

Condition-based monitoring system for rolling element bearing using a generic multi-layer perceptron

Journal of Vibration and Control
2015, Vol. 21(16) 3456–3464
© The Author(s) 2014
Reprints and permissions:
sagepub.co.uk/journalsPermissions.nav
DOI: 10.1177/1077546314524260
jvc.sagepub.com



Luis F de Almeida¹, José WP Bizarria¹, Francisco CP Bizarria²
and Mauro H Mathias³

Abstract

Rolling element bearings are critical mechanical components in rotating machinery and fault detection in the early stages of damage is important to prevent their malfunctioning and failure. Vibration monitoring is the most widely used and cost-effective monitoring technique to detect, locate and distinguish faults in rolling element bearings. This paper proposes single hidden layer architecture for fault diagnosis of rolling element bearings. The particular of this proposed architecture is its ability to generalize for solving both basic classification and fault identification. The network uses the features of time-domain vibration signals with normal and defective bearings. The Multi Layer Perceptron (MLP) was trained and tested with a set of experimental data obtained from previous experiments developed by FEG, CWRU and RANDALL laboratories. The results show the effectiveness of the MLP to diagnose the machine condition for the various data used.

Keywords

Artificial Neural Network, Multi Layer Perceptron, Condition-Based Monitoring, vibration monitoring

1. Introduction

One real problem for the companies that apply industrial machinery is related to stopping the production line because of machine failures or scheduled preventive maintenance. Rolling element bearings are main components of most rotating machinery and the majority of problems in these machines are caused by faulty bearings (Niu et al., 2005, Farshidianfar et al., 2008, Hameed et al., 2009, Randal, 2011, Dong and Chen, 2011). The most common faults in a rolling bearing are: inner race, outer race, cage and rolling-element bearing. Figure 1 illustrate these components.

The bearing defects like spall, pitting, and others, generate impact vibrations over the structure. When a single point defect starts, the strike of the running ball over this fault generates periodic pulses that excite resonances of the structure. The characteristic of such pulses are related with the geometry of the bearing, the kind of fault, the rotation of the shaft and the type of bearing load. The measurements of these signals on the bearing support house is the key point to evaluate the condition of the bearing. The detection of such

signals can be performed using analysis techniques from time domain, frequency domain and time-frequency domain.

An appropriate maintenance policy is a key factor to improve the productivity rate. Usually, the use of mechanisms for detecting or preventing failures in complex machines provides benefits in terms of economy and security (Tavner and Penman, 1987). Condition-Based Maintenance (CBM) techniques have been used as an effective policy for monitoring machines. Vibration monitoring is one of the most important tool applied to CBM (Mobley, 2008, Scheffer and Girdhar, 2004, Kobbacy and Murthy, 2008).

¹Department of Informatics, University of Taubate, Brazil

²Department of Electrical Engineering, University of Taubate, Brazil

³Faculty of Engineering, Sao Paulo State University, Brazil

Received: 22 April 2013; accepted: 23 January 2014

Corresponding author:

Luis F de Almeida, Department of Informatics, University of Taubate, Brazil, 605 Marechal Deodoro Avenue, Taubate, Sao Paulo, 12010000, Brazil.

Email: luis.almeida@unitau.br

Information Technology (IT) has become a fundamental tool for the maintenance. Higgins and Mobley (2002) pointed out four types of benefits for using IT: cost-cutting, easier access to information, better planning and increased control of maintenance operations.

One important IT resource is the software, which allow us to construct useful tools, when associated to advanced techniques, such as Artificial Intelligence (AI). These techniques have shown superior performance compared to conventional approaches (Jardine et al., 2006).

Among AI techniques, Artificial Neural Networks (ANN) stand out as the most widely used, particularly in vibration monitoring. The proposals have fallen into two approaches: detection of presence or absence of failure (Samanta and Al-Balushi, 2003, Samanta, 2004, Rafieea et al., 2007) and identification of the fault type (Jack and Nandi, 2000, Li et al., 2000, Al-Raheem and Abdul-Karem, 2010).

However, each ANN implemented in these works was strictly applied to solve a specific approach. Thus, this paper presents a generic Multi Layer Perceptron (MLP) which, if it is properly parameterized, can help to solve problems for both fault detection and identification of failure.

2. Theoretical foundations

2.1. Artificial Neural Networks (AAN)

An ANN consists in a large number of simple processing elements called neurons, units, nodes or cells, each one connected to other neuron by means of directed communication links (Fausett, 1993). There are various architectures of ANN, and MLP stands out as one of the most widely used.

An ANN has the advantages of adaptive learning, nonlinear generalization, fault tolerance, resistance to noisy data, and parallel computation abilities (Haykin, 1999).

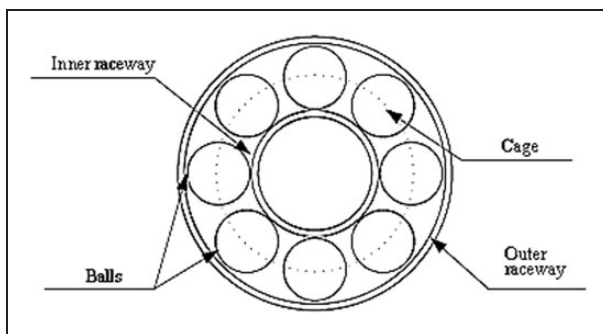


Figure 1. Main element of a rolling bearing.

Its generic architecture (Figure 2) is a directed graph where the neurons are represented by nodes and the weights by edges.

The main component of an ANN is the artificial neuron illustrated in Figure 3. We can identify three basic elements:

1. A set of synapses or weights w_i that determine the effect of one neuronal signal on another.
2. An activation function f that determines the level of neuron activation. The most common activation functions are threshold and sigmoid.
3. An adder Σ for summing the input signals, weighted by respective synapses of the neuron.

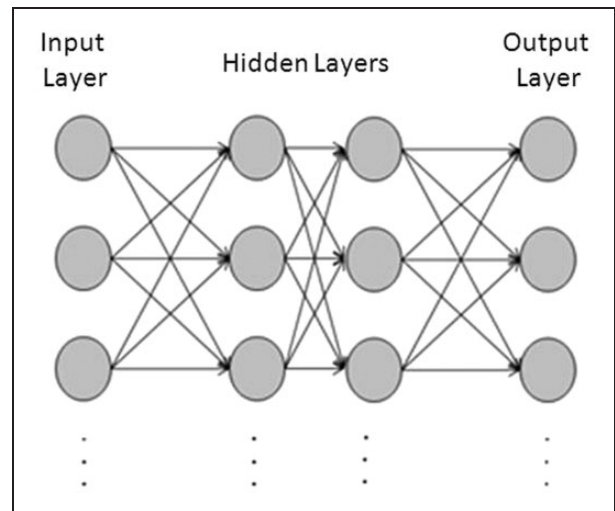


Figure 2. Generic ANN architecture.

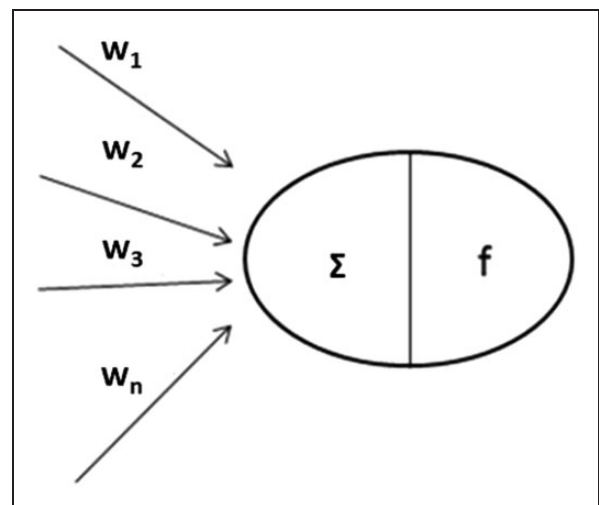


Figure 3. Artificial neuron.

2.2. Time-domain analysis

Analysis techniques applied for processing the vibration signals for condition monitoring can be classified as time-domain, frequency domain and time-frequency domain.

The most commonly used technique in the frequency domain analysis is the spectrum analysis using Fast Fourier Transform (FFT) (McFadden and Smith, 1984, Mechefske and Mathew, 1992, Chen et al., 2005). Other techniques, such as envelope analysis, is also being widely used (Yu et al., 2007, Guo et al., 2009, Ming et al., 2011). Short-time Fourier transform (STFT) and Wigner–Ville distribution are the most popular time–frequency distributions applied to CBM (Jardine et al., 2006, Li et al., 2006, Li and Mechefske, 2006, Lei et al., 2011).

Time-domain analysis is directly based on the time waveform itself and traditional time-domain analysis calculates characteristics called time-domain features. Some authors (Scheffer and Girdhar, 2004) emphasize that a good time-domain analysis depends on the setting of the following parameters: unit of measurement (displacement, velocity or acceleration), time period sample, resolution, averaging and windows.

The most straightforward technique for time domain analysis is the visual inspection of segments of the waveform. Nevertheless, McFadden (1990) highlights its limitation in identifying certain defects in bearings. Some researches have pointed out the statistical features as important parameters for prognostics and diagnostics of failures (Mendel, 1991, Nandi, 1999, Zhang et al., 2005).

Considering a set of samples X of length N , the analyst can calculate the following statistical features: mean, variance, kurtosis, skewness, high order moments, rms, peak value, peak-to-peak value, crest factor, impulse factor, shape factor, clearance factor. In equations (1) to (12) the expression of each feature is shown

$$\mu = \frac{\sum_{i=1}^N x_i}{N} \quad (1)$$

$$\sigma^2 = \frac{\sum_{i=1}^N (x_i - \mu)^2}{N} \quad (2)$$

$$\sigma^3 = \frac{\sum_{i=1}^N (x_i - \mu)^3}{\sigma^3} \quad (3)$$

$$\sigma^4 = \frac{\sum_{i=1}^N (x_i - \mu)^4}{\sigma^4} \quad (4)$$

$$m^k = \frac{\sum_{i=1}^N (x_i - \mu)^k}{\sigma^k} \quad (5)$$

$$rms = \sqrt{\frac{\sum_{i=1}^N (x_i - \mu)^2}{N}} \quad (6)$$

$$pv = \max(x_1, \dots, x_N) \quad (7)$$

$$ppv = \max(x_1, \dots, x_N) - \min(x_1, \dots, x_N) \quad (8)$$

$$crf = \frac{pv}{rms} \quad (9)$$

$$imf = \frac{pv}{\frac{1}{N} \sum_{i=1}^N |x_i - \mu|} \quad (10)$$

$$shf = \frac{rms}{\text{fim} / N} \quad (11)$$

$$clf = \frac{pv}{\frac{1}{N} \left(\sqrt{\sum_{i=1}^N |x_i - \mu|} \right)^2} \quad (12)$$

where x_i is the sampled signal, μ is the signal average, σ^2 is the variance and N is the number of data points.

3. Procedure for development

3.1. Experimental data

The vibration data used in this paper have been taken from three sources: DATA-FEG, DATA-RANDALL and DATA-CWRU.

DATA-FEG. These data were obtained from experiments at Mechanical Engineering Department laboratory of the Sao Paulo State University (UNESP), Campus from Guaratinguetá. The experimental test rig (Figure 4) consists of a three-phase induction motor (WEG 1/2 HP) driving a shaft rotor assembly with a rolling bearing mounted in one end. To collect the data two accelerometers (B&K4371) were attached on the horizontal and vertical positions of the bearing support house, a charge amplifier (ENDEVCO 133) was used for signal conditioning.

The ball bearing tested was the NSK6205, whose specifications are shown in Table 1. The signals were sampled at 12,000 samples/s lasting 10 seconds, and recorded using a digital data recorder system (AQDADOS).

The database was divided in three different datasets, according to motor speed: 1200, 2100 and 3000 rpm (Table 2). Each dataset contains signals which are related to five conditions: normal (NO), inner race with slight and severe defects (IRSL and IRSE), and outer race with slight and severe defects (ORSL and ORSE).

DATA-CWRU. These data are based on the experiments developed by Dr Loparo's laboratory at Case



Figure 4. DATA-FEG test rig.

Table 1. NSK6205 and 6205 SKF specification.

Inside diam.	Outside diam.	Thickness	Ball diam.	Pitch diam.
0.9843	2.0472	0.5906	0.3126	1.537

Table 2. DATA-FEG datasets.

Dataset identification	Motor rotation (rpm)
DF1	1200
DF2	2100
DF3	3000

Western Reserve University. The test setup (Figure 5) includes a three-phase induction motor (Reliance Electric 2HP IQPreAlert), a torque transducer/encoder and a dynamometer. SKF and NTN equivalent bearings were used for the experiments. The bearing specifications are depicted in Tables 1 and 3.

Vibration data was collected using accelerometers placed at the 12 o'clock position at both the drive end and fan end of the motor housing. These data were sampled at 12,000 and 48,000 samples/s with the shaft

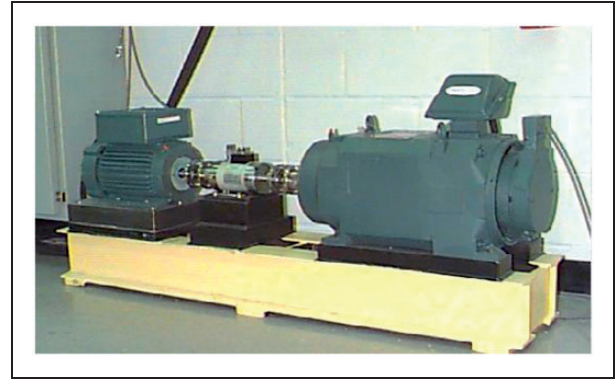


Figure 5. DATA-CWRU test stand (Case Western Reserve University, 2011).

Table 3. 6203 SKF specification.

Inside diam.	Outside diam.	Thickness	Ball diam.	Pitch diam.
0.6693	1.5748	0.4724	0.2656	1.122

Table 4. DATA-CWRU datasets.

Dataset identification	Motor velocity (rpm)
DC1	1797
DC2	1772
DC3	1750
DC4	1730

operating at different speeds (1797, 1772, 1750 and 1730 rpm).

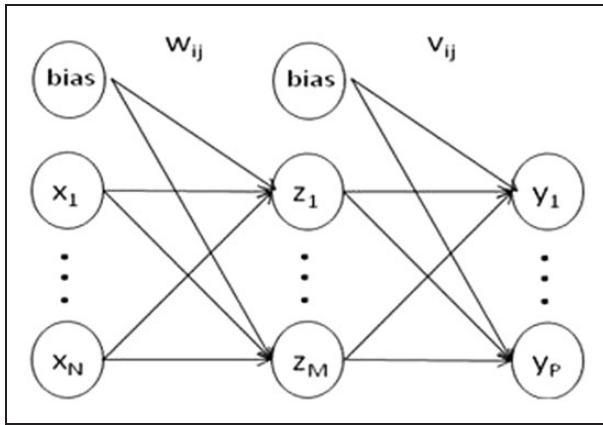
Single point faults were introduced to the test bearings using electro-discharge machining with fault diameters of 0.007 inches, 0.014 inches, 0.021 inches, 0.028 inches and 0.040 inches.

Table 4 presents the datasets generated from CWRU's experiments according to shaft rotational speed. Each dataset was generated taking into account seven different conditions: normal, inner race with slight and severe faults, outer race with slight and severe faults, and rolling element with slight and severe faults.

DATA-RANDALL. This database is proposed by Randall (2011) and contains signals taken from four bearing conditions: normal, inner race fault, outer race fault and rolling element fault. The sampling frequency is 48,000 samples/s and the shaft speed in the test rig was 360 rpm. The bearing parameters are depicted in Table 5 and this database is formed by only one dataset named DR1.

Table 5. Bearing specification of the Randall test.

Ball diam.	Pitch circle diam.	Rolling elements	Contact angle
7.12	38.5	12	0

**Figure 6.** Architecture MLP.

From this database was generated only one dataset containing information about four conditions: normal, inner race fault, outer race fault and rolling element fault.

3.2. MLP implemented

In this paper, we propose a generic ANN architecture able to diagnose the presence or not of a failure, if so, identify the type of fault.

This architecture is based on a MLP, as illustrated in Figure 6, with one hidden layer. It is characterized by the following aspects:

1. The number of N neurons in input layer is specified by amount of time domain features used for classification.
2. The number of P neurons in output layer is specified by amount of faults in dataset.
3. After preliminary investigation the hidden layer was set to a number of 5 neurons since increasing this value did not afford improved results.
4. We use the backpropagation algorithm for training the network.
5. It is used bipolar sigmoid activation function (equation 13) to calculate the output neurons

$$f(x) = \frac{2}{(1 + e^{-x})} - 1 \quad (13)$$

6. Initial weights w_{ij} and v_{ij} are created through the optimization process proposed by Nguyen and Widrow (1990).

7. The input data are normalized according to equation 14. The \min_i and \max_i variables represent the range of "i" normalized features, while d_{\min} and d_{\max} are the lower and upper values for that feature, respectively

$$nd_{i,j} = \frac{\min_i + (\max_i - \min_i) \times (d_{i,j} - d_{\min})}{(d_{\max} - d_{\min})} \quad (14)$$

Therefore, the configurable elements are directly dependent of two points: the problem, detection or identification of the fault, and the solution mode, number of features. The former defines the output layer size while the latter defines the input layer size.

To detect the fault presence, the output layer is set to a single binary neuron with value equals to 1 for fault and 0 for no fault. On the other hand, the number of output neurons is set to amount of different faults that you want to identify.

The input layer configuration is straightforward, i.e., the number of neurons is the same as the amount of the time domain features used for classification.

In order to provide faster learning we use a weight initialization, developed by Nguyen and Widrow (1990). The procedure is divided in the following steps:

1. For the weights w_{ij} between input layer and hidden layer:
 - a) Initialize your values between -0.5 e 0.5 .
 - b) Calculate

$$w_j = \sqrt{w_{1j}^2 + w_{2j}^2 + \dots + w_{Nj}^2} \quad (15)$$

$$\beta = 0.7 \times M^{1/N} \quad (16)$$

- c) Recalculate weights values

$$w_{ij} = \beta \times \frac{w_{ij}}{w_j} \quad (17)$$

- d) Set the bias weights values between β and $-\beta$.
- 2) For the weights between hidden layer and output layer:
 - a) Set your values between -0.5 and 0.5 .

3.3. NeuralNet-CBM architecture

The MLP is implemented in a software named NeuralNet-CBM (Figure 7). Its purpose is to make available a tool that enables the data analysis from different systems of data acquisition. It is divided in two modules: vibration and particles. The first is discussed in this paper.

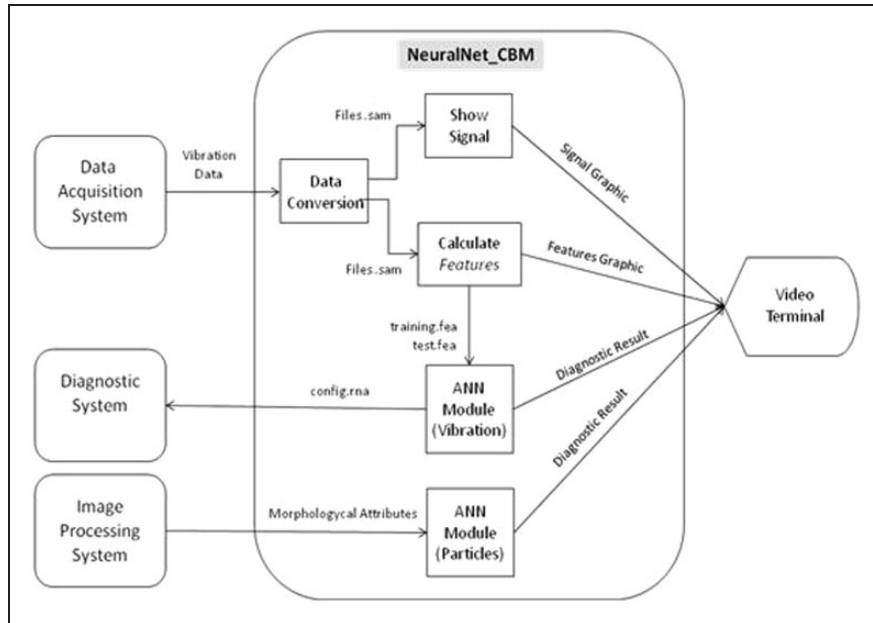


Figure 7. NeuralNet-CBM architecture.

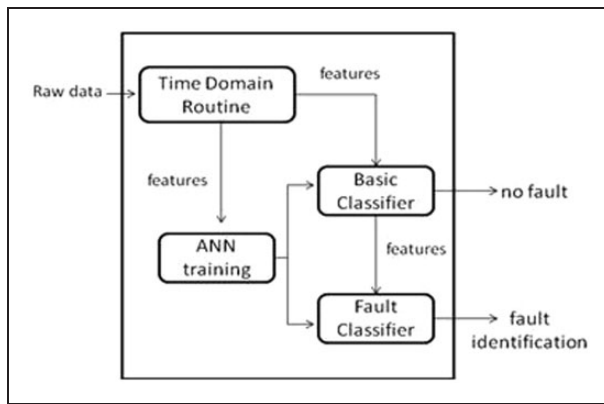


Figure 8. Vibration module.

The CBM process in the vibration module consists in two steps: basic classifier and fault classifier. Its functionality (Figure 8) is in the following way:

1. The time domain routine is responsible for calculating the time domain features.
2. From features the neural classifiers are generated.
3. Having the features, you run the basic classifier for diagnosis of the fault presence. If this happens, you run the fault classifier for fault identification.

The ANN training was performed as follow:

1. Initialize the weights using Nguyen and Widrow (1990) method.

2. While stop condition is false, do steps (3) to (7).
3. Select a pair of training $\{X, class\}$.
4. Calculate the input and output of each hidden layer node

$$z_{in_j} = v_{0j} + \sum_{i=1}^N x_i \times v_{ij} \quad (18)$$

$$z_j = f(z_{in_j}) \quad (19)$$

5. Calculate the input and output of each output layer node

$$y_{in_k} = w_{0k} + \sum_{j=1}^M z_j \times w_{jk} \quad (20)$$

$$y_j = f(y_{in_j}) \quad (21)$$

6. Calculate the output of the ANN.
7. Backpropagation of error:
 - a) Calculate error of the output layer node

$$\delta_{y_k} = (d_k - y_k) \times f'(y_{in_k}) \quad (22)$$

- b) Calculate error of the hidden layer node

$$\delta z_{in_j} = \sum_{k=1}^p w_{jk} \times \delta y_k \quad (23)$$

$$\delta z_j = \delta z_{in_j} \times f'(z_{in_j}) \quad (24)$$

c) Set the learning rate

$$\alpha = \frac{0.4}{\left(1 + \frac{\text{iteration}}{\text{number_of_epochs}}\right)} \quad (25)$$

d) Update the weights

$$w_{jk}^{t+1} = w_{jk}^t + \alpha \delta y_k z_j + \gamma (w_{jk}^t - w_{jk}^{t-1}) \quad (26)$$

$$v_{ij}^{t+1} = v_{ij}^t + \alpha \delta z_j x_i + \gamma (v_{ij}^t - v_{ij}^{t-1}) \quad (27)$$

4. Tests and results

A key aspect about ANN is the performance, i.e., computing time and diagnostic reliability. The first depends on ANN architecture, that is, number of neurons in each layer and training time. On the other hand, the second is related to the ability of ANN in making the classification with the lowest rate possible error. This depends on many factors, such as, number of neurons in input and hidden layers, proper configuration to be used, training time.

The number of the output layer neurons is linked to the faults. Thereby, the basic classifier has one neuron and the fault classifier has four, six and three neurons for DATA-FEG1, DATA-CWRU and DATA-RANDALL, respectively. The number of the hidden layer neurons was set up to five for both the basic and fault classifiers.

For each configuration is performed 200 runs and calculated the average which indicates the classification success rate. The ANN was trained by 3000 and 1000 epochs for the basic and fault classifier, respectively. The following sections present the results for each data set. Table 6 shows a summary of the data used in the ANN.

4.1. DATA-FEG results

For each condition, their samples were divided into 50 segments (S_1, \dots, S_{50}), each with 2048 points. These segments were grouped into two sets, training and test, resulting in 125 segments each. The next step was calculating the time domain features (f_i) for each set. Figure 9 shows the process of creating files of training and testing. This process is similar for all databases.

The ANN's output for the basic classifier is -1 (no fault) or $+1$ (fault) and this is the same for all databases. For the fault classifier is either 0001 or 0010 or 0100 or 1000. The binary value 0 or 1 is associated with activation or not of a particular output neuron, each representing one of four failures.

For each dataset in Table 2, the number of features was set at one and then increased until to reach a stability. This procedure was adopted for three databases. Table 7 presents the final results of all tests for DATA-FEG database.

4.2. DATA-CWRU results

The samples of each condition were divided into 200 segments (S_1, \dots, S_{200}), each with 2048 points. These segments were grouped into two sets, training and test, resulting in 700 segments each. The next step was to calculate the time domain features (f_i) for each set.

The ANN's output for the fault classifier is either 000001 or 000010 or 000100 or 001000 or 010000 or 100000. Table 8 presents the final results of all tests.

Table 6. Summary of the data size.

Dataset	Training data	Test data
DATA-FEG	125	125
DATA-CWRU	700	700
DATA-RANDALL	180	180

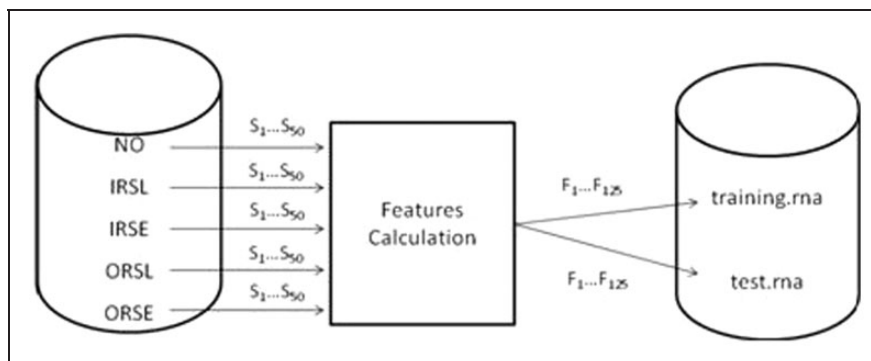


Figure 9. The process of creating files for training and test from DATA-FEG1.

Table 7. Results for DATA-FEG.

Dataset	Basic classifier		Fault classifier	
	Features	%	Features	%
DF1	1	93.0	1	91.3
	2	98.6	2	97.7
	3	99.9	3	99.3
DF2	1	89.1	1	95.7
	2	90.6	2	97.7
	3	92.1	3	98.4
	4	93.7	4	99.4
	5	94.4	–	–
	6	95.8	–	–
DF3	1	100.0	1	88.7
	–	–	2	93.2
	–	–	3	96.3
	–	–	4	98.4

Table 8. Results for DATA-CWRU.

Dataset	Basic classifier		Fault classifier	
	Features	%	Features	%
DC1	1	100.0	1	82.4
	–	–	2	93.9
	–	–	3	95.9
DC2	1	100.0	1	74.2
	–	–	2	93.7
	–	–	3	95.7
DC3	1	100.0	1	86.7
	–	–	2	95.8
	–	–	3	99.6
DC4	1	100.0	1	65.9
	–	–	2	94.4
	–	–	3	99.4

4.3. DATA-RANDALL results

As in previous tests, the ANN output for the basic classifier is -1 and 1 . The fault classifier output is 001 , 010 , 100 , depending on the detected failure.

The samples of each rolling bearing state were divided in 90 segments, $S_1 \dots S_{90}$, each with 1024 points. These segments were grouped into two sets: training and test, resulting in 180 segments in each set. Table 9 presents the final results of all tests.

All results presented in Tables 7–9 show a good performance (greater than 95%) of the ANN using a small amount of features in input layer.

Table 9. Results for DATA-RANDALL.

Dataset	Basic classifier		Fault classifier	
	Features	%	Features	%
DR1	1	94.1	1	88.9
	2	96.9	2	90.8
	3	98.9	3	96.2

5. Conclusion

This paper proposed a generic MLP for solving problems for both fault detection and identification of failure. The results demonstrate its effectiveness for the following aspects: ability to handle signals from different sources; right classification above 95% even using a few nodes (features) in the input layer; and possibility to solve problems for both fault detection and identification of failure.

It is difficult to highlight a single statistical parameter as a determining feature for efficient classification of all databases analyzed. However, the analysis have shown the features with the largest presence in the classifiers generated by the ANN are peak-to-peak value, rms, mean, and kurtosis.

The method proposed has a great potential to be applied in the industrial environment in a way to help maintenance team.

The next challenge in our research would be considering other faults, like cage failure and combined faults into a single sample and also to evaluate the weight of each statistical feature into the analysis.

The result of this approach is deterministic, i.e., given an input signal it indicates the bearing state and, if applicable, the defect. Instead the results indicate the type of defect, it would present values of expected probability for each fault, and the highest value would indicate the most likely fault. In this case the appropriate approach could be the Bayesian networks.

Funding

This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

References

- Al-Raheem KF and Abdul-Karem W (2010) Rolling bearing fault diagnostics using artificial neural networks based on Laplace wavelet analysis. *International Journal of Engineering, Science and Technology* 2: 278–290.
- Case West Reserve University (2011) Bearing Data Center: Apparatus & Procedures. Available at: <http://csegroups>.

- case.edu/bearingdatacenter/pages/apparatus-procedures (accessed 17 September 2010).
- Chen P, Taniguchi M, Toyota T and He Z (2005) Fault diagnosis method for machinery in unsteady operating condition by instantaneous power spectrum and genetic programming. *Mechanical Systems and Signal Processing* 19: 175–194.
- Dong GM and Chen J (2011) Study on cyclic energy indicator for degradation assessment of rolling element bearings. *Journal of Vibration and Control* 17: 1805–1816.
- Farshidianfar A, Abbasi AR, Abbasion S and Moeenfarid H (2008) Dynamic modeling of chaotic response of bearing systems due to surface defects. In: *16th. Annual (International) Conference on Mechanical Engineering-ISME2008*. Shahid Bahonar University, Kerman, Iran, 13 May 2008.
- Fausett L (1993) *Fundamentals of Neural Networks: Architectures, Algorithms and Applications*. New Jersey: Prentice-Hall.
- Guo L, Chen J and Li X (2009) Rolling bearing fault classification based on envelope spectrum and support vector machine. *Journal of Vibration and Control* 15: 1349–1363.
- Hameed Z, Hong YS, Cho YM, Ahn SH and Song CK (2009) Condition monitoring and fault detection of wind turbines and related algorithms: a review. *Renewable and Sustainable Energy Reviews* 13: 1–39.
- Haykin S (1999) *Neural Networks: a Comprehensive Foundation*, 2nd ed. Upper Saddle River: Prentice Hall.
- Higgins LR and Mobley RK (2002) *Maintenance Engineering Handbook*, 6th ed. New York: McGraw-Hill.
- Jack LB and Nandi AK (2000) Genetic algorithms for feature selection in machine condition monitoring with vibration signals. *IEE Proceedings in Vision, Image and Signal Processing* 147: 205–212.
- Jardine AKS, Lin D and Banjevic D (2006) A review on machinery diagnostics and prognostics implementing condition-based maintenance. *Mechanical Systems and Signal Processing* 20: 1483–1500.
- Kobbacy KA and Murthy DNP (2008) *Complex System: Maintenance Handbook*. London: Springer Verlag.
- Lei Y, Lin J, He Z and Zi Y (2011) Application of an improved kurtogram method for fault diagnosis of rolling element bearings. *Mechanical Systems and Signal Processing* 25: 1738–1749.
- Li B, Chow MY, Tipsuwan Y and Hung JC (2000) Neural network-based motor rolling bearing fault diagnosis. *IEEE Transactions on Industrial Electronics* 47: 1060–1069.
- Li H, Zheng H and Tang L (2006) Wigner-Ville distribution based on EMD for faults diagnosis of bearing. *Lecture Notes in Computer Science* 4223: 803–812.
- Li W and Mechefske CK (2006) Detection of induction motor faults: a comparison of stator current, vibration and acoustic methods. *Journal of Vibration and Control* 12: 165–188.
- Mcfadden PD (1990) Condition monitoring of rolling element bearing by vibration analysis. IMechE Paper, Solid Mechanics and Machine Systems Group Seminar, London, UK, 9 January 1990. pp. 49–54.
- Mcfadden PD and Smith LD (1984) Model for the vibration produced by a single point in a rolling element bearing. *Journal of Sound and Vibration* 96: 69–82.
- Mechefske CK and Mathew J (1992) Fault detection and diagnosis in low speed rolling element bearing. Part II: The use of nearest neighbour classification. *Mechanical Systems and Signal Processing* 6: 309–316.
- Mendel JM (1991) Tutorial on higher-order statistics (Spectra) in signal processing and system theory: theoretical results and some applications. *Proceedings of the IEEE* 79: 278–305.
- Ming Y, Chena J and Dong G (2011) Weak fault feature extraction of rolling bearing based on cyclic Wiener filter and envelope spectrum. *Mechanical Systems and Signal Processing* 25: 1773–1785.
- Mobley KM (2008) Predictive Maintenance. In: Mobley KM, Higgins LR and Wikoff DJ (eds) *Maintenance Engineering Handbook*. New York: McGraw-Hill Professional, pp. 2.19–2.34.
- Nandi AK (1999) *Blind Estimation Using High-orders Statistics*. Boston: Kluwer Academic Publishers.
- Nguyen D and Widrow B (1990) Improving the learning speed of 2-layer neural networks by choosing initial values of the adaptive weights. *Proceedings of the International Joint Conference on Neural Networks* 3: 21–26.
- Niu X, Zhu L and Ding H (2005) New statistical moments for the detection of defects in rolling element bearings. *International Journal of Advanced Manufacturing Technology* 26: 1268–1274.
- Rafieea J, Arvania F, Harifib A and Sadeghi MH (2007) Intelligent condition monitoring of a gearbox using artificial neural network. *Mechanical Systems and Signal Processing* 21: 1746–1754.
- Randall RB (2011) *Vibration-Based Condition Monitoring*. West Sussex: John Wiley & Sons, Ltd.
- Samanta B (2004) Artificial neural networks and genetic algorithms for gear fault detection. *Mechanical Systems and Signal Processing* 18: 1273–1282.
- Samanta B and Al-Balushi KR (2003) Artificial neural network based fault diagnostics of rolling element bearings using time-domain features. *Mechanical Systems and Signal Processing* 17: 317–328.
- Scheffer C and Girdhar P (2004) *Practical Machinery Vibration Analysis and Predictive Maintenance*. Boston: Elsevier.
- Tavner PJ and Penman J (1987) *Condition Monitoring of Electrical Machines*. New York: Wiley.
- Yu B, Liu J and Wang C (2007) Rolling bearing fault diagnosis using refinement envelope analysis based on the EMD method. *Advanced Intelligent Computing Theories and Applications* 2: 562–570.
- Zhang L, Jack LB and Nandi AK (2005) Fault detection using genetic programming. *Mechanical Systems and Signal Processing* 19: 271–289.