THE PUBLISHING HOUSE
OF THE ROMANIAN ACADEMY

PROCEEDINGS OF THE ROMANIAN ACADEMY, Series A,
Volume 23, Number 2/2022, pp. 153–164

FAULT DETECTION ON A ROTATING TEST RIG BASED ON VIBRATION ANALYSIS AND MACHINE LEARNING

Iulian LUPEA[1], Mihaiela LUPEA[2]

[1] Technical University of Cluj-Napoca, Faculty of Industrial Engineering, Robotics and Production Management
[2] Babes-Bolyai University, Faculty of Mathematics and Computer Science, Cluj-Napoca
Corresponding author: Iulian LUPEA, e-mail: iulian.lupea@mep.utcluj.ro

Abstract. A fault detection system based on vibration analysis and machine learning techniques is proposed in this paper. A test rig comprising a flexible rotor supported by oscillating ball bearings with a central disc driven by a DC motor and a timing belt has been build. Six artificial faults are imposed on the central disc and on the timing belt transmission. The input dataset is a balanced one, containing 80 observations for each of the seven test rig health states (classes), corresponding to one healthy state and six faults. Twenty one features (in time and frequency domains) are extracted from only one uniaxial accelerometer and the tachometer sensor. An in-depth data analysis, by applying supervised and unsupervised processing techniques, aims at selecting the most relevant features used further as predictors in the multi-class classification task. A large set of classifiers from Matlab were trained and tested in order to find the best classification model that predicts the seven health states. The best results were provided by Quadratic Discriminant and Neural Network (Wide) with the accuracy 95% and 94% respectively.

Key words: fault detection, vibration analysis, timing belt, slender shaft, machine learning, multi-class classification.

1. INTRODUCTION

Rotating mechanical systems are important components in the manufacturing and automation industries. Health Condition Monitoring is important to sustain and maintain the machines at good condition and available for production. It is vital to detect, diagnose and analyze the severity of faults associated to the machine parts or elements. To predict the faults and to estimate the remaining time until failure will occur is important too. Vibration signals give early indication of mechanical failures such as misalignment, unbalance, bent shaft, looseness and faults on transmission belts, bearings and gears defects etc. [1, 2]. The fault detection problem aims to distinguish between healthy state and one or more faulty states of a machine.

Feature extraction methods from vibration signals are categorized into statistical time-domain features extraction, frequency-domain features extraction, time-frequency representation, phase-space dissimilarity measurement, complexity measurement and others [3]. The statistical time-domain features extraction method is found to be the most dominant method in typical rolling element bearings. The extraction of the frequency domain features, especially by using the fast Fourier transform (FFT) is very important, often sufficient for stationary signals processing. The time-frequency representation and phase-space dissimilarity measurement methods are suitable for non-stationary, non-linear and chaotic signals [4]. Time-frequency analysis methods are used for feature extraction in non-stationary signals of rotating machinery, like the short-time Fourier transform (STFT), continuous wavelet transform (CWT), wavelet packet decomposition (WPD) and the Wigner–Ville distribution [5, 6, 7]. In [8] a time-domain vibration signal analysis and an artificial Neural Network (ANN) for fault diagnosis are proposed for health monitoring of a pulley-belt rotating system. In [9] the maintenance of the transmission belts is observed. Various belt defects and belt diagnosis by using vibration time and frequency domain features are studied in [10]. Faults on belt conveyors with the focus on the detection of idler faults with WPD and support vector machine (SVM) are presented in [7]. Timing belts wear, misalignment and failure originating from change in the geometrical dimensions are

observed in [11]. We noticed from literature that the research in the belts health monitoring and fault detection with machine learning techniques is not extensive and for toothed belt even less.

In the present article six artificial faults are introduced in a rotating machinery test rig with slender shaft and timing belt transmission from the motor reducer. A fault detection system based on vibration analysis and machine learning techniques is proposed, with the aim of discriminating between the healthy state and the six possible faulty states of the test rig. This approach is an automatic machinery fault detection by using a vibration sensor versus a classical approach which is mainly based on frequency spectrum pattern peculiar to each defect and recognized by the specialist. Also, the proposed detection system is knowledge-based and learns a predictive model from specific vibration features extracted from data previously measured for the machinery states under observation. For a new measured vibration signal the detection system will recognize the health state (healthy state or one of the six faults) of the test rig.

The observed data is acquired from only one uniaxial accelerometer, unlike many other approaches [3], and very good results for fault prediction were obtained. The detection of the health machinery state is modeled as a multi-class classification problem, each state being represented as a class. There are seven health states: the healthy (no fault) state and six faults, characterized by twenty one statistical features extracted from the vibrational data. By applying supervised and unsupervised processing techniques, the most relevant features are selected and used further as predictors in classification. Different supervised machine learning algorithms (classifiers) from Matlab 2021 are applied in order to find the best classification model for the fault detection problem.

The common supervised machine learning techniques [12] based on manually extracted features include the classifiers: *decision trees* (DT) and DT *ensembles*, *support vector machine* (SVM), *Naïve Bayes* (NB), *neural networks* (NN), *k – Nearest Neighbor* (kNN), *discriminant-based classifiers*. Based on artificial neural networks with more hidden layers, deep learning approaches (Convolutional Neural Network, autoencoder) [6, 13, 14] are able to encode the input data in a good representation, suitable for prediction, thus the manual feature engineering step is not needed anymore.

The remainder of the paper is structured as follows. Section 2 presents the stand used for experiments, the data acquisition process, the faults and the features extracted from vibration signals. Section 3 is dedicated to the data analysis and the classification process. In the next section, the performance results of the classifiers employed in fault detection are presented. The last section contains conclusions and directions for future work.

2. TEST RIG SET UP

The test rig used for experiments comprises some common machine elements like a DC motor, a reducer, a timing belt connecting the reducer output with the flexible shaft sustained on ball bearings (Fig. 1). Some machine elements (like a gear pair) are possible to be added, hence the test rig is reconfigurable for future experiments.

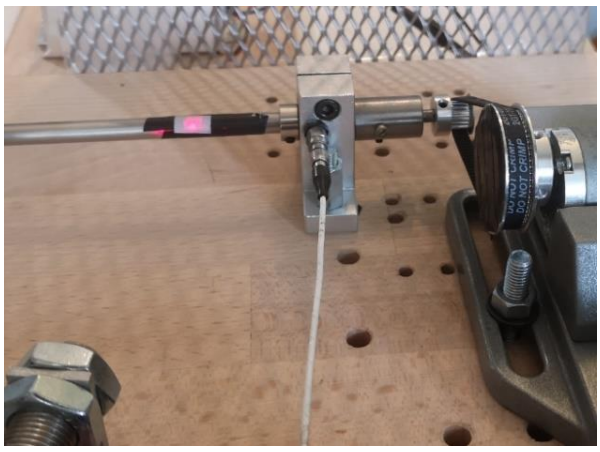

Fig. 1 – The test rig and two sensors.

The observed transmission chain is animated by a brushed Pololu DC motor (12 Volts)**,** coupled with the reducer of three stages (helical and spur gear pairs). The metal gearbox reducer is characterized by the 19:1 transmission ratio. The first stage of the reducer is a helical gear pair for reducing noise and vibrations and improving efficiency, followed by two stages of spur gears. The motor speed controller input voltage range is 9V to 60V and the output current range is from 0 to 20A. The rotation from the motor reducer output shaft is transferred to the flexible shaft through a timing synchronous belt of 2 mm tooth pitch, the length of 200 mm and the width of 6 mm. The belt has a special profile with rounded teeth which reduces backlash. It is a Neoprene – synthetic rubber reinforced with fiberglass cords. A timing belt aluminum pulley of $z_1 = 60$ teeth is mounted on the gearbox output shaft. This is the driver sheave of the timing belt; the rotation speed is n_1. A timing belt aluminum pulley of $z_2 = 18$ teeth is mounted on the driven shaft (rotation speed: n_2). The belt tensioning is done by manually translating the motor reducer with the driver sheave caught in a vice. The calibrated flexible shaft of 0.008 m diameter and 0.9 m length between the two double row oscillating ball bearings, has a disc at the center. The central disc is 450 mm apart from the support with the accelerometer. The shaft support's lowest resonance frequency is higher than the operating rotational speed during the experiment, hence we have rigid supports.

2.1. Data Acquisition

A dynamic acquisition board NI USB-4431 (24 bit, 102.4 kS/s) with the first two analog channels (out of four simultaneous acquisition channels) is used. The tachometer laser beam is pointing on the rotating shaft (Fig. 1). A uniaxial PCB shear piezoelectric accelerometer (352C68, sensitivity: 101 mV/g, frequency range: 0.5 Hz to 10 kHz, mass: 2 grams) is glued horizontally on the bearing support. The load is obtained by friction imposed on the opposite end of the shaft (outside the two bearings) by a tensioned plastic ribbon wrapped around the shaft.

A log data LabVIEW application with acquisition on two channels, one analog channel for the accelerometer and the second analog channel for the tachometer voltage pulses acquisition (one pulse per shaft revolution) is employed. The sampling rate of the analog signal is 5000 sampling per second, resulting the frequency theoretical bandwidth of 2.5 kHz. The ADC (converter) of USB-4431 board places the samples in an intermediate memory buffer. From the LabVIEW acquisition application the DAQmx Read.vi function is executed in a while loop in order to read the samples from the intermediate buffer. Data is saved in a log file from which is processed later. The acquisition and recording time on the two channels is 800 seconds for each health state. This acquisition time totals eighty observations, each of 10 seconds duration from which the features of one health state are extracted.

2.2. Faults and feature extraction

Six faults have been proposed. The health states of the observed system can be categorized in seven classes (C0, C1, …, C6). The first one is the healthy state, with no fault, the disc on the shaft considered with minimum eccentric mass. The other six states (classes) correspond to faults that will be described in the sequel. Eighty (80) observations have been recorded for each state. Each registration lasts for ten seconds and gathers 50000 samples for a sampling frequency of 5000 samples per second.

The first flexural mode of vibration or critical speed of the reference shaft (on vertical stiff bearings) supported by pin joints at each end, with a central disc, is at the frequency $f_1 = \omega_1/(2\pi)$, resulting a value of 17.56 Hz. The first natural circular frequency is given by Equation (1), according to the Ritz method [15], where the mode shape is approximated by a sum of orthogonal (sinusoidal) functions that satisfy the boundary conditions:

$$\omega_1 = \frac{\pi^2}{l}\sqrt{\frac{EI}{2l\,(M/2+m)}}, \tag{1}$$

where: EI (42.22 Nm^2) is the bending rigidity of the shaft (E is the Young modulus of elasticity, I is the area moment of inertia), M (0.35 kg) and m (0.055 kg) are the shaft and respectively the disc mass and l (0.9 m) is the shaft length between bearings. The shaft rotation speed (n_2) during the experiments is 700 rpm

(11.66 rps), with small variations about the nominal value. Being relatively far under the first critical speed, the behavior of the system is considered to not be affected by the first flexural mode of vibration of the shaft (even that the central mass *m* will be slightly modified by some of the faults observed in the study).

The initial state of the system, with no modifications, is considered with no fault or healthy, being denoted by class zero (C0). The central disc is integral with the shaft by using two M3 screws.

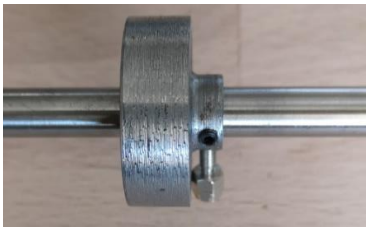

Fig. 2 – The rotor and eccentric mass.

The *first fault* (class C1) is obtained by a small eccentric mass attached radially to the central disc (Fig. 2). The *second fault* (class C2) is described by a slightly larger eccentric mass (one M3 nut, added on a longer M3 screw) attached radially to the same central disc. An eccentric mass is changing the radial vibration amplitude of the system and the statistical features of the signal measured at the bearing support and as well the power spectrum lines mainly at the level of the first order 1× turning speed n_2, where n_2 is the rotation speed (rps) of the shaft.

The *third system fault* (class C3) is referring to the timing belt, more exactly the offset misalignment of sheaves, without eccentric mass on the central disc. The offset is about 3.5 mm outward, obtained by the motor reducer translation from the initial (nominal) position. This fault type produces high vibration amplitude at 1x f_{sv} [rps] of the particular sheave rotational speed (f_{sv}), predominantly in the axial direction [2]. However, the defect is partially caught by the radially measuring accelerometer. We recall the width of the timing belt which is 6 mm and the width of the teeth of the sheaves which is about 7.3 mm. The effect of the fault is an increased overall vibration level causing accelerated wear for the belt and the sheaves. Wear or pulley/sheave misalignment is regularly indicated by high vibration amplitude at the timing belt frequency f_b, which is sub-synchronous with respect to n_2:

$$f_b = D_1 \pi n_1 / L_b = D_2 \pi n_2 / L_b \,, \tag{2}$$

where D_1, D_2 are the sheaves pitch diameters, n_1, n_2 are the rotational speeds of the sheaves and L_b is the belt length. The sheave – timing belt mesh frequency (Hz) is $f_{mesh} = z_2 n_2$, where n_2 is measured by the tachometer. At least 6 revolutions of the belt should be contained in an analyzed block data [16]. The block data lasts for 10 seconds and $f_b = 0.18\, n_2$, where n_2 is about 11.7 Hz.

The *fourth system fault* (class C4) is summing up the belt misalignment fault specific to the previous class (C3) with the small eccentric mass added to the disc (specific to class C1). The *fifth fault* (class C5) is the mentioned belt fault with the larger eccentric mass already used in class C2.

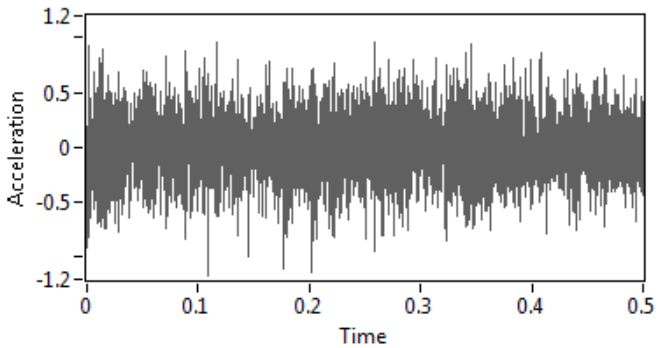

Fig. 3 – A signal segment from a C4 observation.

For the *sixth fault* (class C6) belt misaligned fault and any eccentric mass are removed, hence we are at the initial state specific to C0 and a thin and relatively soft ring is added transversally to the timing belt filling 80% the channel between two adjacent teeth. The defect can be seen as a foreign body between two adjacent belt teeth which is caught and compressed periodically between the belt and the sheaves. The frequency of the ring rotation is equal to the belt passing frequency f_b.

A small friction load at the shaft level is maintained for all states. A short portion (0.5s) of the measured acceleration signal belonging to the C4 class, is shown in Fig. 3. Portions of the signal belonging to other classes look similar, being generally difficult to visually distinguish the defect from signals in the time domain. Fault classes are depicted in Fig. 4.

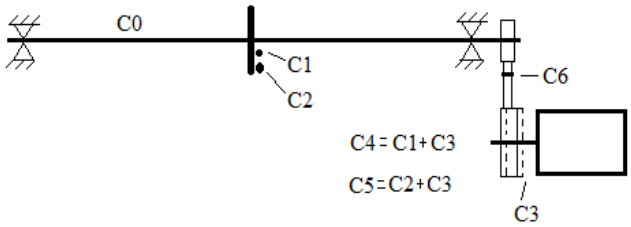

Fig. 4 – Test rig sketch with fault classes.

A number of time domain parameters has been employed as a first set of parameters. Spectral analysis, an important diagnostic tool providing information about the amplitude and phase content of the vibration at various frequencies, is represented by a second set of parameters. A LabVIEW graphical code has been programmed for the feature extraction from the observations. Totally, twenty one parameters (features) for each registration have been proposed in order to characterize and identify the faults, from which later a subset of the most important parameters has been selected and used for classification.

The first parameter is the average of the shaft rotation frequency n_2 (Hz) from the tachometer channel for the time window of a registration (10 seconds). Follows the set of time domain statistical parameters: median value, crest factor, peak to peak of the signal, mean value (μ), mean of absolute values, standard deviation (σ), variance (σ^2), root mean square (RMS), 3-rd moment about mean, 4-th moment about mean, skewness and kurtosis, all of them from the accelerometer time signal [17]. RMS level includes the mean value (the DC value of the signal). When DC value of the signal is zero, RMS equals standard deviation. Crest factor is the ratio of the maximum absolute value of the signal to the RMS value. For a signal with N sample values the m-th order moment is found in Table 1. For $m = 2$, the moment about the mean (μ) equals the population variance (σ^2). Skewness (x_{SK}) expresses the asymmetrical behavior of the vibration signal through its probability density function (PDF), Table 1. A negative value of skewness indicates that the left side of the probability density function graph is longer than the right side. A positive value indicates that the right side of the probability density function graph is longer than the right side. For a symmetric signal the skewness is zero. Kurtosis (x_{KU}) is formulated as shown in Table 1 and expresses the peak measurement (peakedness) of the distribution of the input sequence. Normal distribution has a kurtosis value of 3. A value less than 3 indicates a flatter distribution than normal and a value greater than 3 indicates a sharper distribution than normal.

Table 1

Time domain features

m-th order moment	Skewness	Kurtosis
$\sigma_x^m = \frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^m$	$x_{SK} = \frac{\sum_{i=1}^{N}(x_i - \mu)^3}{N\sigma_x^3}$	$x_{KU} = \frac{\sum_{i=1}^{N}(x_i - \mu)^4}{N\sigma_x^4}$

The first spectral parameter is the peak (P1) of the power spectrum in a narrow interval around the shaft rotational frequency n_2 followed by three harmonics (P2, P3, P4). The harmonics are obtained by

taking the peaks of the power spectrum in a narrow interval (of spectral lines) centered at two times, three times and four times the first peak value P1. Finally, four spectral features [18], the spectral centroid (*C*) or frequency-weighted sum normalized by the unweighted sum (Table 2), spectral spread (*S*) or standard deviation around the spectral centroid (Table 2), spectral skewness – computed from the third order moment and spectral kurtosis – computed from the fourth order moment, were added. A_k is the *k*-th power spectral coefficient magnitude and f_k is the associated frequency.

Table 2

Spectral centroid (*C*) and spectral spread (*S*)

$C = \sum_{k=1}^{N/2} f_k A_k \Big/ \sum_{k=1}^{N/2} A_k$	$S = \sqrt{\sum_{k=1}^{N/2} (f_k - C)^2 A_k \Big/ \sum_{k=1}^{N/2} A_k}$

With a sampling rate (f_s) of 5 kHz and the analysis time window of 10 seconds, the spectrum resolution is $\Delta f = f_s / N = 0.1$, where N = 50000 samples.

The fault detection problem will be modeled and solved as a multi-class classification problem. Based on the measured observations and the extracted features (parameters), different classifiers from Matlab will learn how to distinguish between the seven health states (classes).

3. DATA ANALYSIS AND CLASSIFICATION

The proposed fault detection system based on machine learning techniques is composed of the following phases: (1) Data acquisition presented in Section 2.1; (2) Feature extraction described in Section 2.2; (3) Data analysis: data preprocessing, feature selection and data visualization; (4) Train a classifier and build a predictive model; (5) Testing – fault prediction based on the predictive model.

The input dataset contains 560 observations, 80 observations for each of the seven classes (health states). The 21 features extracted from the vibration signals observes the following notations: F1 – shaft rotation frequency, F2 – median, F3 – peak to peak of the signal, F4 – mean of absolute values, F5 – mean, F6 – standard deviation, F7 – RMS, F8 – crest factor, F9 – 3-rd moment, F10 – 4-th moment, F11 – skewness, F12 – kurtosis, F13 – power spectrum PS1, F14 – PS2, F15 – PS3, F16 – PS4, F17 – spectral centroid, F18 – spectral spread, F19 – spectral skewness, F20 – spectral kurtosis and F21 – power spectrum at mesh frequency (f_{mesh}).

Three sets of features were analyzed and used in the experiments. The first one, S1, comprises the statistical time-domain features: {F1,...,F12} and the power spectrum at shaft first order F13, which is important for eccentric masses faults. The second set S2 extends S1 with four statistical frequency-domain features: F17, F18, F19, F20. The third set S3 = {F1,...,F21} contains all the parameters (features) mentioned before.

Feature scaling [19] is an important preprocessing step in some machine learning algorithms (kNN, SVM, clustering algorithms, neural networks) and its aim is to transform data that have large differences between their ranges in order to obtain comparable scales for all features. *Tree-based algorithms* (e.g. Decision trees, Ensembles based on DT) are not affected by feature scaling. *Standardization* is a scaling technique of an input vector using its statistical profile (μ, σ), where μ is the mean and σ is the standard deviation**.** After applying standardization on the input dataset, the statistical profile of all features is (0, 1), so the same scale. In the following the largest set of features, S3, standardized, is used to present the details of the data analysis phase.

3.1. Feature selection

Supervised feature analysis is a processing step prior to modeling, with the aim of scoring the relevance of each feature to the output/predicted results. Relief-based algorithms [20, 21] belong to the filter methods employed for feature selection and are based on a ‘proxy measure’. The k-Nearest Neighbor instances (Euclidian or Manhattan distances are applied) contribute to the calculation of the features’

weights. The weights range from –1 to 1, with positive values assigned to important, relevant, informative features, and negative values assigned to irrelevant features. Based on the weights, the features are ranked and the most relevant ones are selected to be included in the classification model. Because this technique is distance-based, a feature scaling/standardization preprocessing step is required.

The Matlab function *relieff* was applied to the initial dataset for ranking the 21 extracted features. The result is a descending order of the features with respect to their relevance in the prediction of the health states: F13, F17, F14, F19, F18, F1, F16, F15, F12, F20, F5, F21, F2, F11, F3, F9, F4, F7, F6, F10, F8.

In the features selection process the ones with negative weights were eliminated because they are not good predictors in classification and the dimensionality is also reduced. From the initial 21 features, the first 12 features in the previous order, {F13, F17, F14, F19, F18, F1, F16, F15, F12, F20, F5, F21}, having positive weights, were selected and used as predictors by the classifiers. Two of them, F5 (mean value) and F12 (kurtosis), are in the time domain. The most relevant features are in the frequency domain.

3.2. Data exploration and visualization

For the exploration and visualization of the input data *t-SNE* (*t-Distributed Stochastic Neighbor Embedding*) [22], an unsupervised, non-linear technique has been applied. The relatively high-dimensional space defined by the features is embedded into a low-dimensional space (2D or 3D) preserving similarities between points. This technique is based on an optimization procedure that generates low-dimensional points, such that the Kullback-Leibler divergence between a Gaussian distribution in the high-dimensional space and a *t* distribution in the low-dimensional space is minimized. The natural clusters in the original data are mapped into corresponding clusters in the low-dimensional space. The graphical representations in the low-dimensional space help to check whether the output classes are well separated and to compare the natural clusters with those classes.

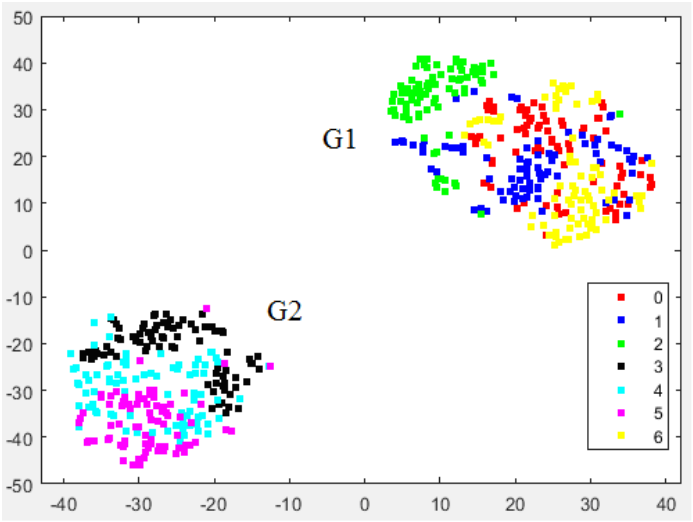

Fig. 5 – *t-SNE* visualization of data.

The new points in 2D, can be generated by the Matlab function *tsne* using different distance metrics (Euclidian, Cosine, Mahalanobis, Chebychev) starting from the raw or standardized input data. Figure 5 depicts a 2D *t-SNE* projection (Euclidian distance) of the initial dataset of 560 observations, standardized and restricted to the 12 selected features from the set S3.

One can observe two well separable groups (clusters). The cluster G1 contains the classes (C0, C1, C2, C6) in which the motor reducer is not translated. G2 cluster comprises the faults/classes C3, C4 and C5, in which the motor is translated. This seems to be true because of the friction (belt - sheave), although the accelerometer is measuring the bearing support radial vibrations instead of the axial vibration. Hence, the eccentric small mass fault (C1, C4) and slightly larger eccentric mass fault (C2, C5) have to be detected inside each of the two clusters G1 and G2, being less detectable in comparison to the timing belt pulleys

offset. The fault of the added ring on the timing belt (yellow dots – class C6) and the healthy state (red dots – class C0) are other classes, from G1 cluster, to be detected by the classification system.

Another unsupervised exploration of the initial dataset is based on the *classification difficulty measure* for the classes. The *classification difficulty* (CD) for a class is obtained as the percentage of the number of instances of that class that have the nearest neighbor (calculated using the Euclidian distance in the 12-dimensional space of the standardized selected features) pertaining to another class. The bigger the CD for a class, the less distinguishable is the class.

The values in Table 3 are consistent with the *t-SNE* visualization of data. In G1 the class C2 has CD = 0.05 and correlated with Figure 5 one estimate a very good detectability of the class. The class C3 with CD = 0.025 is well separable from the other two classes in G2. On the other side for C0 (CD = 0.275, the highest classification difficulty) one expect the lowest classification result in the G1 group. As well, C4 (CD = 0.26) is estimated to be the least detectable class inside the G2 cluster. These classification difficulties will be compared with the classification results.

Table 3

Classification difficulties

cluster	G1			G2			G1
class	C0	C1	C2	C3	C4	C5	C6
CD	0.275	0.225	**0.05**	**0.025**	0.26	0.15	0.17

3.3. Data classification

In the proposed supervised learning approach the training, validation and testing steps are followed. The input dataset is a balanced one containing 80 observations (instances) for each of the seven classes (health states), a total of 560 observations.

After the transformations in the preprocessing and processing steps a new dataset is obtained: 560 observations (instances) and 12 selected features (predictors). The transformed data was split into two subsets:

- the *training + validation subset* containing 87.5% of the initial dataset, meaning 490 instances (70 from each class). A 5 fold cross-validation was performed to prevent the overfitting and a classification model was built.
- the *testing subset* containing 12.5% of the initial dataset, meaning 70 instances (the remaining observations, 10 from each class). This set was used to predict the health states and evaluate the performance (accuracy) of the learned classifier.

Classifiers from Matlab have been applied to solve the proposed multi-class classification task. During the training stage, an internal validation of the model is needed. The dataset used to build a classification model is split into *training subset + validation subset*. The model learned in this phase is then improved by training it on the full dataset (training+validation) and a *final model* is obtained. In the testing phase the final model is used for prediction on the testing subset.

The results in the testing phase are reported using the confusion matrix and based on it different evaluation metrics [23] are calculated. *Accuracy* is an overall metric which expresses the probability that an arbitrary observation (instance) is correctly classified. To compare the performance results for the classes, three metrics, *Precision* (P), *Recall* (R) and *F*1-*score* are computed. F1-score is a balanced measure between Precision and Recall, and it is defined as their harmonic mean.

4. RESULTS AND DISCUSSIONS

This section presents the performance results of the applied classifiers from Matlab. Details and interpretations of the classification results at the level of classes for the initial dataset restricted to 12 selected features from the 21 extracted features (set S3) are also provided.

Table 4 (same colors for classes as in Fig. 5) contains the confusion matrix for Quadratic discriminant classifier and the calculated metrics. A test accuracy of 0.943 was obtain in a run, which proved to be a "favorable" random split of the training+validation/testing subsets. Recall (R) for the classes C1, C2, C3 and C6 is 1.0 (perfect sensitivity) so, the instances from these classes are always correctly classified by the classification model. The true classes C0 and C5 are correctly identified 90% of the time (R = 0.9). The lowest recall (R = 0.8) was obtained for C4. Precision (P) for the classes C0, C1 and C2 is 1.0 (perfect precision), meaning that when the classifier predicts that an instance belongs to one of the classes C0, C1, C2 it is correct 100% of the time. The lowest precision (P = 0.89) was obtained for C4. According to the F1-score the best distinguishable classes are C1, C2 with F1 = 1.0 (perfect classification), followed by C0, C3, C6 with F1 = 0.95 and C5 with F1 = 0.9. We note that classes with high F1-scores have low classification difficulties (Table 3). The class C4 is the least detectable one with F1 = 0.84, being confused with C3 and C5, fact that can be observed in cluster G2 (Fig. 5) and correlated with a high classification difficulty (Table 3). There are also small prediction errors for the class C0 part of G1 cluster together with C1, C2, C6. As a conclusion, the classification results at the level of classes are consistent with the *t-SNE* representation in Fig. 5 and the classification difficulties for the classes (Table 3).

Table 4

Confusion matrix for Quadratic Discriminant

70 tested instances		**Predicted Classes**							
True Classes		**C0**	**C1**	**C2**	**C3**	**C4**	**C5**	**C6**	**R**
	C0	**9**						1	0.9
	C1		**10**						1.0
	C2			**10**					1.0
	C3				**10**				1.0
	C4				1	**8**	1		0.8
	C5					1	**9**		0.9
	C6							**10**	1.0
	P	1.0	1.0	1.0	0.91	0.89	0.9	0.91	
F1-score		0.95	1.0	1.0	0.95	0.84	0.9	0.95	
Accuracy = (9+10+10+10+8+9+10) / 70 = 0.943									

From a practical perspective, for a predicted class (Table 4) one can conclude and give some examples as follows: a) for the predicted C3 class the operator has for sure to check and mend the timing belt misalignment offset and for 9% of the cases the small eccentric mass is added to the misalignment fault; hence, the misaligned timing belt is definitely present; b) for the predicted C4 class the timing belt misalignment is definitely a fault; in 89% of the cases the small eccentric mass defect is present, while in 11% of the cases the larger eccentric mass defect is added to the misalignment; c) for the predicted C0 class, definitely the system is in the initial healthy state but the system can be in the initial state for C6 class prediction as well, for 9% of the cases. Finally, for a C6 predicted class, in 91% of the cases a foreign body between two adjacent belt teeth is present and for 9% the system is with no fault.

For a correct evaluation of the performance of the classifiers we repeated the split of the data 15 times, and we ran the classifiers. In each split the instances used for training+validation and testing were selected randomly in the proportions of 87.5% and 12.5% respectively. For $n = 15$ runs, 15 test accuracies, a mean accuracy (*ma*), and a standard deviation (σ) were obtained. A confidence interval (*CI*) of the mean accuracy at the 95% confidence level was calculated using the formula: $CI = ma \pm 1.96\,\sigma / \sqrt{n}$ [24]. The classifiers with the best results (accuracies >85%) are reported in Table 5.

After a detailed analysis of the results provided by the classifiers in the testing phase, in all 15 runs, we note that all the classifiers, including those with lower accuracy, classify correctly the instances of the true classes C2, C3, C6, in at least 95% of the time. These results correlate well with the low classification

difficulties for these classes (Table 3). The best classifiers, with higher accuracy, succeeded at discriminating better the classes inside the two clusters G1 and G2 depicted in Figure 5.

For comparison, if all 21 extracted features are used as predictors in classification the accuracies of SVM classifiers are about the same, but for the other classifiers are smaller with 1% – 3%. The conclusion is that the feature selection is an important processing step prior to modeling.

The configurations of the classifiers employed in the experiments are as follows: (1) Discriminant Quadratic – covariance structure: full; (2) Neural Network wide and medium – one fully connected layer, first layer size 100 and 25 respectively, iteration limit: 1000, activation function: ReLU; (3) SVM – kernel functions: Quadratic/ Cubic/ Medium Gaussian, box constraint level:1; (4) Bagged trees – method: bag, number of learners: 30, maximum number of splits: 489; (5) Boosted trees – method: AdaBoost, number of learners: 30, maximum number of splits: 20, learning rate: 0.1; (6) Naïve Bayes – distribution: Kernel/ Gaussian; and (7) kNN weighted – number of neighbors: 10, distance metric: Euclidean, distance weight: squared inverse.

Table 5

Classifiers' performance (S3, 12 selected features)

Classifier type	Classification model	*CI* (%) of the mean accuracy over 15 runs
Discriminant	Quadratic	95.14±0.86
Neural Networks (NN) Standardized data	Wide	94.01±0.96
	Medium	91.05±1.1
Support Vector Machine (SVM) Standardized data	Quadratic	93.15±1.45
	Cubic	92.97±1.55
	Medium Gaussian	92.97±1.55
Ensemble Decision trees	Bagged trees	91.73±1.77
	Boosted trees	87.39±2.28
Naïve Bayes (NB)	Kernel	88.87±1.73
	Gaussian	87.89±1.47
k-Nearest Neighbor (kNN) Standardized data	Weighted	85.85±2.5

The experiments were conducted on the three sets of features, S1, S2 and S3 presented in Section 3 and the same set of classifiers has been employed. The results can be summarized as follows:

- For the set S1 (time-domain features), from 13 initial features, 8 were selected as relevant in prediction and they were able to correctly separate the health states (classes) in two big clusters (G1 and G2). The best accuracies (~83%) were obtained by Ensembles Decision Trees and SVM classifiers.
- In the set S2, 10 features (including the new 4 features in frequency domain) from 17 proved to be informative in the classification. SVM classifiers (accuracy = 87%) and Neural Networks (accuracy = 86%) provided the best performance results.
- Very good classification results were obtained by combining features from both time and frequency domains. From the largest set of features (S3, 21 features), 12 features and from these 9 in frequency-domain were the most relevant ones in the prediction of the faults. According to Table 5, seven of the employed classifiers provided accuracies from 90% to 95%. Quadratic Discriminant and the Wide Neural Network proved to be the best suited classifiers for the proposed fault detection task.
- At the level of health states (classes), the second, the third and the sixth faults are the best distinguishable ones regardless the configuration of the experiment: the set of features + the classifier. The features in the frequency domain helped at discriminating better the classes inside the two clusters.

5. CONCLUSIONS

A test rig comprising a slender shaft supported by oscillating ball bearings with a central disc, driven by a DC motor and a timing belt has been build. Six faults are imposed on the central disc and/or on the timing belt and a fault detection system based on machine learning techniques with Matlab was proposed. A uniaxial accelerometer is recording the horizontal vibration on the bearing support, close to the shaft drive system. Twenty one features were extracted from the accelerometer signal and the tachometer sensor. Some of the features are statistical time-domain features and the others are in the frequency domain. Three sets of features were gathered and used for fault prediction, observing the increase of the prediction performance.

To summarize, the contributions of the paper are as follows:

- A fault detection system based on vibration signals (from one uniaxial accelerometer), for a rotating test rig, was introduced as a proof of concept with application in real time systems.
- The in-depth vibration data analysis phase, as proposed in this approach, provided very important insights (relevance, correlation, visualization, classification difficulties) on the initial dataset represented by the extracted features. Relief-based algorithms were used to select the most relevant features as predictors in the multi-class classification task. The t-*SNE* representation of the dataset and the calculated classification difficulties for the classes were compared later with the classification results and they were consistent.
- An exploration of the effectiveness of a large set of classifiers employed on the three feature sets was performed. The Quadratic Discriminant and the Wide Neural Network, with accuracies of 95% and 94% respectively, applied to the 12 selected features (most of them in frequency domain), are the best suited classifiers for the proposed fault detection task.

To improve the fault detection one can mention techniques like a time synchronous averaged signal previous to the power spectrum derivation, considering the slight change in the shaft speed during experiment. As well, the sub harmonics of the shaft rotational speed associated to the belt frequency and its harmonics could discriminate better between the sixth fault state (a timing belt fault) and the no fault state. An additional axis of the vibration measuring accelerometer helps. A larger palette of machine element faults (including gears and bearings faults) and an increased number of accelerometer axis on the reconfigurable rotating machinery test rig is the subject of future research. As another further work direction we aim to develop the proposed methodology such as to fulfill the required attributes for industrial applications.

REFERENCES

1. H. AHMED, A.K. NANDI, *Condition monitoring with vibration signals – Compressive sampling and learning algorithms for rotating machines*, John Wiley & Sons Ltd., 2020.
2. C. SCHEFFER, P. GIRDHAR, *Practical machinery vibration analysis and predictive maintenance*, Elsevier, 2004.
3. W. CAESARENDRA, T. TJAHJOWIDODO, *A review of feature extraction methods in vibration-based condition monitoring and its application for degradation trend estimation of low-speed slew bearing*, Machines, **5**, *21*, art. 21, 2017.
4. A.M. UMBRAJKAAR, A. KRISHNAMOORTHY, R.B. DHUMALE, *Vibration analysis of shaft misalignment using machine learning approach under variable load conditions*, Hindawi Shock and Vibration, **4**, art. 1650270, 2020.
5. A.K.S. JARDINE, D. LIN, D. BANJEVIC, *A review on machinery diagnostics and prognostics implementing condition-based maintenance*, Mechanical Systems and Signal Processing, **20**, pp. 1483–1510, 2006.
6. C.E. RODRIGUES, C.L. NASCIMENTO, D.A. RADE, *Machine learning techniques for fault diagnosis of rotating machines using spectrum image of vibration orbits*, Annals of XXII Congresso Brasileiro de Automática (CBA), **2**, *1*, pp. 1–7, 2020.
7. W. LI, Z. WANG, Z. ZHU, G. ZHOU, G. CHEN, *Design of online monitoring and fault diagnosis system for belt conveyors based on wavelet packet decomposition and support vector machine*, Advances in Mechanical Engineering, **5**, pp. 1–10, 2013.
8. A. JABER, K.M. ALI, *Artificial neural network based fault diagnosis of a pulley-belt rotating system*, International Journal on Advanced Science, Engineering and Information Technology (IJASEIF), **9**, *2*, pp. 544–551, 2019.
9. https://industrialbeltdrives.com/wp-content/uploads/2017/11/20087_e2_preventive_maintenance_manual.pdf
10. A. BULUSHI, G. RAMESHKUMAR, M. LOKESHA, *Fault diagnosis in belts using time and frequency based signal processing techniques*, International Journal of Multidisciplinary Sciences and Engineering (IJMSE), **6**, *11*, pp. 12–20, 2015.
11. C. STOJANOVIC, S. TANASIJEVIC, N. MARJANOVIC, M. BLAGOJEVIC, *Failure analysis of the timing belt drives*, 12th International Conference on Tribology, pp. 210–215, 2011
12. E. ALPAYDIN, *Introduction to machine learning*, 2nd edition, MIT Press, Cambridge, 2010.
13. A. BRICIU, G. CZIBULA, M. LUPEA, *AutoAt: A deep autoencoder-based classification model for supervised authorship attribution*, 25th International Conference on Knowledge-Based and Intelligent Information & Engineering Systems, Procedia Computer Science, **192**, pp. 119–128, 2021.
14. I. GOODFELLOW, Y. BENGIO, A. COURVILLE, *Deep learning*, MIT Press, Cambridge, 2017.

15. G. BUZDUGAN, L. FETCU, M. RADEȘ, *Vibrațiile sistemelor mecanice*, Editura Academiei RSR, Bucharest, 1975.
16. www.emerson.com/ams, Vibration analysis for machinery health diagnosis, brochure-vibration-analysis-for-machinery-health-diagnosis-ams-en-6652272, 2019.
17. A. BRANDT, *Noise and vibration analysis*, John Wiley & Sons, 2011.
18. www.mathworks.com/help/audio/ug/spectral-descriptors.html
19. www.analyticsvidhya.com/blog/2020/04/feature-scaling-machine-learning-normalization-standardization
20. K. KIRA, L.A. RENDELL, *A practical approach to feature selection*, Proceedings of the Ninth International Workshop on Machine Learning, 1992, pp. 249–256.
21. R.J. URBANOWICS, M. MEEKER, W. LA CAVA, R.S. OLSON, J.H. MOORE, *Relief-based feature selection: Introduction and review*, Journal of Biomedical Informatics, **85**, pp. 189–203, 2018.
22. L. VAN DER MAATEN, G. HINTON, *Visualizing data using t-SNE*, Journal of Machine Learning Research, **9**, pp. 2578–2605, 2008.
23. M. GRANDINI, E. BAGLI, G. VISANI, *Metrics for multi-class classification: an overview*, arxiv.org/abs/2008.05756v1, 2020.
24. L. BROWN, T. CAI, A. DASGUPTA, *Interval estimation for a binomial proportion*, Statistical Science, **16**, *2*, pp. 101–133, 2001.

Received March 23, 2022