection algorithm. *IEEE Transactions on Signal Processing*, 53(8), 2961–2974.
- Dragomiretskiy, K., & Zosso, D. (2013). Variational mode decomposition. *IEEE Transactions on Signal Processing*, 62(3), 531–544.
- Ghods, A., & Lee, H. H. (2016). Probabilistic frequency-domain discrete wavelet transform for better detection of bearing faults in induction motors. *Neurocomputing*, 188, 206–216.
- Gupta K, et al. (2019) Daubechies wavelets: Theory and applications. Master's thesis, Thepar Institute of engineering and technology
- Haykin SS, et al. (2009) Neural networks and learning machines/simon haykin.
- Hoang, D. T., & Kang, H. J. (2019). A motor current signal based bearing fault diagnosis using deep learning and information fusion. *IEEE Transactions on Instrumentation and Measurement*, 69(6), 3325–3333.
- Irfan, M. (2019). Modeling of fault frequencies for distributed damages in bearing raceways. *Journal of Nondestructive Evaluation*, 38(4), 98.
- Jiang, F., Zhu, Z., & Li, W. (2018). An improved vmd with empirical mode decomposition and its application in incipient fault detection of rolling bearing. *IEEE Access*, 6, 44483–44493.
- Kamiel BP, Howard I (2019) Ball bearing fault diagnosis using wavelet transform and principal component analysis. In: AIP Conference Proceedings, vol 2187, AIP Publishing LLC, p 50031
- Leite, V. C., da Silva, J. G. B., Veloso, G. F. C., da Silva, L. E. B., Lambert-Torres, G., Bonaldi, E. L., & de Oliveira, LEd. L. (2014). Detection of localized bearing faults in induction machines by spectral kurtosis and envelope analysis of stator current. *IEEE Transactions on Industrial Electronics*, 62(3), 1855–1865.
- Lessmeier C, Kimotho JK, Zimmer D, Sextro W (2016) Condition monitoring of bearing damage in electromechanical drive systems by using motor current signals of electric motors: A benchmark data set for data driven classification. In: *Proceedings of the European conference of the prognostics and health management society*, Citeseer, pp. 05–08
- Maragos, P., & Sun, F. K. (1993). Measuring the fractal dimension of signals: Morphological covers and iterative optimization. *IEEE Transactions on Signal Processing*, 41(1), 108.
- Merizalde, Y., Hernández-Callejo, L., & Duque-Perez, O. (2017). State of the art and trends in the monitoring, detection and diagnosis of failures in electric induction motors. *Energies*, 10(7), 1056.
- Moshrefzadeh, A., & Fasana, A. (2018). The autogram: An effective approach for selecting the optimal demodulation band in rolling element bearings diagnosis. *Mechanical Systems and Signal Processing*, 105, 294–318.
- Narendiranath, B. T., Himamshu, H., Prabin, K. N., Rama, P. D., & Nishant, C. (2017). Journal bearing fault detection based on daubechies wavelet. *Archives of Acoustics*, 42(3), 401–414.
- Noumir Z, Honeine P, Richard C (2012) On simple oneclass classification methods. In: IEEE International Symposium on Information Theory Proceedings, pp 2022–2026
- Randall, R. B. (2011). *Vibration-based condition monitoring: Industrial, aerospace and automotive applications*. Wiley.
- Rao SG, Lohith S, Gowda PC, Singh A, Rekha S (2019) Fault analysis of induction motor. In: 2019 IEEE International Conference on Intelligent Techniques in Control, Optimization and Signal Processing (INCOS), IEEE, pp. 1–4
- San, M. G., López, D. E., Meruane, V., & das Chagas Moura M. (2019). Deep variational auto-encoders: A promising tool for dimensionality reduction and ball bearing elements fault diagnosis. *Structural Health Monitoring*, 18(4), 1092–1128.
- Shen Z, Li J, Shen J, Zhang B (2020) Research on fault diagnosis of rolling bearings based on fractal dimension. In: 2020 7th International Conference on Dependable Systems and Their Applications (DSA), IEEE, pp. 441–446
- Wang, Z., Zhou, J., Wang, J., Du, W., Wang, J., Han, X., & He, G. (2019). A novel fault diagnosis method of gearbox based on maximum kurtosis spectral entropy deconvolution. *IEEE Access*, 7, 29520–29532.

- Witten IH, Frank E, Hall MA, Pal CJ (2016) *Data Mining: Practical machine learning tools and techniques*. Morgan Kaufmann
- Xia, M., Li, T., Xu, L., Liu, L., & de Silva, C. W. (2018). Fault diagnosis for rotating machinery using multiple sensors and convolutional neural networks. *IEEE/ASME Transactions on Mechatronics*, 23(1), 101–110. <https://doi.org/10.1109/TMECH.2017.2728371>
- Yu, K., Lin, T. R., Tan, J., & Ma, H. (2019). An adaptive sensitive frequency band selection method for empirical wavelet transform and its application in bearing fault diagnosis. *Measurement*, 134, 375–384.
- Zhang, W., Li, C., Peng, G., Chen, Y., & Zhang, Z. (2018). A deep convolutional neural network with new training methods for bearing fault diagnosis under noisy environment and different working load. *Mechanical Systems and Signal Processing*, 100, 439–453.
- Zhang, X., Luan, Z., & Liu, X. (2019). Fault diagnosis of rolling bearing based on kurtosis criterion VMD and modulo square threshold. *The Journal of Engineering*, 1(23), 8685–8690.
- Zhao, X., Jia, M., Ding, P., Yang, C., She, D., & Liu, Z. (2020). Intelligent fault diagnosis of multichannel motor rotor system based on multi manifold deep extreme learning machine. *IEEE/ASME Transactions on Mechatronics*, 25(5), 2177–2187. <https://doi.org/10.1109/TMECH.2020.3004589>

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.