# 12

# Artificial Neural Networks (ANNs)

## 12.1 Introduction

An artificial neural network (ANN) is a mathematical model that mimics biological neurons. A *neuron* is a special biological cell that process information. It receives signals using dendrites, processes the received signals using a cell body, and send signals to other neurons using an axon (Jain et al. 1996; Graupe 2013). An ANN is a group of connected nodes called *artificial neurons*, which mimic the neurons in a biological neural network (NN). In fact, an ANN often consists of a series of algorithms that work together to recognise underlying relationships in a set of data. The first model of neurons was formulated by McCulloch and Pitts (1943). It was a binary threshold unit computational model that computes a weighted sum of the input signals, $x_1, x_2, \ldots, x_n$, and produces an output of 1 if the weighted sum is above a given threshold. Otherwise, it produces an output of 0 (Jain et al. 1996). This model can be represented mathematically using the following equation,

$$y = f\left(\sum_{i=1}^{n}(w_i x_i - \tau\right) \tag{12.1}$$

where $f(.)$ is a unit step function at 0, $w_i$ is the weight of $x_i$, and $\tau$ is a threshold. This model has been developed in many ways. For example, Rosenblatt argued that the McCulloch and Pitts neuron model is not capable of learning because its parameters, i.e. the weight and the threshold coefficients, are fixed. Hence, the perceptron learning algorithm for the McCulloch and Pitts neuron model is introduced in (Rosenblatt 1958). Widrow and Hoff (1960) introduced the delta rule, also known as adaline, which is an adaptive linear neuron learning algorithm. Moreover, Cowan (1990) presented a short account of various investigations of NN properties in the period from 1943–1968. In 1982, Hopfield presented a model called neural networks based on the McCulloch and Pitts neuron and some characteristics of neurobiology and adapted to integrated circuits (Hopfield 1982). As described by Chua and Yang (1988), the basic characteristics of NNs are synchronous parallel processing, continuous-time dynamics, and global interaction of network elements.

The literature on NNs has highlighted various types of ANNs, e.g. multilayer perceptron (MLP), radial basis function (RBF), probabilistic neural network (PNN), etc. This chapter introduces some widely used ANN algorithms that have already been used for machine fault diagnosis using vibration signals. To begin with, this chapter presents

*Condition Monitoring with Vibration Signals: Compressive Sampling and Learning Algorithms for Rotating Machines,* First Edition. Hosameldin Ahmed and Asoke K. Nandi.

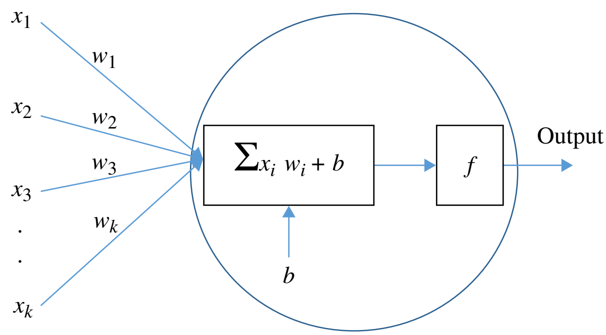


**Figure 12.1** Model of an artificial neuron (Ahmed and Nandi 2018).

essential concepts of ANNs; then the chapter describes three different types of ANN (i.e. MLP, RBF, and Kohonen) that can be used for fault classification. In addition, the applications of these methods in machine fault diagnosis are described.

## 12.2 Neural Network Basic Principles

An ANN is a group of connected nodes called artificial neurons, which mimic the neurons in a biological NN. It is a supervised learning algorithm that has the ability to learn real, discrete, and vector-valued target functions (Nandi et al. 2013). In fact, an ANN often consists of a series of algorithms that work together to recognise underlying relationships in a set of data. The neuron receive inputs, multiplies it by the weights of each input, and combines the results of the multiplication. Then, the combined multiplications of the signals and weights are passed to a transfer function to generate the output of the neuron (Figure 12.1).

As illustrated in Figure 12.1, the artificial neuron is a computational model that transform a set of input signals $X = x_1, x_2, \ldots, x_k$ into a single output using a structure of two functions as follows:

- A weighted sum function, which is a net value function that uses the inputs and their corresponding weights to produce a value ($v$) that summarises the input data such that,

$$v = \sum_{i=1}^{k} (w_i x_i + b) \tag{12.2}$$

- An activation function, which transfers $v$ into the output of the neuron such that

$$output = f(v) \tag{12.3}$$

The literature on ANNs has highlighted several types of artificial neurons based on the types of net value functions and activation functions. For example, a linear neuron that uses a weighted sum function as a net function and linear or piecewise linear function as an activation function, a sigmoid-based neuron that uses a weighted sum as a net function and a sigmoid function as an activation function, and a distance-based neuron that uses a distance function as a net function and a linear or piecewise linear as an activation function.

An ANN is a group of connected artificial neurons. Based on the type of connectivity between these neurons, various types of ANN architectures can be defined, as follows:

(1) *Layered network*. A layered network organises its neurons into hierarchical layers, which involves an input layer, hidden layer(s), and output layer. Each layer consists of a number of neurons that perform specific functions. Examples of layered networks include the feedforward neural network (FFNN) and multilayer perceptron (MLP), which involve an input layer, one to several hidden layers, and an output layer, and which often use a weighted sum function as a net function and a linear or sigmoid function as an activation function; RBF networks, which involve an input layer, a hidden radial basis layer, and an output linear layer that uses a Gaussian function as an activation function (Lei et al. 2009); and learning vector quantisation (LVQ) networks, which have a feedforward structure with a single computational layer of neurons where input neurons are connected directly to output neurons (Kohonen 1995).
(2) *Feedback network*. The feedback network, also called a recurrent or interactive network, often involves loops in the network. Typical examples of feedback networks include recurrent neural networks that consist of both feedforward and feedback connections between layers and neurons (Chow and Fang 1998).

The following subsections describe three different types of ANN (i.e. MLP, RBF, and Kohonen) that can be used for fault classification. In addition, the applications of these methods in machine fault diagnosis are described.

### 12.2.1 The Multilayer Perceptron

The MLP, also called a multilayer FFNN, involves an input layer, one to several hidden layers, and an output layer, and is one of the most commonly used NNs. It often uses a weighted sum function as a net function and a linear or sigmoid function as the activation function. An example of the MLP model, with an input layer, one hidden layer, and output layer, is illustrated in Figure 12.2. As shown in the figure, the layers are organised successively following the flow of data from the input layer that receives the input data $X = x_1, x_2, \ldots, x_k$ using its neurons and passes it to each neuron in the hidden layer, then processed using the hidden layer neurons and ending at the output layer. The process in

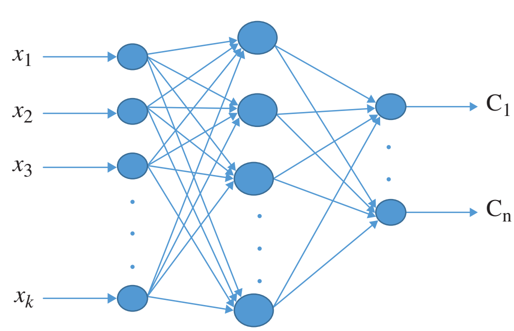


**Figure 12.2** A multilayer perceptron model for ANN (Ahmed and Nandi 2018).

each neuron of MLP can be represented mathematically using the following equation,

$$y = f(v(x)) = \sum_{i=1}^{k}(w_i x_i + b) \tag{12.4}$$

where $y$ is the output, the neuron has $k$ inputs, $w_i$ is a weight corresponding to the $i$th input $x_i$, and $b$ is a bias term. Then, the produced value of each neuron is sent to each neuron in the output layer, which acts as a classification step. In order to minimise the error between the MLP's output and the target output, many error metrics can be used to train the MLP, e.g. minimum classification error (MCE), mean square error (MSE), and least mean log square (Gish 1992; Liano 1996). The optimisation of the MLP training is based on minimising the following objective function:

$$J(w) = \int J(x; w)p(x)dx \tag{12.5}$$

Most techniques include selecting some initial value $w_0$ for the weight vector and then moving through the weight space in a sequence of steps that can be represented using Eq. (12.6) (Bishop 2006),

$$w_{\eta+1} = w_\eta + \Delta w_\eta \tag{12.6}$$

where $\eta$ is the iteration step. Different techniques use a different selection for the update of $\Delta w_\eta$. Most techniques utilise the gradient information. As described in Chapter 11, gradient descent optimisation can be used when the objective function is minimised. This can be done by selecting $\Delta w_\eta$ in Eq. (12.6) to encompass a small step in the direction of the negative gradient such that,

$$w_{\eta+1} = w_\eta - \mathcal{L}\nabla w_\eta \tag{12.7}$$

where $\mathcal{L} > 0$ is the learning rate. Here, after each update, the gradient is re-examined for the updated weight, and this process is repeated. Also, there are other methods such as conjugate gradients and quasi-Newton methods that can be used to optimise the objective function. Unlike gradient decent, these methods, i.e. conjugate gradient and quasi-Newton, decreases the error function at each iteration unless the weight vector has reached a local or global minimum.

To evaluate the gradient of error function $E(w)$, the backpropagation, which describes the evaluation of derivatives, is used. It is used to find a local minimum of the error function $E(w)$. Let $E_n(w)$ denote the error function for a specific input pattern $n$ such that

$$E_n(w) = \frac{1}{2}\sum_{i=1}^{k}(y_{nk} - t_{nk})^2 \tag{12.8}$$

where $y_{nk} = y_k(x_n, w)$. The gradient of this error function with respect to a weight $w_{ji}$ can be defined using the following equation:

$$\frac{\partial E_n(w)}{\partial w_{ji}} = \frac{1}{2}\sum_{i=1}^{k}(y_{nk} - t_{nk})^2 \tag{12.9}$$

Generally, each neuron computes a weighted sum of its inputs such that

$$a_j = \sum_i w_{ji} z_i \tag{12.10}$$

Here, $z_i$ is the activation of a neuron that sends a connection to unit $j$, and $w_{ji}$ is the weight corresponding to that connection. Also, we can define Eq. (12.9) as follows,

$$\frac{\partial E_n(w)}{\partial w_{ji}} = \frac{\partial E_n(w)}{\partial a_j}\frac{\partial a_j}{\partial w_{ji}} \tag{12.11}$$

where

$$\delta_j = \frac{\partial E_n(w)}{\partial a_j} \tag{12.12}$$

$$z_i = \frac{\partial a_j}{\partial w_{ji}} \tag{12.13}$$

Then,

$$\frac{\partial E_n(w)}{\partial w_{ji}} = \delta_j z_i \tag{12.14}$$

Similarly, we can define $\delta_j$ as

$$\delta_j = \frac{\partial E_n(w)}{\partial a_j} = \sum_k \frac{\partial E_n(w)}{\partial a_k}\frac{\partial a_k}{\partial a_j} \tag{12.15}$$

More efficiently,

$$\delta_j = h'(a_j)\sum_k w_{kj}\delta_k \tag{12.16}$$

Here, $h(a_j) = z_j$. To evaluate the derivatives, we first compute $\delta_j$ and then apply it to Eq. (12.14).

### 12.2.2 The Radial Basis Function Network

The RBF network is a layered network that has the same structure as MLP. It uses a Gaussian kernel function as the activation function. Assume a $c$-class classification problem with $K$ hidden neurons in the RBF and $c$ output nodes. The output function, i.e. the Gaussian function, of the neuron can be computed using the following equation,

$$f(x) = e^{\left(-\frac{\|x-\mu_k\|^2}{2\sigma^2}\right)} \tag{12.17}$$

where $\|x - \mu_k\|$ is the activity of the $k$th neuron, $\mu_k$ is the centre of the NN in the $k$th hidden neuron, and $\sigma$ is the width of the Gaussians. In the RBF NN, the input nodes pass the input signals to the connections arcs, and the first layer connections are not weighted, i.e. each hidden node (RBF unit) receives all the input values unchanged. The output can be calculated using Eq. (12.18):

$$y_c = \sum_{k=1}^{K} w_{k,c} f(x) \tag{12.18}$$

The training methodology of the RBF network is based on minimising the MSE between the output vector and the target vector. The RBF training process often consists of two steps: (i) define the hidden layer parameters, i.e. the set of centres and the number of hidden nodes, and (ii) determine the connection weights between the hidden layer and the output layer (Maglogiannis et al. 2008).

### 12.2.3 The Kohonen Network

The Kohonen neural network (KNN), also called a self-organising map (SOM), is a NN model and algorithm that implements characteristic nonlinear projections from high-dimensional space onto a low-dimensional array of neurons (Kohonen et al. 1996). In engineering, the most straightforward applications of the SOM are in the identification and monitoring of complex machine and process conditions. The SOM can be defined as a nonlinear, ordered, smooth mapping of a high-dimensional input data space onto the low-dimensional array. Let $X = \{x_1, x_2, \ldots, x_N\}^T$ be a set of input signals and $r_i = \{v_{i,1}, v_{i,2}, \ldots, v_{i,n}\}^T$ be a parametric real vector related to each element in the SOM array. We define the 'image' index $b$ of input vector x on the SOM array using the following equation,

$$b = arg \min_i \{d(x, r_i)\} \tag{12.19}$$

where $d(x, r_i)$ is the distance between $x$ and $r_i$. Our main task here is to determine $r_i$ such that the mapping is ordered and descriptive of the $x$ distribution. One way to determine $r_i$ is by using an optimisation process following the idea of vector quantisation (VQ). The basic idea is to place a codebook vector $r_c$ into the space of $x$ signals where $r_c$ is closest to $x$ in the signal space. VQ then minimises the expected quantisation error $E$ such that,

$$E = \int f[d(x, r_c)]p(x)dx \tag{12.20}$$

where $d(x, r_c)$ is the distance between the codebook vector and the signal, $x$, and $p(x)$ is the probability density function of $x$. The parameter $r_i$ that minimises $E$ is the solution of the VQ problem. An alternative technique to estimate $r_i$ that minimises $E$ is by using stochastic approximation such that,

$$E'(t) = \sum_i h_{ci} f[d(x(t), r_i(t))] \tag{12.21}$$

The approximate optimisation algorithm can be represented as

$$r_i(t+1) = r_i(t) - \frac{1}{2\lambda(t)\frac{\partial E'}{\partial r_i(t)}} \tag{12.22}$$

Various SOM algorithms can be represented using Eq. 12.20.

There are also many other methods, including fuzzy-ANNs (Filippetti et al. 2000), learning vector quantization (LVQ) (Wang and Too 2002), etc. Because of limitations on space, we cannot detail them all here. Readers who are interested are referred to the more comprehensive review (Meireles et al. 2003).

ANNs are being considered in a large diversity of applications, e.g. image processing (Egmont-Petersen et al. 2002), medical diagnosis (Amato et al. 2013), stock market index prediction (Guresen et al. 2011), etc. The following subsection briefly describes their application in machine fault diagnosis.

## 12.3 Application of Artificial Neural Networks in Machine Fault Diagnosis

For the application of ANN algorithms in machine fault diagnosis problems, the following selections concerning ANN system parameters have to be considered first. These include,

- Type of ANN to be used (distance-based or weighted sum–based NN)
- Number of layers and number of nodes in each layer
- Activation function to be used
- Training technique to be used
- Number of epochs to be used
- Validation technique to be used (different validation techniques will be described in Chapter 15)

Many research efforts have been dedicated to the application of NNs in condition monitoring and fault diagnosis. In an early example of research into machine diagnosis using ANNs, Knapp and Wang proposed a back-propagation NN for machine fault diagnosis (Knapp and Wang 1992). D'Antone (1994) implemented the MLP in a parallel architecture based on a transputer-based system to speed up the computation time of the network in an expert system used for fault diagnosis. Peck and Burrows (1994) presented an approach based on ANNs to predict and identify machine component conditions in advance of complete failure. In 1966, McCormick and Nandi described the use of ANNs as a method for automatically classifying the machine condition from the vibration time series. For example, a method for extracting features to use as ANN inputs is described in McCormick and Nandi (1996a). In this method, a combination of the horizontal and vertical vibration time series is used to produce a complex time series. Then, the zero-lag higher-order moments of the magnitude of this time series, its derivative, and its integral are estimated. Moreover, in (McCormick and Nandi 1996a,b, 1997b), statistical estimates of vibration signals such as the mean and variance are also used as indications of faults in rotating machinery. The authors stated that using these estimates jointly can give a more robust classification than using them individually as an input to ANNs.

Furthermore, McCormick and Nandi (1997a, b) described the usage of ANNs as a classifier in machine condition monitoring using vibration time series. In these studies, several methods for the extraction of features that used as inputs to ANNs are described and compared. Real vibration signals of a rotating shaft with four different machine condition – no faults (NO), only rub applied (N-R), only weight added, and both rub and weight applied (W-R) – are used to validate the proposed method. Based on their experimental results, the authors compared the ANN as a classification system with other methods – the thresholding classification and the nearest centroid classification – and found that thresholding was not an appropriate method as it could only use one feature to make a decision. The nearest-neighbour classification as a simple multi-feature technique performed slightly better; however, it requires large storage space and is very computationally expensive and is therefore not suitable for real-time implementation. Therefore, for a continuous monitoring system, the authors recommended the use of ANNs.

Subrahmanyam and Sujatha (1997) developed two NN techniques based on the supervised error back propagation (EBP) learning algorithm and unsupervised adaptive resonance theory-2 (ART2) based training paradigm, for recognizing bearing conditions. Real vibration signals acquired from a normal bearing and two different faulty bearings under several load and speed conditions are used to validate the proposed methods. A number of statistical parameters are obtained from the original vibration signals. The trained NNs are used for the identification of bearing conditions. The experimental results demonstrated the effectiveness of the proposed methods in roller bearing fault diagnosis. Li et al. (1998) presented an approach of using frequency-domain vibration signals and NN to detect common bearing faults from motor vibration data. In this study, vibration signals with varying severity caused by bearing looseness, IR fault, RE fault, and combinations of them are used to validate the proposed approach. The fast Fourier transform (FFT) is employed to extract features from these vibration signals. With these extracted features, a three-layer NN with 10 hidden nodes is used to deal with the fault-detection problem. The experimental results showed that NNs can be used effectively in the detection of various bearing faults using vibration signals from motor bearings.

Jack and Nandi (1999, 2000) examined the use of a genetic algorithm (GA) to select the most-significant input features for ANNs from a large set of possible features in machine condition monitoring contexts. Real vibration data acquired from bearings with a total of six different conditions – two normal conditions and four fault conditions – are used to validate their proposed method. The experimental results showed that the GA is able to selecting a subset of 6 inputs from a set of 156 features that allow the ANN to perform with 100% accuracy. In the same vein, Samanta et al. (2001) presented a procedure for diagnosing gear conditions using GAs and ANNs. In this study, the time domain vibration signals of a rotating machine with normal and faulty gears are processed for feature extraction using the mean, the root mean square, the variance, skewness, and kurtosis. The selection of input features and the number of nodes in the hidden layer of the ANN are optimised using a GA-based technique in combination with the ANN. The output layer consists of two binary nodes indicating the condition of the machine, i.e. normal or faulty gears. The experimental results showed the effectiveness of the proposed method in machine fault detection. Also, Saxena and Saad (2007) presented the results of their investigation into the use of GAs and ANNs for condition monitoring of mechanical systems. In this study, the raw vibration signals are first normalised, and then five feature sets are considered: (i) statistical features from the raw vibration signals; (ii) statistical features from the sum signal; (iii) statistical features from the difference signals; (iv) spectral features; and (v) all the features considered together. Then, a GA is used to select the best features of all the considered features. With these selected features, the ANN is used to deal with the fault classification problem. Vibration signals acquired from bearings are used to examine the effectiveness of the proposed method. The results showed that GA-evolved ANNs clearly outperform the stand-alone ANNs.

An alternative method to optimise ANN-based fault diagnosis techniques is to use the UTA feature-selection algorithm. To investigate the effectiveness of UTA compared to a GA, Hajnayeb et al. (2011) designed an ANN-based fault-diagnosis system for a gearbox using a number of vibration features. In this study, several statistical features – maximum value, root mean square, kurtosis, crest factor, etc. – extracted from the time domain vibration signals are used as characteristics features. Then, UTA and a GA are used to select the best subset features from the extracted features. The results showed that the UTA method is as accurate as the GA, despite its simple algorithm.

Filippetti et al. (2000) presented a review of artificial intelligence (AI) based methods for diagnosing faults in induction motor drives. This review covers the application of expert systems, ANNs, and fuzzy logic systems and combined systems, e.g. (fuzzy-ANNs). Moreover, Li et al. (Li et al. 2000) presented a method for motor bearing fault diagnosis using NNs and time/frequency domain bearing vibration analysis. In this study, the authors used real and simulated vibration signals with looseness, IR, and ball faults to validate their proposed method. The experimental results demonstrated that ANNs can be effectively used in the fault diagnosis of motor bearings through appropriate measurement and interpretation of motor bearing vibration signals.

Vyas and Satishkumar (2001) presented an ANN-based design for fault recognition in a rotor-bearing system. In this study, a back-propagation learning algorithm (BPA) and a multilayer network, which contained layers created of nonlinear neurons and a normalisation technique, are used. Real vibration signals with several fault conditions – rotor with no fault, rotor with mass unbalance, rotor with bearing cap loose, rotor with misalignment, play in spider coupling, and rotor with both mass unbalance and misalignment – collected from a laboratory rotor rig are used to validate the proposed method. Statistical moments of the collected vibration signals are employed to train the ANNs. Several experiments are conducted to investigate the adaptability of different network architectures, and an overall success rate up to 90% is achieved.

Kaewkongka et al. (2001) described a method for rotor dynamic machine condition monitoring using the continuous wavelet transform (CWT) and ANNs. In this method, the CWT is used to extract features from collected vibration signals. The extracted features are then converted into a greyscale image that is used as a characteristic feature of each signal. Four types of machine operating conditions are investigated: balanced shaft, unbalanced shaft, misalignment shaft, and faulty bearing. The backpropagation neural network (BPNN) is used as a classification tool. The experimental results showed that the proposed method achieved a classification accuracy of 90%. Yang et al. (2002b) introduced third-order spectral techniques for the diagnosis of motor bearing conditions using ANNs. In this study, seven techniques based on the power spectrum, bispectrum, bicoherence spectrum, bispectrum diagonal slice, bicoherence diagonal slice, summed bispectrum, and summed bicoherence are examined as signal preprocessing techniques in fault diagnosis of induction motor roller bearings. Four health conditions of the bearings – normal, cage fault, IR fault, and OR fault – are considered in this study. The obtained features are used as inputs to the ANN to identify the bearing conditions. The experimental results demonstrated that the method based on the summed bispectrum achieved better results than the other six methods introduced in this study.

Hoffman and Van Der Merwe evaluated the performance of three different neural classification techniques – Kohonen SOMs, nearest neighbour rule (NNR), and RBF networks – on multiple fault-diagnosis problems. In this study, the authors used real vibration measurements acquired from roller bearings with six classes: three classes with no bearing fault and imbalance masses of 0, 12, and 24 g; and three classes with bearing OR faults and imbalance masses of 0, 12, 24 g. Then, several features are extracted from the collected vibration signals, including six time domain features – mean, root mean square, crest factor, variance, skewness, and kurtosis – and four frequency domain features – amplitude at rotational frequency in horizontal and vertical directions, ball pass frequency of the OR (BPFO) of the faulty bearing, higher-frequency domain components (HFDs), and BPFO obtained from the envelope spectra. These features are then normalised and used as input to the SOM, NNR, and RBF. Based on the experimental results, the authors demonstrated that the SOM can be used for multiclass problems if the class labels are known and the features are chosen and normalised correctly. NNR and RBF can accurately discriminate between different combinations of multiple fault conditions as well as identify the severity of fault conditions. Also, RBF classifiers have a speed advantage over NNR (Hoffman and Van Der Merwe 2002).

Wang and Too (2002) proposed a method for machine fault detection using higher-order statistics (HOS) and ANNs. In this method, HOS is used to extract features from vibration signals. With these extracted features, two types of NN (SOM and LVQ) are used to deal with the classification problem. SOM is used for collecting data at the initial stage, and LVQ is used at the recognition stage. Vibration signals acquired from a rotating pump with eight different health conditions are used to validate the proposed method. The experimental results demonstrated the effectiveness of the proposed method in recognizing rotating machine faults. Yang et al. (2002a) presented a review of a variety of AI-based diagnosis methods used for rotating machinery condition monitoring. Of these methods, the authors presented a review of eight types of NNs applied in diagnosing rotating machinery faults: backpropagation for feed-forward networks (BPFFN), FFNN, recurrent neural network (RNN), RBF, back propagation (BP), MLP, Kohonen SOM, and LVQ.

Shuting et al. introduced an adaptive RBF network called a two-level cluster algorithm for fault diagnosis of a turbine-generator using vibration signals (Shuting et al. 2002). This method can automatically calculate the number, centre, and width of hidden layer neurons in the RBF network. In this study, practical acquired vibration data with three conditions – normal operation, rotor excitation winding short circuit, and stator winding fault – are used to validate the introduced method. The results verified that the proposed method achieved high diagnosis precision in turbo-generator vibration fault diagnosis. Samanta et al. (2003) presented a study that compared the performance of bearing fault detection using ANNs and support vector machines (SVMs) with all the considered signal features and fixed parameters of the two classifiers. In this study, several statistical features of both the original vibration signals and some with preprocessing – such as differentiation and integration, low- and high-pass filtering, and spectral data of the signals – are used. GA is used for selection of input features and classifier parameters. Experimental vibration signals collected from roller bearings are used to validate the proposed methods. The results demonstrated that ANNs and SVMs achieved 100% classification accuracy using only six GA-based features.

Wu and Chow (2004) developed an RBF NN-based fault-detection technique for induction machine fault detection. In this study, after the collected vibration signals are transformed into the frequency domain using FFT, four feature vectors are extracted from the power spectra of the machine vibration signal. Then, the extracted features are used as inputs for the RBF NN for fault detection and classification. The authors proposed a cell-splitting grid algorithm to automatically determine the optimal network architecture of the RBF network. Unbalanced electrical faults and mechanical faults operating at different rotating speeds are used to validate the proposed method. The results showed the effectiveness of the proposed method for induction machine fault detection.

Yang et al. (2004) proposed a method for diagnosing faults in rotating machinery based on NN, which synthesises ART and the learning strategy of KNN. In this study, first, the discrete wavelet transform (DWT) is used to decompose the time domain signal into three levels. Then the transformed signal and the original signal are estimated using eight parameters: mean, standard deviation, root mean square, shape factor, skewness, kurtosis, crest factor, and entropy. Vibration signals acquired from a machinery fault simulator are used to test the proposed method. The experimental results showed the proposed method's effectiveness.

Guo et al. (2005) proposed a genetic programming (GP) based method for feature extraction from raw vibration data recorded from a rotating machine with six different conditions. Then, the extracted features are used as the inputs to ANN and SVM classifiers for the recognition of the six bearing conditions. The experimental results demonstrated that the proposed method achieved improved classification results compared with those using extracted features by classical methods. Castro et al. (2006) presented a method for diagnosing bearing faults based on ANNs and DWT. In this study, the authors used DWT to extract features from the original signals. Then, the extracted features are used as inputs to three different NNs – MLP, RBF, and PNN – to identify the bearing conditions. Vibration signals collected from motor bearings with a normal condition, IR fault, OR fault, and ball fault are used to validate the proposed method. The results showed that the PNN achieved better classification results than MLP and RBF.

Rafiee et al. (2007), presented an ANN-based process for fault detection and identification in gearboxes using a feature vector extracted from the standard deviation of wavelet packet coefficients of vibration signals. In this process, first, the collected vibration signals are preprocessed using the following steps: (i) synchronisation of vibration signals, achieved by applying interpolation using piecewise cubic Hermite interpolation (PCHI), and (ii) feature extraction using the standard deviation of wavelet packet coefficients. Then, a two-layer MLP made up of an input layer, a hidden layer, and an output layer is used to deal with the fault-diagnosis problem. Experimental vibration signals collected from experimental testing of the gearbox are used to test the proposed method. The results showed that an MLP network with a 16 : 20 : 5 structure achieved 100% accuracy for identifying gear failures.

Sanz et al. (2007) presented a technique for rotating machine condition monitoring, combining auto-associative neural networks (AANNs) and a wavelet transform (WT) using vibration analysis. In this study, DWT is used to extract features from the collected vibration signals; then, from these features, the identification of significant changes is performed using AANNs. Real vibration signals obtained under two different operational conditions corresponding to high and low loads are used to test the proposed method. The results demonstrated the effectiveness of the proposed method in online fault diagnosis for rotating machinery.

Yang et al. (2008), presented a fault-diagnosis method for wind turbine gearboxes, based on ANN. In this method, to separate noise signals from tested vibration signals, the collected vibration signals are decomposed into four-level details by using wavelet decomposition and then reconstructing the real signals. Then, a three-layer BPNN is used to diagnosis the health condition of the gearbox. Experimental vibration signals with four kinds of typical patterns of gearbox faults, collected from a gearbox diagnostic testbed, are used to examine the proposed method. The results indicated that the BPNN is an efficient tool to solve complicated state-identification problems in gearbox fault diagnosis.

Al-Raheem et al. (2008) introduced a technique for detecting and diagnosing faults in rolling bearings using the Laplace-wavelet transform and ANNs. In this method, the Laplace-wavelet transform is used to extract features from the time domain vibration signals of rolling bearings. Then, the extracted features are used as input to ANNs for rolling bearing fault classification. The parameters of the Laplace-wavelet shape and the ANN classifier are optimised using a GA algorithm. Real and simulated bearing vibration data are used to test the effectiveness of the proposed method. The results showed the effectiveness of the proposed method in diagnosing rolling bearing faults.

Tyagi (2008) presented a comparative study of SVM classifiers and ANNs' application for rolling bearing fault diagnosis using WT as a preprocessing technique. In this study, (i) the features of the collected vibration signals are extracted using statistical features such as standard deviation, skewness, kurtosis, etc. Then, (ii) the extracted features are used as inputs to the SVM and ANN classifiers. Moreover, the effects of a preprocessing step using DWT prior to the feature-extraction step is also studied. Experimental vibration signals collected from bearings with four health conditions – normal, OR fault, IR fault, and roller fault – are used to validate the proposed method. The results showed that using the simple statistical features extracted from the time domain vibration signals and ANN or SVM, the bearing conditions can be correctly identified. Also, the results demonstrated that preprocessing with DWT improved the performance of both ANN and SVM classifiers in diagnosing rolling bearing faults.

Sreejith et al. (2008), introduced a fault-diagnosis method for rolling bearings using time domain features and ANNs. In this method, normal negative log-likelihood (Nnl) and kurtosis (KURT) are extracted from the time domain vibration signals, and then Nnl and KURT are used as input features for FFNN to diagnose faults in rolling bearings. The time domain vibration signals used in this study are acquired for four different health conditions of the rolling bearings: normal, RE fault, OR fault, and IR fault. The results demonstrated the effectiveness of the proposed method in diagnosing rolling bearing faults.

Li et al. (2009) applied order cepstrum and RBF NN for gear fault detection during a speed-up process. In this method, the time domain vibration signal during a gearbox speed-up process is sampled at constant time increments and then resampled at constant angle increments. Then, cepstrum analysis is used to process the resampled signal. For feature extraction, the order cepstrum with normal condition, wear, and a crack fault is processed. The extracted features are used as inputs to RBF for fault recognition. Experimental vibration data acquired from a gearbox are used to examine the proposed method. The results demonstrated the effectiveness of the proposed method in gear fault detection and identification.

Lei et al. (2009), introduced a method for intelligent fault diagnosis of rotating machinery based on the wavelet packet transform (WPT), empirical mode decomposition (EMD), dimensionless parameters, and RBF NN. In this method, WPT and EMD are used to preprocess the time domain vibration signals. Then, the dimensionless parameters are extracted from the original vibration signal and the preprocessed signals, which are formed into a combined feature set; their sensitivities are evaluated using the distance-evaluation technique. The sensitive features are selected and used as inputs to the RBF NN classifier. Vibration signals acquired from rolling bearings with a normal condition, IR fault, OR fault, and roller fault; and from a heavy oil catalytic cracking unit with a normal condition, a large area of rub fault, and a slight rub fault are used to validate the proposed method. The results showed the effectiveness of the proposed method in fault diagnosis for the rotating machine.

Saravanan et al. (2010) investigated the effectiveness of Morlet wavelet-based features for fault diagnosis of a gearbox using ANNs and proximal support vector machines (PSVMs), and the ability to use a Morlet wavelet in feature extraction. In this study, the Morlet wavelet is used to extract features from the collected vibration signals in the time domain; several statistical features such as kurtosis, standard deviation, peak, etc., are extracted from the Morlet coefficients. Then, the J48 is used to select the best features from the extracted features. Finally, the selected features are used as input to the ANN and PSVM for fault classification. Real vibration signals collected from a gearbox with good condition, gear tooth breakage (GTB), a gear with a crack at the root (GTC), and a gear with face wear are used to examine the proposed methods. The results showed that PSVM has an edge over ANN in the classification of features.

Castejón et al. (2010) developed a method for classifying bearing conditions using multiresolution analysis (MRA) and ANNs. In this study, the authors argued that the WT cannot be used practically by using analytical equations; hence, a discretisation process is needed. MRA is used to perform the discretisation. The mother wavelet, Daubechies-6, is used for feature extraction, and the fifth-level detail coefficients (cD5) are selected as characteristic features, which are normalised in the range [−1 1]. With these features, the MLP NN is used for bearing conditions classification. Four sets of experimental vibrations acquired from a roller bearing experimental system with a normal condition, IR fault, OR fault, and ball fault are used to validate the proposed method. The results showed that the proposed method is sound and detects four bearing conditions in a very incipient stage.

De Moura et al. (2011) evaluated principal component analysis (PCA) and ANN performance for diagnosing rolling bearing faults using vibration signals preprocessed by detrended-fluctuation analysis (DFA) and rescaled-range analysis (RSA). In this study, PCA and ANN are combined with DFA and RSA in a total of four approaches for diagnosing bearing faults. Vibration signals acquired from bearings with four health conditions are used to examine the proposed methods. The results demonstrated that the ANN-based classifier presented performance slightly better that the one based on PCA.

Bin et al. (2012) proposed a method for fault diagnosis of rotating machinery based on wavelet packets, EMD, and ANN. In this method, using vibration signals collected from an experimental rotor-bearing system with ten rotor failures, four stages of investigations are carried out. These are (i) wavelet packet decomposition is used to denoise the collected vibration signal; (ii) EMD is employed to obtain a series of intrinsic mode functions (IMFs) from the denoised signal; (iii) the moment of the energy of IMFs is calculated to express the failure feature; and (iv) a three-layer BPNN with fault feature from the frequency domain is employed as the target input of the NN. The energy in five spectral bandwidths of the vibration spectrum is taken as characteristic parameters, and ten types of representative rotor faults are taken as the output. The results showed the effectiveness of the proposed method in early fault diagnosis of rotating machinery.

Liang et al. (2013) introduced the application of the power spectrum, cepstrum, bispectrum, and ANN for fault pattern extraction from induction motors. This study compared the effectiveness of the power spectrum, cepstrum, and bispectrum for vibration, phase current, and transient speed analyses for detection and diagnostics of induction motor faults, based on experimental results. The authors found that for vibration signals, the power spectrum, cepstrum, and bispectrum presented a better ability to identify induction motors faults if the fault symptoms demonstrated characteristics in rich sidebands and harmonics. In addition, the authors stated that a combination of the power spectrum, cepstrum, and HOS methods along with ANN analysis should undoubtedly provide a better tool for condition monitoring and fault diagnosis of induction motors.

Ertunc et al. (2013) proposed a multistage method for detection and diagnosis of rolling bearing faults based on ANN and adaptive neuro-fuzzy inference system (ANFIS) techniques. In this study, both time and frequency domain parameters extracted from the vibration and current signals are used as inputs to ANN and ANFIS models. Experimental data acquired from a shaft-bearing system is used to examine the performance of the two approaches for diagnosing rolling bearing faults. The experimental results demonstrated that the ANFIS-based approach is superior to the ANN-based approach in diagnosing fault severity.

Zhang et al. (2013) proposed a method for classification of faults and prediction of degradation of components and machines in manufacturing systems using wavelet packet decomposition, FFT, and BPNN. In this method, the collected vibration signals were decomposed into several signals using wavelet.. Then, these signals were transformed to the frequency domain using FFT. Finally, the features extracted in the frequency domain were used as input to the BPNN. The peak values of FFT are selected as features to judge the degradation of the monitored machine. A case study was used to illustrate the proposed method, and the results showed the effectiveness of the proposed method.

Unal et al. (2014), presented a method for diagnostics of rolling bearing faults using envelope analysis, the FFT, and the FFNN. In this study, the authors suggested some methods to extract features using envelope analysis accompanied by the Hilbert transform (HHT) and FFT. Then, the extracted features are used as inputs to GA-based FFNN. Vibration signals, collected from rolling bearings in the experimental setup, are used to validate the proposed method. The experimental results verified the effectiveness of the proposed method in diagnosing faults in the rolling bearing.

Ali et al. (2015) introduced an approach for diagnosing bearing faults based on statistical features, EMD energy entropy, and ANN. In this techniques, (i) 10 statistical features are extracted from the time domain vibration signals. (ii) In addition to these 10-time domain features, EMD is used to extract some other features to form robust and reliable features. Finally, (iii) ANN is adopted to identify bearing health conditions. Also, the authors proposed a health index (HI) for online damage detection at an early stage. Three bearing run-to-failure vibration signals with a roller fault, an IR fault, and an OR fault are used to examine the effectiveness of the proposed method. The results showed that the proposed method achieved high classification accuracy in diagnosing bearing faults.

Bangalore and Tjernberg (2015) introduced a self-evolving maintenance-scheduler framework for maintenance management of wind turbines and proposed an ANN-based condition monitoring technique using data from a supervisory control and data acquisition system (SCADA). The Levenberg–Marquardt back-propagation (LM) training algorithm is used for training the NN. This ANN-based condition monitoring technique is applied to gearbox bearings with real data from onshore wind turbines. The results showed that the proposed technique is capable of identifying damage in the gearbox bearings almost a week before the vibration-based condition monitoring system (CMS) raised an alarm.

Janssens et al. (2016) proposed a feature-learning model for condition monitoring based on convolutional neural networks (CNNs). The CNN model is not applied to extracted features such as kurtosis, skewness, mean, etc. but to the raw amplitudes of the frequency spectrum of the vibration data. By applying CNN on the raw data, the network learns transformations on the data that result in better representation of the data for the fault-classification task in the output layer. Vibration signals collected from rolling bearings with eight health conditions are used to validate the proposed technique. The results showed that the feature-learning system based on CNN significantly outperformed the classical feature-engineering based approach that utilises manually engineered features and a random forest classifier. Another feature-learning technique for condition monitoring is introduced by Lei et al. (2016). In this study, the authors introduced a two-stage learning method for intelligent diagnosis of machines. In the first stage of this method, sparse filtering, which is an unsupervised two-layer NN, is used to directly learn features from the collected mechanical vibration signals. In the second stage, softmax regression, i.e. logistic regression, is employed to classify health conditions based on the learned features. Two vibration datasets acquired from a motor bearing and a locomotive bearing are used to validate the proposed method. The results demonstrated that the proposed method obtained high diagnosis accuracies.

Recently, Han et al. (2018) explored the performances of random forest, SVM, and two advanced ANNs – an extreme learning machine (ELM) and PNN – with different features using two datasets from rotating machinery. The results showed that random forest outperformed the comparative classifiers in terms of recognition accuracy, stability, and robustness of features, in particular with a small training set. Moreover, Ahmed and Nandi (2018) proposed combined compressive sampling (CS) based on a multiple-measurement vector (MMV) and feature-ranking framework to learn optimally fewer features from a large amount of vibration data from which bearing health conditions can be classified. Based on this framework, the authors investigated different combinations of MMV-based CS and feature-ranking techniques to learn features from vibration signals acquired from rolling bearings. Three classification algorithms (multinomial logistic regression (MLR), ANN, and SVM) are tested to evaluate the proposed framework for the classification of bearing faults. The experimental results showed that the proposed framework is able to achieve high classification accuracy in all the faults considered in this study. In particular, results from this proposed framework with MLR and ANN showed the efficiency of this framework with high classification accuracies using different values of the sampling rate ($\alpha$) and a number of selected features ($k$) for all the considered CS and feature-selection technique combinations.

## 12.4 Summary

This chapter has presented essential concepts of ANNs and has described three different types of ANN (i.e. MLP, RBF network, and Kohonen network) that can be used for fault classification. In addition, the applications of these methods and several other types of ANN-based methods in machine fault diagnosis were described. A considerable

**Table 12.1** Summary of some of the introduced techniques and their publically accessible software.

| Algorithm name | Platform | Package | Function |
|---|---|---|---|
| Perceptron | MATLAB | Deep Learning Toolbox – Define shallow neural network architecture | Perceptron |
| MLP neural network trained by backpropagation | | (Chen 2018) | mlpReg mlpRegPred |
| Radial basis function with k means clustering | | (Shujaat 2014) | RBF |
| Design probabilistic neural network | | Deep Learning Toolbox – Define shallow neural network architecture | newpnn |
| Train shallow neural network | | Deep Learning Toolbox – Function Approximation and Clustering | train |
| Self-organising map | | Deep Learning Toolbox – Function Approximation and Clustering-Self-organising maps | selforgmap |
| Gradient descent backpropagation | | | net.trainFcn = 'traingd' |

amount of literature has been published on the application of ANNs and variants in machine fault diagnosis. Most of these studies introduced many preprocessing techniques that include normalisation, feature selection, transformation, and feature extraction. The produced data of the preprocessing step represent the final training set that is used as input to ANNs. In order to learn more useful features for machine fault diagnosis, most of the proposed methods combine two or more analysis techniques. For example, the GA algorithm along with various types of time domain statistical features, frequency domain, and time-frequency domain features have been widely used with different types of ANNs. Hence, the challenge will always remain to produce possible approaches to machine condition monitoring capable of improving fault diagnosis accuracy and reducing computations. Most of the introduced techniques and their publicly accessible software are summarised in Table 12.1.

## References

Ahmed, H. and Nandi, A.K. (2018). Compressive sampling and feature ranking framework for bearing fault classification with vibration signals. *IEEE Access* 6: 44731–44746.

Ali, J.B., Fnaiech, N., Saidi, L. et al. (2015). Application of empirical mode decomposition and artificial neural network for automatic bearing fault diagnosis based on vibration signals. *Applied Acoustics* 89: 16–27.

Al-Raheem, K.F., Roy, A., Ramachandran, K.P. et al. (2008). Application of the Laplace-wavelet combined with ANN for rolling bearing fault diagnosis. *Journal of Vibration and Acoustics* 130 (5): 051007.

Amato, F., López, A., Peña-Méndez, E.M. et al. (2013). Artificial neural networks in medical diagnosis. *Journal of Applied Biomedicine* 11: 47–58.

Bangalore, P. and Tjernberg, L.B. (2015). An artificial neural network approach for early fault detection of gearbox bearings. *IEEE Transactions on Smart Grid* 6 (2): 980–987.

Bin, G.F., Gao, J.J., Li, X.J., and Dhillon, B.S. (2012). Early fault diagnosis of rotating machinery based on wavelet packets—empirical mode decomposition feature extraction and neural network. *Mechanical Systems and Signal Processing* 27: 696–711.

Bishop, C. (2006). *Pattern Recognition and Machine Learning*, vol. 4. New York: Springer.

Castejón, C., Lara, O., and García-Prada, J.C. (2010). Automated diagnosis of rolling bearings using MRA and neural networks. *Mechanical Systems and Signal Processing* 24 (1): 289–299.

Castro, O.J.L., Sisamón, C.C., and Prada, J.C.G. (2006). Bearing fault diagnosis based on neural network classification and wavelet transform. In: *Proceedings of the 6th WSEAS International Conference on Wavelet Analysis & Multi-Rate Systems*, 16–18. http://www.wseas.us/e-library/conferences/2006bucharest/papers/518-473.pdf.

Chen, M. (2018). MLP neural network trained by backpropagation. Mathworks File Exchange Center. https://uk.mathworks.com/matlabcentral/fileexchange/55946-mlp-neural-network-trained-by-backpropagation.

Chow, T.W. and Fang, Y. (1998). A recurrent neural-network-based real-time learning control strategy applying to nonlinear systems with unknown dynamics. *IEEE Transactions on Industrial Electronics* 45 (1): 151–161.

Chua, L.O. and Yang, L. (1988). Cellular neural networks: theory. *IEEE Transactions on Circuits and Systems* 35 (10): 1257–1272.

Cowan, J.D. (1990). Neural networks: the early days. In: *Advances in Neural Information Processing Systems*, 828–842. http://papers.nips.cc/paper/198-neural-networks-the-early-days.pdf.

D'Antone, I. (1994). A parallel neural network implementation in a distributed fault diagnosis system. *Microprocessing and microprogramming* 40 (5): 305–313.

De Moura, E.P., Souto, C.R., Silva, A.A., and Irmao, M.A.S. (2011). Evaluation of principal component analysis and neural network performance for bearing fault diagnosis from vibration signal processed by RS and DF analyses. *Mechanical Systems and Signal Processing* 25 (5): 1765–1772.

Egmont-Petersen, M., de Ridder, D., and Handels, H. (2002). Image processing with neural networks—a review. *Pattern Recognition* 35 (10): 2279–2301.

Ertunc, H.M., Ocak, H., and Aliustaoglu, C. (2013). ANN-and ANFIS-based multi-staged decision algorithm for the detection and diagnosis of bearing faults. *Neural Computing and Applications* 22 (1): 435–446.

Filippetti, F., Franceschini, G., Tassoni, C., and Vas, P. (2000). Recent developments of induction motor drives fault diagnosis using AI techniques. *IEEE transactions on industrial electronics* 47 (5): 994–1004.

Gish, H. (1992). A minimum classification error, maximum likelihood, neural network. In: *1992 IEEE International Conference on Acoustics, Speech, and Signal Processing, 1992. ICASSP-92*, vol. 2, 289–292. IEEE.

Graupe, D. (2013). *Principles of Artificial Neural Networks*, vol. 7. World Scientific.

Guo, H., Jack, L.B., and Nandi, A.K. (2005). Feature generation using genetic programming with application to fault classification. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)* 35 (1): 89–99.

Guresen, E., Kayakutlu, G., and Daim, T.U. (2011). Using artificial neural network models in stock market index prediction. *Expert Systems with Applications* 38 (8): 10389–10397.

Hajnayeb, A., Ghasemloonia, A., Khadem, S.E., and Moradi, M.H. (2011). Application and comparison of an ANN-based feature selection method and the genetic algorithm in gearbox fault diagnosis. *Expert Systems with Applications* 38 (8): 10205–10209.

Han, T., Jiang, D., Zhao, Q. et al. (2018). Comparison of random forest, artificial neural networks and support vector machine for intelligent diagnosis of rotating machinery. *Transactions of the Institute of Measurement and Control* 40 (8): 2681–2693.

Hoffman, A.J. and Van Der Merwe, N.T. (2002). The application of neural networks to vibrational diagnostics for multiple fault conditions. *Computer Standards & Interfaces* 24 (2): 139–149.

Hopfield, J.J. (1982). Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences of the United Stated of America* 79 (8): 2554–2558.

Jack, L.B. and Nandi, A.K. (1999). Feature selection for ANNs using genetic algorithms in condition monitoring. In: *ESANN European Symposium on Artificial Neural NetworksBruges (Belgium)*, 313–318. https://www.researchgate.net/publication/221165522_Feature_selection_for_ANNs_using_genetic_algorithms_in_condition_monitoring.

Jack, L.B. and Nandi, A.K. (2000). Genetic algorithms for feature selection in machine condition monitoring with vibration signals. *IEE Proceedings-Vision, Image and Signal Processing* 147 (3): 205–212.

Jain, A.K., Mao, J., and Mohiuddin, K.M. (1996). Artificial neural networks: a tutorial. *Computer* 29 (3): 31–44.

Janssens, O., Slavkovikj, V., Vervisch, B. et al. (2016). Convolutional neural network based fault detection for rotating machinery. *Journal of Sound and Vibration* 377: 331–345.

Kaewkongka, T., Au, Y.J., Rakowski, R., and Jones, B.E. (2001). Continuous wavelet transform and neural network for condition monitoring of rotodynamic machinery. In: *Instrumentation and Measurement Technology Conference, 2001. IMTC 2001. Proceedings of the 18th IEEE*, vol. 3, 1962–1966. IEEE.

Knapp, G.M. and Wang, H.P. (1992). Machine fault classification: a neural network approach. *International Journal of Production Research* 30 (4): 811–823.

Kohonen, T. (1995). Learning vector quantization. In: *Self-Organizing Maps*, 175–189. Berlin, Heidelberg: Springer.

Kohonen, T., Oja, E., Simula, O. et al. (1996). Engineering applications of the self-organizing map. *Proceedings of the IEEE* 84 (10): 1358–1384.

Lei, Y., He, Z., and Zi, Y. (2009). Application of an intelligent classification method to mechanical fault diagnosis. *Expert Systems with Applications* 36 (6): 9941–9948.

Lei, Y., Jia, F., Lin, J. et al. (2016). An intelligent fault diagnosis method using unsupervised feature learning towards mechanical big data. *IEEE Transactions on Industrial Electronics* 63 (5): 3137–3147.

Li, B., Chow, M.Y., Tipsuwan, Y., and Hung, J.C. (2000). Neural-network-based motor rolling bearing fault diagnosis. *IEEE Transactions on Industrial Electronics* 47 (5): 1060–1069.

Li, B., Goddu, G., and Chow, M.Y. (1998). Detection of common motor bearing faults using frequency-domain vibration signals and a neural network based approach. In: *American Control Conference, 1998. Proceedings of the 1998*, vol. 4, 2032–2036. IEEE.

Li, H., Zhang, Y., and Zheng, H. (2009). Gear fault detection and diagnosis under speed-up condition based on order cepstrum and radial basis function neural network. *Journal of Mechanical Science and Technology* 23 (10): 2780–2789.

Liang, B., Iwnicki, S.D., and Zhao, Y. (2013). Application of power spectrum, cepstrum, higher order spectrum and neural network analyses for induction motor fault diagnosis. *Mechanical Systems and Signal Processing* 39 (1–2): 342–360.

Liano, K. (1996). Robust error measure for supervised neural network learning with outliers. *IEEE Transactions on Neural Networks* 7 (1): 246–250.

Maglogiannis, I., Sarimveis, H., Kiranoudis, C.T. et al. (2008). Radial basis function neural networks classification for the recognition of idiopathic pulmonary fibrosis in microscopic images. *IEEE Transactions on Information Technology in Biomedicine* 12 (1): 42–54.

McCormick, A.C. and Nandi, A.K. (1996a). A comparison of artificial neural networks and other statistical methods for rotating machine condition classification. In: *Colloquium Digest-IEE*, 2-2. IEE Institution of Electrical Engineers.

McCormick, A.C. and Nandi, A.K. (1996b). Rotating machine condition classification using artificial neural networks. In: *Proceedings of COMADEM*, vol. 96, 85–94. Citeseer.

McCormick, A.C. and Nandi, A.K. (1997a). Classification of the rotating machine condition using artificial neural networks. *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science* 211 (6): 439–450.

McCormick, A.C. and Nandi, A.K. (1997b). Real-time classification of rotating shaft loading conditions using artificial neural networks. *IEEE Transactions on Neural Networks* 8 (3): 748–757.

McCulloch, W.S. and Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *The Bulletin of Mathematical Biophysics* 5 (4): 115–133.

Meireles, M.R., Almeida, P.E., and Simões, M.G. (2003). A comprehensive review for industrial applicability of artificial neural networks. *IEEE Transactions on Industrial Electronics* 50 (3): 585–601.

Nandi, A.K., Liu, C., and Wong, M.D. (2013). Intelligent vibration signal processing for condition monitoring. In: *Proceedings of the International Conference Surveillance*, vol. 7, 1–15.

Peck, J.P. and Burrows, J. (1994). On-line condition monitoring of rotating equipment using neural networks. *ISA Transactions* 33 (2): 159–164.

Rafiee, J., Arvani, F., Harifi, A., and Sadeghi, M.H. (2007). Intelligent condition monitoring of a gearbox using an artificial neural network. *Mechanical Systems and Signal Processing* 21 (4): 1746–1754.

Rosenblatt, F. (1958). The perceptron: a probabilistic model for information storage and organization in the brain. *Psychological Review* 65 (6): 386.

Samanta, B., Al-Balushi, K.R., and Al-Araimi, S.A. (2001). Use of genetic algorithm and artificial neural network for gear condition diagnostics. In: *Proceedings of COMADEM*, 449–456. Elsevier.

Samanta, B., Al-Balushi, K.R., and Al-Araimi, S.A. (2003). Artificial neural networks and support vector machines with genetic algorithm for bearing fault detection. *Engineering Applications of Artificial Intelligence* 16 (7–8): 657–665.

Sanz, J., Perera, R., and Huerta, C. (2007). Fault diagnosis of rotating machinery based on auto-associative neural networks and wavelet transforms. *Journal of Sound and Vibration* 302 (4–5): 981–999.

Saravanan, N., Siddabattuni, V.K., and Ramachandran, K.I. (2010). Fault diagnosis of spur bevel gear box using artificial neural network (ANN), and proximal support vector machine (PSVM). *Applied Soft Computing* 10 (1): 344–360.

Saxena, A. and Saad, A. (2007). Evolving an artificial neural network classifier for condition monitoring of rotating mechanical systems. *Applied Soft Computing* 7 (1): 441–454.

Shujaat, K. (2014). Radial basis function with k mean clustering. Mathworks File Exchange Center. https://uk.mathworks.com/matlabcentral/fileexchange/46220-radial-basis-function-with-k-mean-clustering?s_tid=FX_rc2_behav.

Shuting, W., Heming, L., and Yonggang, L. (2002). Adaptive radial basis function network and its application in turbine-generator vibration fault diagnosis. In: *International Conference on Power System Technology, 2002. Proceedings. PowerCon 2002*, vol. 3, 1607–1610. IEEE.

Sreejith, B., Verma, A.K., and Srividya, A. (2008). Fault diagnosis of rolling element bearing using time-domain features and neural networks. In: *2008 IEEE Region 10 and the Third International Conference on Industrial and Information Systems*, 1–6. IEEE.

Subrahmanyam, M. and Sujatha, C. (1997). Using neural networks for the diagnosis of localized defects in ball bearings. *Tribology International* 30 (10): 739–752.

Tyagi, C.S. (2008). A comparative study of SVM classifiers and artificial neural networks application for rolling element bearing fault diagnosis using wavelet transform preprocessing. *Neuron* 1: 309–317.

Unal, M., Onat, M., Demetgul, M., and Kucuk, H. (2014). Fault diagnosis of rolling bearings using a genetic algorithm optimized neural network. *Measurement* 58: 187–196.

Vyas, N.S. and Satishkumar, D. (2001). Artificial neural network design for fault identification in a rotor-bearing system. *Mechanism and Machine Theory* 36 (2): 157–175.

Wang, C.C. and Too, G.P.J. (2002). Rotating machine fault detection based on HOS and artificial neural networks. *Journal of Intelligent Manufacturing* 13 (4): 283–293.

Widrow, B. and Hoff, Macian E., 1960. *Adaptive switching circuits*, pp. 96–104.

Wu, S. and Chow, T.W. (2004). Induction machine fault detection using SOM-based RBF neural networks. *IEEE Transactions on Industrial Electronics* 51 (1): 183–194.

Yang, B.S., Han, T., and An, J.L. (2004). ART–KOHONEN neural network for fault diagnosis of rotating machinery. *Mechanical Systems and Signal Processing* 18 (3): 645–657.

Yang, D.M., Stronach, A.F., MacConnell, P., and Penman, J. (2002b). Third-order spectral techniques for the diagnosis of motor bearing condition using artificial neural networks. *Mechanical Systems and Signal Processing* 16 (2–3): 391–411.

Yang, H., Mathew, J. and Ma, L., 2002a. Intelligent diagnosis of rotating machinery faults-a review.

Yang, S., Li, W., and Wang, C. (2008). The intelligent fault diagnosis of wind turbine gearbox based on artificial neural network. In: *International Conference on Condition Monitoring and Diagnosis, 2008. CMD 2008*, 1327–1330. IEEE.

Zhang, Z., Wang, Y., and Wang, K. (2013). Fault diagnosis and prognosis using wavelet packet decomposition, Fourier transform and artificial neural network. *Journal of Intelligent Manufacturing* 24 (6): 1213–1227.