



Blind Application of Developed Smart Vibration-Based Machine Learning (SVML) Model for Machine Faults Diagnosis to Different Machine Conditions

Natalia F. Espinoza Sepúlveda¹ · Jyoti K. Sinha¹

Received: 11 February 2020 / Revised: 21 August 2020 / Accepted: 16 September 2020 / Published online: 7 October 2020
© The Author(s) 2020

Abstract

Purpose The development and application of intelligent models to perform vibration-based condition monitoring in industry seems to be receiving attention in recent years. A number of such research studies using the artificial intelligence, machine learning, pattern recognition, etc., are available in the literature on this topic. These studies essentially used the machine vibration responses with known machine faults to develop smart fault diagnosis models. These models are yet to be tested for all kinds of machine faults and/or different operating conditions. Therefore, the purpose is to develop a generic machine faults diagnosis model that can be applied blindly to any identical machines with high confidence level in accuracy of the predictions.

Methods In this paper, a supervised smart fault diagnosis model is developed. This model is developed using the available measured vibration responses for the different rotor faults simulated on an experimental rotating rig operating at a constant speed. The developed smart vibration-based machine learning (SVML) model is then blindly tested to identify the healthy and faulty conditions of the rig when operating at different speeds.

Results and conclusions Several scenarios are proposed and examined during the development of the SVML model. It is observed that scenario of the vibration measurements simultaneously from all bearings from a machine is capable to fully map the machine dynamics in the VML model. Therefore, this developed when applied blindly to the sets of data at a different machine speed, the results are observed to be encouraging. The results clearly show a possibility for a centralised vibration-based condition monitoring (CVCM) model for identical machines operating at different rotating speeds.

Keywords Machine fault diagnosis · Vibration analysis · Machine learning · Artificial neural network · Pattern recognition

Introduction

Vibration-based condition monitoring (VCM) in rotating machines has been successfully applied in industry for fault detection and diagnosis. However, the current and future approaches in the VCM are the research topic due to rapid changes in the technologies and instrumentation, including techniques for data processing and analysis.

Knowledge-based approaches, such as machine learning (ML) models stand out among the developed methods, since their lack of dependency on the expertise or knowledge level of a person to generate the correlations between the identified symptoms and their associated faults. They are characterised by their capability of exploring and learning from empirical data [1]. This ability is a powerful tool for fault identification, for instance through the performance of pattern recognition. These methods can also be classified according their learning process, which could be either unsupervised or supervised.

A wide range of techniques with a supervised learning have been used for pattern recognition. For instance, the application of support vector machine (SVM) in condition monitoring and fault identification has been extended to several machines and their components: induction motors [2], pumps [3], gearboxes [4], and others. However, most

✉ Jyoti K. Sinha
jyoti.sinha@manchester.ac.uk

Natalia F. Espinoza Sepúlveda
natalia.espinozasepulveda@postgrad.manchester.ac.uk

¹ Dynamics Laboratory, Department of Mechanical, Aerospace and Civil Engineering (MACE), School of Engineering, The University of Manchester, Manchester M13 9PL, UK

of the studies are concentrated in rolling element bearings fault diagnosis [5–7].

Hybrid methods for fuzzy logic have been developed to overcome the shortcomings that the individual techniques could present. An improved learning ability is achieved by neuro-fuzzy classifiers. This technique is applied for fault detection in several components such as bearings [8] and gearboxes [9].

The application of principal component analysis (PCA) at multiple steady speed rotors has allowed the acquirement of clearly spaced groups of data points, representing by them different defects [10], as well the differentiation from healthy condition points against data with introduced failures. Extension of the classical PCA have been developed to improve the limitation of the method regarding the linearity, such as kernel principal component analysis (KPCA) [11] and evolving kernel principal component (EKPCA) [12]. The drawback associated with the non-linear versions is the high usage of computational resources for being processed.

Regarding their capability to deal with complex data processing problems, artificial neural networks (ANN) have been positively implemented in mechanical systems. Nahvi and Esfahanian [13] used three-layer feedforward networks to detect faults in rotating machines through vibration data, considering more than 40 different faults. Their study evidenced the capability of this method of dealing with a high number of features and provide acceptable results in the fault detection. However, the 100% of accuracy in the diagnosis is not achieved at any of the conducted experiments.

The study made by Vyas and Satishkumar [14] in rotor-bearing system considers the use of moments of time series as the input to identify the faults, obtaining, under laboratory conditions, over a 90.0% of success in their method. Ben et al. [15] also studied bearing fault, using in this case a combination of features extracted in time domain with features extracted in time–frequency domain to create the inputs for the ANN, demonstrating experimentally the ability of the method for damage detection at an early stage.

Walker et al. [16] used the sub-synchronous nonlinear features of the vibration signal in frequency domain in addition of ANN to localise the existent unbalance in a rotating machine. Mohammed et al. [17] conducted the crack in shaft identification and damage quantification, considering depths from 0 to 60% of the shaft diameter, obtaining 100% of accuracy at some of the measuring locations using a four-layer feedforward

perceptron in the classification and peak position component analysis for the features extraction.

An extensive number of ML techniques have been applied to assess the fault detection and diagnosis in rotors, through a vibration-based approach. Learning assessment, introduced by pattern recognition models, has simplified the fault identification stage but generally the earlier studies have either not included different faults simultaneously in their ML model or blindly tested the developed ML at different machine conditions or both. This is required if the ML model to be used in the industries. Therefore, the current objective of the paper is to develop a smart vibration-based machine learning (SVML) model for a machine at an operating condition for the fault diagnosis, and then blindly applied and tested the adaptability of the model when machine is operating at different conditions. The paper presents the model development, results and its blind application to test robustness and reliability of the model.

Experimental Rig and Experimental Data

The used data in this study are pre-existent data acquired in an experimental rig in laboratory conditions. The rig has been used to conduct previous researches at the University of Manchester [18–20]. The rig, in Fig. 1, consists in two shafts connected by a rigid coupling (C2). The driven shaft (Sh1) has a length of 1.0 m and it is coupled by a flexible unit (C1) to a three-phase electric motor (0.75 kW). This shaft, has installed two balancing discs (D1, D2), while the second shaft (Sh2) of a length of 0.5 m has one balancing disc (D3). The assembly is mounted over a total

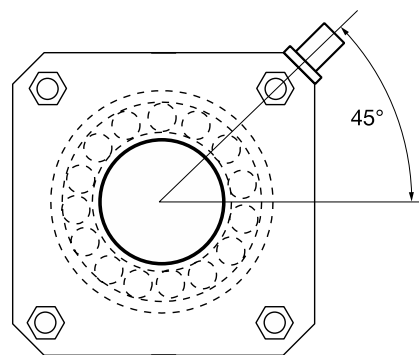


Fig. 2 Accelerometer's mounting position at bearing

Fig. 1 Experimental rig [19]

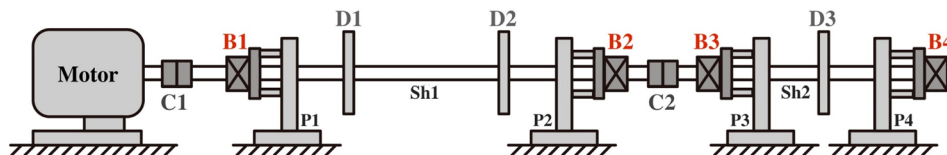


Table 1 Summary of number of measurements (samples) per fault (rotor condition) at machine operating at 1800 RPM and 2400 RPM

Rotor condition	Number of samples (runs)	
	1800 RPM	2400 RPM
Healthy (residual misalignment and residual unbalance)	86	45
Misalignment	121	82
Bow	204	104
Looseness	192	92
Rub	114	58
Total of samples	717	381

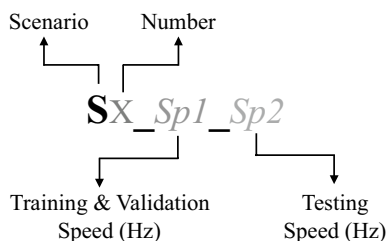


Fig. 3 Nomenclature of the scenarios used

of four grease lubricated ball bearings (B1, B2, B3, B4). Bearings are mounted on the flexible bearing pedestals (P1, P2, P3, P4), which are secured by bolts to a steel base that acts as foundation of the machinery within a high mass. The measured natural frequencies of the rig are 50.66 Hz, 56.76 Hz, 59.2 Hz and 127 Hz [19].

Vibration data are acquired at the sampling frequency of 10 kHz from four uniaxial accelerometers simultaneously [18–20]. The sensors are located at 45° from the horizontal line in anticlockwise direction as shown in Fig. 2. The sampling frequency of 10,000 Hz is used so that the measured frequency range can cover both rotor faults and bearing defects related to high frequency range. The accelerometer with a sensitivity of 100 mV/g and the measurement frequency range upto 10 kHz is used. The measured vibration data are available at the rotor speeds of 1800 RPM (30 Hz) and 2400 RPM (40 Hz) for following conditions.

- *Healthy condition* this is the baseline and it is subject of residual misalignment and residual unbalance. It is because both are difficult remove fully.
- *Faulty conditions* the four simulated faults are misalignment, looseness, bow and rub. They are considered to occur independently from each other. Each type of fault is introduced, separately, at two different locations in the rig. In the current study, each fault type irrespective of their locations in the rig is grouped into one fault class.

Data Preparation and Feature Extraction

The summary of number of vibration measurements (samples) per fault (machine condition) is listed in Table 1. Four scalar features are extracted from each of the data samples. The extracted parameters are arranged in the three different proposed scenarios, 1–3. The nomenclature used to name these scenarios is shown in Fig. 3.

Spectrum analyses are also carried to observe the dynamic behaviour of the rig with different machine conditions. Spectrum is showing dominant vibration peak at $1 \times$ (one times machine speed) but followed by its harmonics depending upon the rig fault conditions. However, the only time features are considered here to avoid extensive computing time and to find the effectiveness of the ML method in the VCM. The first selected feature is root mean square (RMS), typically used to represents the overall vibration amplitude, while the remaining ones are statistical parameters—variance, skewness and kurtosis. The variance, V , represents the signal power [13] and also useful if the mean of the signal is not zero. This behaviour is very much likely in the vibration response due to presence of faults. The asymmetry of the measured signal is represented by skewness, S . The kurtosis, K provides information about the shape distribution of the sample. These features provide useful information both qualitatively and quantitatively for any time domain data. Therefore, these parameters are selected as the representative features of the vibration data for the different machine fault conditions.

The four parameters are estimated by following Eqs. (1)–(4), where N is the number of points contained in the measured signal, $z(t_i) = z_i$.

$$RMS = \sqrt{\frac{1}{N} \sum_{i=1}^N z_i^2}, \tag{1}$$

$$V = \frac{1}{N} \sum_{i=1}^N (z_i - \bar{z})^2, \tag{2}$$

$$S = \frac{\frac{1}{N} \sum_{i=1}^N (z_i - \bar{z})^3}{\sigma^3}, \tag{3}$$

$$K = \frac{\frac{1}{N} \sum_{i=1}^N (z_i - \bar{z})^4}{\sigma^4}, \tag{4}$$

where $\bar{z} = \frac{1}{N} \sum_{i=1}^N z_i$ is the mean and standard deviation, $\sigma = \sqrt{\frac{1}{N} \sum_{i=1}^N (z_i - \bar{z})^2}$.

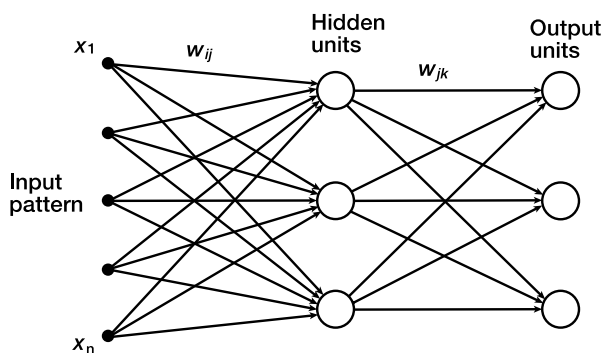


Fig. 4 Typical multi-layer perceptron neural network [21]

Machine Learning Model Construction

The ANN approach is used to develop the VML model for the fault diagnosis in rotating machines. The steps used and proposed architecture are discussed in this section.

ANN Method

Artificial neural networks are systems based on knowledge, which is generated by a training process that creates a correlation, in this case, between symptoms and their correspondent causes [14]. Since ANNs are not programmed, their performance depends on the quality and pre-processing of the acquired data, the network architecture and its design.

One of the first ANN used for simple tasks in classification problems is the perceptron, which with a simple-layer presents a narrow capability regarding its lack of ability to solve problems where the processed data are not linearly separable [21]. Because of this, a multi-layered network structure is proposed by introducing a hidden layer of weights between the inputs and the outputs. Based on backpropagation rule, a multi-layer perceptron (MLP) algorithm is developed (Fig. 4), with nonlinear activation functions which helps also to solve the noise–saturation dilemma due a network handling both small and large signals [13].

Proposed Architecture

In this study, a MLP is used to perform the pattern recognition and classification of acquired vibration data. By iterations, the network parameters, such as number of layers, number of neurons and types of functions, are adjusted and defined to obtain an accurate performance. It results into a feedforward network with four hidden layers, each of them

Table 2 Targets associated to each class for supervised pattern recognition

Class	Target vector
Healthy condition (residual misalignment and unbalance)	[1 0 0 0 0]'
Misalignment	[0 1 0 0 0]'
Bow	[0 0 1 0 0]'
Looseness	[0 0 0 1 0]'
Rub	[0 0 0 0 1]'

with a variable quantity of nonlinear neurons. The number of neurons varies according to the conducted experiments in the different scenarios. Signal components pass the input layer and move forward along the hidden layers, finishing with the result delivery from the decision layer, which have five possible classes as output (target vector, Table 2).

The neural transfer function for the hidden layers is hyperbolic tangent sigmoid, given in Eq. (5) with x_j an input vector and y_j the output returned [22]. In the output layer, it is used the normalised exponential transfer function (softmax), Eq. (6), which finally assigns a class to the input provided [23]. These functions are the most typically implementation for their respective tasks when conducting pattern recognition through ANN.

$$y_{j,\text{internal}} = \frac{2}{(1 + e^{-2 * x_j})} - 1, \quad (5)$$

$$y_{j,\text{output}} = \frac{e^{x_j}}{\sum_{k=1}^n e^{x_k}}. \quad (6)$$

The weights are obtained from the training process. A supervised learning is used for training the network; therefore, the processed data have already set the wanted outputs. The weights are adjusted according to the actual error of the outputs against the desired ones during this process. The error for each pair of single output (y_i) and its corresponding target value (t_i) is calculated through the Matlab cross-entropy function (ce) in Eq. (7). This function penalises the outputs depending on their accuracy obtained. Consequently, the closer are the outputs from the targets, the smaller the penalty applied.

$$ce = -t_i * \log(y_i). \quad (7)$$

Generalisation and Regularisation

Early stopping method is used in this work to avoid the overfitting in the network, improving the generalisation through a cross-validation of independent data [13]. By this process,

available data are into three sets. 70% of the samples are used for training the network, modifying the weights according the learning rule. 15% of the samples are used for validation, which is conducted by testing the trained network with these samples until their classification error reaches a desired point of minimum error, giving the order to stop the training process. At this point, the weights are the optimal for the network and, the last group of unknown data, 15%, is tested, providing the generalisation of the network [21].

Depending on the performance reached at the specific scenario, either scaled conjugate gradient or Bayesian regularisation functions are used to train the network. They differ in the manner they update the weights and bias values. While the first function uses the scaled conjugate gradient method, the second function uses the Levenberg–Marquardt optimisation, which allows the regularisation of the network through the obtainment of the smallest possible weights.

Target Matrix

As a supervised network, it is possible to identify five different binary target vectors, each one of them related to just one specific machine condition (Table 2). The target matrix in the ANN method is constructed such that where number of rows equals to the target vectors (defined machine conditions) and the number of columns depends on as samples per defined machine condition. For example, if three samples are used per each defined class (machine condition), the target matrix is given by Eq. (8).

$$\text{Target} = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \end{bmatrix}. \tag{8}$$

Model Outputs

The classifier performance is calculated as in Eq. (9). When error occurs, samples could be classified as false negative or false positive. Both errors have different consequences, being a ‘faulty as healthy’ the most critical diagnosis in any rotating machine.

$$\% \text{ Performance (or \% Diagnosis)} = \frac{\text{no. correct classifications}}{\text{total of inputs}} * 100\%. \tag{9}$$

Application of the Proposed Method with Machine Operating at 1800 RPM

Studied Scenarios

The three scenarios as per Fig. 3 are considered here. The scenarios are differentiated by the way the input vectors (i.e., 4 extracted features—RMS, variance, skewness and kurtosis per bearing) are used in the training for the proposed method of the ML model. The information related to the data management and arrangement is summarised in Table 3, while the specifications for the pattern recognition ML model for each scenario are in Table 4.

Scenario 1 (S1_30_30) In this scenario, each sample is taken from just an accelerometer randomly from only one bearing. This may be any one bearing from B1 to B4. This random data samples are used to generate the inputs for training and testing the ANN. The four features from Eqs. (1) to (4) are extracted from each sample to build up the input vectors. Their structure and nomenclature due the acquisition point are shown in Table 5. This scenario could be considered as the most complex, but may be useful if it is successfully detect the faults even with random measurement locations.

Scenario 2 (S2_30_30) This scenario consider the measurement again from one bearing only but the bearing

Table 3 Studied scenarios and data management

Scenario	Description of input’s construction and studied scenarios						
	Data origin	Input vector size (no. of parameters)	Measurement points	Total of samples (runs)	Training samples (70%)	Validation samples (15%)	Testing samples (15%)
S1	Randomly measured at only one bearing out of B1, B2, B3, B4	4	1	717	501	108	108
S2	B1 only	4	1	717	501	108	108
	B2 only						
	B3 only						
	B4 only						
S3	B1, B2, B3, B4 combined	16	4	717	501	108	108

Table 4 Specifications for pattern recognition ANN, all 3 Scenarios

	Scenario 1		Scenario 2		Scenario 3	
	Hidden layer number	No. of neurons	Hidden layer number	No. of neurons	Hidden layer number	No. of neurons
ANN parameter	1	1000	1	1000	1	1000
	2	1000	2	100	2	100
	3	100	3	100	3	100
	4	10	4	10	4	10
Transfer function hidden neurons	Hyperbolic tangent sigmoid		Hyperbolic tangent sigmoid		Hyperbolic tangent sigmoid	
Transfer function output neurons	Normalised exponential function (Softmax)		Normalised exponential function (Softmax)		Normalised exponential function (Softmax)	
Training function	Scaled conjugate gradient backpropagation		Scaled conjugate gradient backpropagation		Bayesian regulation backpropagation	
Performance function	Cross-entropy		Cross-entropy		Cross-entropy	

location is always fixed. Therefore, there are four cases (sub-scenarios), one per bearing with 717 input data samples, considered separately for analysis. The same four input vectors listed in Table 5 are again used for each case related to each bearing. For example, the input matrix for B1 sub-scenario, $Input_{B1}$, is shown in Eq. (10). Inputs at the other sub-scenarios for bearing B2, B3 and B4 are similarly constructed.

$$Input_{B1} = \begin{bmatrix} RMS1_1 & \dots & \dots & RMS1_{717} \\ V1_1 & \ddots & & V1_{717} \\ S1_1 & & \ddots & S1_{717} \\ K1_1 & \dots & \dots & K1_{717} \end{bmatrix}. \tag{10}$$

The four sub-scenarios for the scenario S2 have the same network architecture which is listed in Table 4. This allows to evaluate, through their performances, the quality of the information contained by the samples at each bearing location for the experimental rig.

Scenario 3 (S3_30_30) This scenario consider the simultaneous collection of vibration signal from all four bearings (B1–B4). In this scenario, each input vector is compounded by the combination of the features extracted from the four bearings, shown in Table 5. Therefore, each input contains 16 parameters per measurement. The *i*-input vector for scenario S3 is given by Eq. (11).

$$input_i = [RMS1_i \ RMS2_i \ RMS3_i \ RMS4_i \ V1_i \ V2_i \ V3_i \ V4_i \ S1_i \ S2_i \ S3_i \ S4_i \ K1_i \ K2_i \ K3_i \ K4_i]'. \tag{11}$$

Results and Discussion

Overall performances achieved at all the studied scenarios are presented with the machine conditions in Fig. 5 and also diagnoses by each VML model are summarised in Tables 6, 7, 8, 9, 10, 11 and 12. Scenario S1_30_30 is easiest approach for vibration measurements on any industrial machine but

Table 5 Input vectors built by features extracted at each measurement location in time domain

Measurement point	Input vector from <i>i</i> th sample (<i>i</i> th run)
B1	$[RMS1_i \ V1_i \ S1_i \ K1_i]'$
B2	$[RMS2_i \ V2_i \ S2_i \ K2_i]'$
B3	$[RMS3_i \ V3_i \ S3_i \ K3_i]'$
B4	$[RMS4_i \ V4_i \ S4_i \ K4_i]'$

the results in Table 6 indicate that the VML model was not good enough for all machine conditions during the data training, validation and testing. The diagnoses for different machine conditions for this scenario are listed in Table 7. It is observed that nearly 25% chance is that the healthy machine conditions may be diagnosed as the faulty machine conditions.

Similar results are also observed in case of Scenario S2_30_30 at different bearing numbers (Table 6) during the training, validation and testing process for the development of the VML model. However the performance of the Scenario S2_30_30 is observed to be much better compared to Scenario S1_30_30. It is because the consistent information from a particular bearing location provides certain features

of the machine dynamic behavior, which helps the VML model to perform better than the Scenario S1_30_30. The diagnosis performances of the VML model for each case of the Scenario S2_30_30 are also listed in Tables 8, 9, 10 and 11. It is also important to note that the performance of the VML model, S2_30_30_B3 is showing much better compared to other models at bearings B1, B2 and B4. This

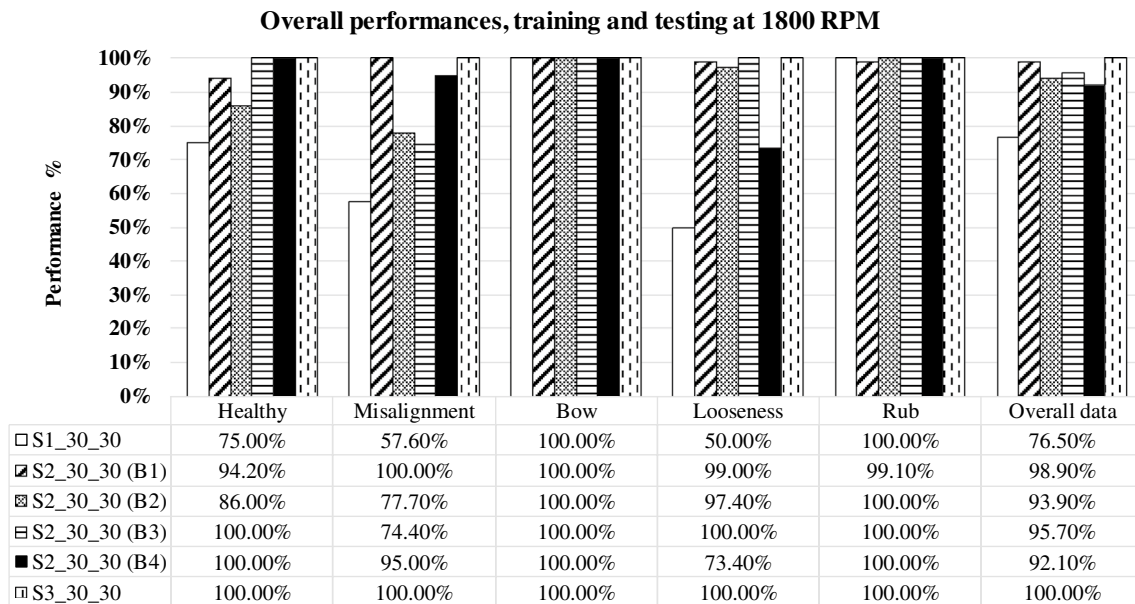


Fig. 5 Overall performances by fault type classification, all scenarios with testing, validation and training at 1800 RPM

simply indicates the bearing B3 contains better information about the machine dynamics behaviour.

The Scenario S3_30_30 shows the best performance with a 100.00% of accuracy in all of the stages of learning and training processes, as shown in Table 6 and the diagnosis of the machine conditions are listed in Table 12. Table 12 shows 100% accurate diagnosis of each machine condition; therefore, model is appropriate for the required purpose. These observations simply indicate that the dynamics from all bearing locations are required to map the machine dynamics accurately and hence accurate fault diagnosis is possible.

Blind Application at a Different Speed

It is very much likely that the prediction capability of the VML model is generally good enough if the machine and the operating conditions remain same. This is demonstrated in “Application of the proposed method with machine operating at 1800 RPM”. But the challenge is whether this developed ML model can be applied blindly to (a) the same machine but different operating conditions, (b) another identical machines but same operating conditions that was used for the VML model development, and (c) the combination of both (a) and (b).

Now the developed VML model for the rig at the rotating speed of 1800 RPM is blindly applied to the rig data at 2400 RPM (i.e., different operation condition) without

any training at 2400 RPM. This test is carried on all 2400 RPM data listed in Table 1. The model predicts the machine conditions accurately in two categories (Table 13)—healthy as 100% healthy machine condition and the remaining rotor faults as 100% machine faulty condition. The blind application of the VML model accurately predicted the faults of looseness, bow and rub except the misalignment. The healthy and fault wise results are listed in Table 14. However, the classification in two categories (Healthy and faulty) on the blind application is also an useful information for any plant maintenance. Therefore, the results are very encouraging for the future development of the VCM using the Industry 4.0 IoT.

Conclusions

A smart vibration-based machine learning (SVML) model is developed for the rotor faults diagnosis. The model is based on a multi-layer perceptron artificial neural network, which showed a successful performance in the diagnosis of faulty states in a rotating machine. Several scenarios are proposed and examined. It is concluded that the VML models based on the partial information about the machine (such as the scenarios S1 and S2) are not accurate enough for the industrial application. However the VML model in the scenario S3 provides 100% accuracy in the diagnoses of the machine conditions. This concludes that inclusion of the vibration measurements simultaneously

Table 6 Performances (%) achieved in all scenarios by stage

	Healthy	Misalignment	Bow	Looseness	Rub	All data
Performances %						
S1_30_30						
Training	74.1	58.9	100.0	51.1	100.0	77.5
Validation	74.4	55.4	100.0	48.1	100.0	74.0
Testing	80.4	54.5	100.0	47.7	100.0	74.2
Overall	75.0	57.6	100.0	50.0	100.0	76.5
S2_30_30—B1						
Training	93.8	100.0	100.0	100.0	100.0	99.2
Validation	85.7	100.0	100.0	96.6	100.0	98.1
Testing	100.0	100.0	100.0	96.8	93.1	98.1
Overall	94.2	100.0	100.0	99.0	99.1	98.9
S2_30_30—B2						
Training	82.3	79.5	100.0	97.7	100.0	93.6
Validation	91.7	75.0	100.0	92.6	100.0	93.5
Testing	100.0	70.6	100.0	100.0	100.0	95.4
Overall	86.0	77.7	100.0	97.4	100.0	93.9
S2_30_30—B3						
Training	100.0	77.1	100.0	100.0	100.0	96.2
Validation	100.0	73.7	100.0	100.0	100.0	95.4
Testing	100.0	63.2	100.0	100.0	100.0	93.5
Overall	100.0	74.4	100.0	100.0	100.0	95.7
S2_30_30—B4						
Training	100.0	93.0	100.0	71.7	100.0	90.6
Validation	100.0	100.0	100.0	82.6	100.0	96.3
Testing	100.0	100.0	100.0	75.0	100.0	94.4
Overall	100.0	95.0	100.0	73.4	100.0	92.1
S3_30_30						
Training	100.0	100.0	100.0	100.0	100.0	100.0
Validation	100.0	100.0	100.0	100.0	100.0	100.0
Testing	100.0	100.0	100.0	100.0	100.0	100.0
Overall	100.0	100.0	100.0	100.0	100.0	100.0

Table 7 Diagnoses (%) by the VML model, S1_30_30

Diagnosis	Actual				
	Healthy	Misalignment	Bow	Looseness	Rub
Healthy	75.0	0.0	0.0	0.0	0.0
Misalignment	25.0	57.6	0.0	28.5	0.0
Bow	0.0	17.4	100.0	21.5	0.0
Looseness	0.0	25.0	0.0	50.0	0.0
Rub	0.0	0.0	0.0	0.0	100.0

Table 8 Diagnoses (%) by the VML model, S2_30_30, B1

Diagnosis	Actual				
	Healthy	Misalignment	Bow	Looseness	Rub
Healthy	94.2	0.0	0.0	1.0	0.0
Misalignment	1.2	100.0	0.0	0.0	0.0
Bow	0.0	0.0	100.0	0.0	0.9
Looseness	4.7	0.0	0.0	99.0	0.0
Rub	0.0	0.0	0.0	0.0	99.1

from all bearings from a machine is capable to fully map the machine dynamics in the VML model. Furthermore, when this model is blindly tested with sets of data at a

different speed, the results provide an accurate separation of the samples into two categories—faulty and healthy. This observation encourages a possibility for centralised vibration-based condition monitoring (CVCM) for identical

Table 9 Diagnoses (%) by the VML model, S2_30_30, B2

Diagnosis	Actual				
	Healthy	Misalignment	Bow	Looseness	Rub
Healthy	86.0	15.7	0.0	1.0	0.0
Misalignment	10.5	77.7	0.0	1.6	0.0
Bow	0.0	0.0	100.0	0.0	0.0
Looseness	3.5	6.6	0.0	97.4	0.0
Rub	0.0	0.0	0.0	0.0	100.0

Table 10 Diagnoses (%) by the VML model, S2_30_30, B3

Diagnosis	Actual				
	Healthy	Misalignment	Bow	Looseness	Rub
Healthy	100.0	0.0	0.0	0.0	0.0
Misalignment	0.0	74.4	0.0	0.0	0.0
Bow	0.0	25.6	100.0	0.0	0.0
Looseness	0.0	0.0	0.0	100.0	0.0
Rub	0.0	0.0	0.0	0.0	100.0

Table 11 Diagnoses (%) by the VML model, S2_30_30, B4

Diagnosis	Actual				
	Healthy	Misalignment	Bow	Looseness	Rub
Healthy	100.0	0.0	0.0	0.0	0.0
Misalignment	0.0	95.0	0.0	26.6	0.0
Bow	0.0	0.0	100.0	0.0	0.0
Looseness	0.0	5.0	0.0	73.4	0.0
Rub	0.0	1.0	0.0	0.0	100.0

Table 12 Diagnoses (%) by the VML model, S3_30_30

Diagnosis	Actual				
	Healthy	Misalignment	Bow	Looseness	Rub
Healthy	100.0	0.0	0.0	0.0	0.0
Misalignment	0.0	100.0	0.0	0.0	0.0
Bow	0.0	0.0	100.0	0.0	0.0
Looseness	0.0	0.0	0.0	100.0	0.0
Rub	0.0	0.0	0.0	0.0	100.0

Table 13 Performances (%) achieved at blind testing of the VML model at 2400 RPM

Diagnosis	Actual	
	Healthy	Faulty
Healthy	100.0	0.0
Faulty	0.0	100.0

Table 14 Diagnoses (%) by the VML model in blind testing at 2400 RPM, S3_30_40

Diagnosis	Actual				
	Healthy	Misalignment	Bow	Looseness	Rub
Healthy	100.0	0.0	0.0	0.0	0.0
Misalignment	0.0	0.0	0.0	0.0	0.0
Bow	0.0	67.1	100.0	0.0	0.0
Looseness	0.0	0.0	0.0	100.0	0.0
Rub	0.0	32.9	0.0	0.0	100.0

machines operating at different rotating speeds. This smart CVCM model can be realised under the concept of Industry 4.0 Internet of Things (IoT) which is likely to overcome the current limitations of the experience and engineering judgements required for the machine faults diagnosis. Therefore, it allows an optimisation of the resources and offering a standard procedure for all identical machines across many worldwide plants within an organisation.

Acknowledgements Jyoti K. Sinha acknowledges his Ph.D. students then Dr Keri Elbhah for the development of the rig, and Dr Akilu Kaltungo and Dr Adrian Nembhard for the experiments and experimental data that are used in this study. Natalia Fernanda Espinoza Sepúlveda acknowledges the support by CONICYT (Comisión Nacional de Investigación Científica y Tecnológica/Chilean National Commission for Scientific and Technological Research) “Becas Chile” Doctorate’s Fellowship programme; Grant no. 72190062 for her Ph.D. study.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Dai X, Gao Z (2013) From model, signal to knowledge: a data-driven perspective of fault detection and diagnosis. *IEEE Trans Ind Inform* 9(4):2226–2238
2. Gangsar P, Tiwari R (2017) Comparative investigation of vibration and current monitoring for prediction of mechanical and electrical faults in induction motor based on multiclass-support vector machine algorithms. *Mech Syst Signal Process* 94:464–481
3. Nourmohammadzadeh A, Hartmann S (2015) Fault classification of a centrifugal pump in normal and noisy environment with artificial neural network and support vector. In: Dediu AH,

- Magdalena L, Martín-Vide C (eds) 4th International Conference on Theory and practice of Natural Computing. Springer, Cham, pp 58–70
4. Li C, Sanchez R, Zurita G, Cerrada M, Cabrera D, Vásquez RE (2015) Neurocomputing Multimodal deep support vector classification with homologous features and its application to gearbox fault diagnosis. *Neurocomputing* 168:119–127
 5. Liu Z, Cao H, Chen X, He Z, Shen Z (2013) Multi-fault classification based on wavelet SVM with PSO algorithm to analyze vibration signals from rolling element bearings. *Neurocomputing* 99:399–410
 6. Shen C, Wang D, Liu Y, Kong F, Tse PW (2014) Recognition of rolling bearing fault patterns and sizes based on two-layer support vector regression machines. *Smart Struct Syst* 13(3):453–471
 7. Ao H, Cheng J, Li K, Truong TK (2014) A roller bearing fault diagnosis method based on LCD energy entropy and ACROA-SVM. *Shock Vib*. <https://doi.org/10.1155/2014/825825>
 8. Tiwari R, Gupta VK, Kankar PK (2015) Bearing fault diagnosis based on multi-scale permutation entropy and adaptive neuro fuzzy classifier. *J Vib Control* 21(3):461–467
 9. Watany M, Eltantawie MA, Abouel-seoud SA (2015) Application of an adaptive neuro fuzzy inference system for low speed planetary gearbox vibration control. *J Low Freq Noise Vib Active Control* 34(3):323–341
 10. Nembhard AD, Sinha JK (2015) Unified multi-speed analysis (UMA) for the condition monitoring of aero-engines. *Mech Syst Signal Process* 64–65:84–99
 11. Shao R, Hu W, Wang Y, Qi X (2014) The fault feature extraction and classification of gear using principal component analysis and kernel principal component analysis based on the wavelet packet transform. *Measurement* 54:118–132
 12. Sun R, Tsung F, Qu L (2007) Evolving kernel principal component analysis for fault diagnosis. *Comput Ind Eng* 53(2):361–371
 13. Nahvi H, Esfahanian M (2005) Fault identification in rotating machinery using artificial neural networks. *Proc Inst Mech Eng Part C J Mech Eng Sci* 219(2):141–158
 14. Vyas N (2001) Artificial neural network design for fault identification in a rotor-bearing system. *Mech Mach Theory* 36(2):157–175
 15. Ben J, Fnaiech N, Saidi L, Chebel-morello B, Fnaiech F (2015) Application of empirical mode decomposition and artificial neural network for automatic bearing fault diagnosis based on vibration signals. *Appl Acoust* 89:16–27
 16. Walker RB, Vayanat R, Perinpanayagam S, Jennions IK (2014) Unbalance localization through machine nonlinearities using an artificial neural network approach. *Mech Mach Theory* 75:54–66
 17. Mohammed AA, Neilson RD, Deans WF, MacConnell P (2014) Crack detection in a rotating shaft using artificial neural networks and PSD characterisation. *Meccanica* 49(2):255–266
 18. Nembhard AD, Sinha JK (2016) Comparison of experimental observations in rotating machines with simple mathematical simulations. *Measurement* 89:120–136
 19. Nembhard AD, Sinha JK, Yunusa-Kaltungo A (2015) Development of a generic rotating machinery fault diagnosis approach insensitive to machine speed and support type. *J Sound Vib* 337:321–341
 20. Yunusa-kaltungo A, Sinha JK (2016) Sensitivity analysis of higher order coherent spectra in machine faults diagnosis. *Struct Health Monit* 15(5):555–567
 21. Tarassenko L (1998) A guide to neural computing applications. Arnold, London
 22. Vogl TP, Mangis JK, Rigler AK, Zink WT, Alkon DL (1988) Accelerating the convergence of the backpropagation method. *Biol Cybern* 59:257–263
 23. Bishop CM, Recognition P, Learning M (2006) New York. Springer, N.Y

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