Article

Towards Developing an Automated Faults Characterisation Framework for Rotating Machines. Part 1: Rotor-Related Faults

Akilu Yunusa-Kaltungo [1,*] **and Ruifeng Cao** [1,2]

[1] Department of Mechanical, Aerospace and Civil Engineering, University of Manchester, Manchester M13 9PL, UK; ruifeng.cao@outlook.com

[2] School of Mechanical and Vehicle Engineering, Taiyuan University of Technology, Taiyuan 030024, China

[*] Correspondence: akilu.kaltungo@manchester.ac.uk

check for updates

Abstract: Rotating machines are pivotal to the achievement of core operational objectives within various industries. Recent drives for developing smart systems coupled with the significant advancements in computational technologies have immensely increased the complexity of this group of critical physical industrial assets (PIAs). Vibration-based techniques have contributed significantly towards understanding the failure modes of rotating machines and their associated components. However, the very large data requirements attributable to routine vibration-based fault diagnosis at multiple measurement locations has led to the quest for alternative approaches that possess the capability to reduce faults diagnosis downtime. Initiatives aimed at rationalising vibration-based condition monitoring data in order to just retain information that offer maximum variability includes the combination of coherent composite spectrum (CCS) and principal components analysis (PCA) for rotor-related faults diagnosis. While there is no doubt about the potentials of this approach, especially that it is independent of the number of measurement locations and foundation types, its over-reliance on manual classification made it prone to human subjectivity and lack of repeatability. The current study therefore aims to further enhance existing CCS capability in two facets—(1) exploration of the possibility of automating the process by testing its compatibility with various machine learning techniques (2) incorporating spectrum energy as a novel feature. It was observed that artificial neural networks (ANN) offered the most accurate and consistent classification outcomes under all considered scenarios, which demonstrates immense opportunity for automating the process. The paper describes computational approaches, signal processing parameters and experiments used for generating the analysed vibration data.

Keywords: spectrum energy; machine learning; data fusion; composite spectrum; vibration-based condition monitoring; rotating machines

1. Introduction

The past few decades have been characterised by very significant population growths (i.e., from 2.53 billion in 1950 to 7.16 billion in 2011 and predicted to reach 14.4 billion within the next 6 decades) across the world [1]. As a direct consequence of this unprecedented growth, the global primary energy consumption has correspondingly risen from 3701 Mtoe in the mid-1960s to approximately 13511 Mtoe in 2017 [1,2]. Besides the strains exerted by current energy demands on existing energy generation and distribution systems across the world, the sustainability of present scenarios is questionable, owing to the rapid depletion of primary energy resources such as crude oil, coal, and natural gas (estimated to be completely depleted in 50.2, 52.6 and 134 years, respectively) [2]. Additionally, the global upsurge

in initiatives aimed at shifting the energy paradigm towards renewable and more sustainable sources is gradually increasing the attention on alternative sources including wind, solar, etc. However, irrespective of energy generation source, the existence of rotating machines is almost inevitable along the entire supply chain, owing to their versatility and criticality. For instance, wind turbines are basically rotating machines with bearings, blades, gearbox and rotors accounting for incessant failure modes due to variable loading. Such rotating components are also associated with more established energy generation systems such as tidal, geothermal, hydroelectric, wave, etc. Routine health assessment of these systems is often challenged by high cost of downtime, due to their often remote locations. Based on this premises, the ability to cost-effectively guarantee the reliability of such critical energy generating physical industrial assets (PIAs) could be immensely enhanced through the development of potentially autonomous health monitoring methodologies that reduce human intervention.

The complexity and compactness of modern day industrial rotating machines has soared significantly over the past decade due to the advent of high power computational technologies as well as smart design initiatives. On one hand, these smart configurations have in some instances proven useful for reducing routine human interventions and subjectivity. On the other hand however, the resulting close proximity of such highly dynamic components (especially rotors, bearings, electric motors, gears, etc.) require an equally smart faults detection and diagnosis system if operational reliability of systems is to be sustained. Among most of the popular industrial rotating machinery health monitoring techniques, vibration analysis, in particular, has offered profound knowledge about the characteristics of individual components that make up a typical machine train. The guiding principle of any vibration monitoring system is that all machines offer some sort of warning prior to failure. Adequate understanding and trending of such warnings allows for the implementation of commensurate repair/replace actions prior to catastrophic failures.

Vibration signals from typical rotation machines is usually measured in the time domain and subsequently analysed in time [3,4], frequency or time-frequency domains. Central to most of the universally accepted vibration-based condition monitoring (VCM) is the frequency domain approach especially amplitude spectrum analysis [3], due to its interpretation simplicity as well as ease of computation. However, conventional amplitude spectra are often criticised for loss of phase information during the magnitude squared operation that precedes its generation of diagnostic features. This is perhaps why research efforts explored the feasibility of using approaches that possess the capabilities of retaining both amplitude and phase information, including higher order spectra (HOS) [5–21] and higher order coherence (HOC) [8,22–25]. While HOS and HOC components are able to retain phase and magnitude information, their predominant applications have been centred on individualised computation of features from distinct measurement locations. This approach often leads to heavy data burden especially when dealing with large industrial rotating machines that are often characterised by numerous bearings, with a knock-on effect on the downtime associated with fault diagnosis of such critical machines.

Efforts aimed at rationalising vibration-based condition monitoring data (VCM) through sensor reduction gave rise to the development of data fusion approaches such as coherent composite spectra (CCS) [26,27] and poly coherent composite spectra (pCCS) [7,18]. The principles guiding each of these approaches involve combining Fourier transformations (FT) at all measurement locations so as to generate a single coherent composite spectrum (CCS) [7,18,26–29]. The fundamental merit of this approach is that the single CCS provides a representation of the entire rotating machine, irrespective of the number of measurement locations, which significantly eases machine characterisation. While previous research endeavours on CCS have yielded very promising results, it is fair to highlight that the characterisation of faults have been manually implemented. The current study therefore aims to build on existing knowledge around CCS and pCCS data fusion application through two fundamental premises—(1) exploration of the possibility of automating existing approaches by testing its compatibility with several machine learning tools (2) incorporation of spectrum energy as a novel feature so as to account for the overall energy content of the CCS, thereby enhancing machine state

distinction. This paper therefore describes the computational approaches, signal processing parameters as well as the description of experiments used for generating the analysed vibration data. Additionally, the present study provides key observations on the compatibility of earlier developed CCS with machine learning, through a detailed comparison of the performance of five distinct machine learning tools.

2. Theoretical Background of the Modified Faults Diagnosis Architecture

The proposed faults diagnosis framework is based on a hybrid of various techniques so as to enable the strength(s) of one technique to compensate for the limitation(s) of others [28,29]. Figure 1 provides a flowchart of the proposed automated fault classification approach. The initial stage of the architecture is the collection of vibration data from all bearing locations, followed by the fusion of all such data into a single spectrum using the earlier developed CCS [7,18,26–29]. Once the CCS is developed, the spectrum energy beneath each of the CCS harmonics is then computed and centralised, prior to dimensionality reduction through principal component analysis (PCA). Feature sets with pre-existing labels are then trained through suitable supervised machine learning models. For feature sets without pre-existing labels, repeat the initial steps up until dimensionality reduction and then project onto the same linear space as the earlier PCA. Finally, the trained model is then used for routine fault classification. Full details of the mathematical computations are provided in the remainder of this section.
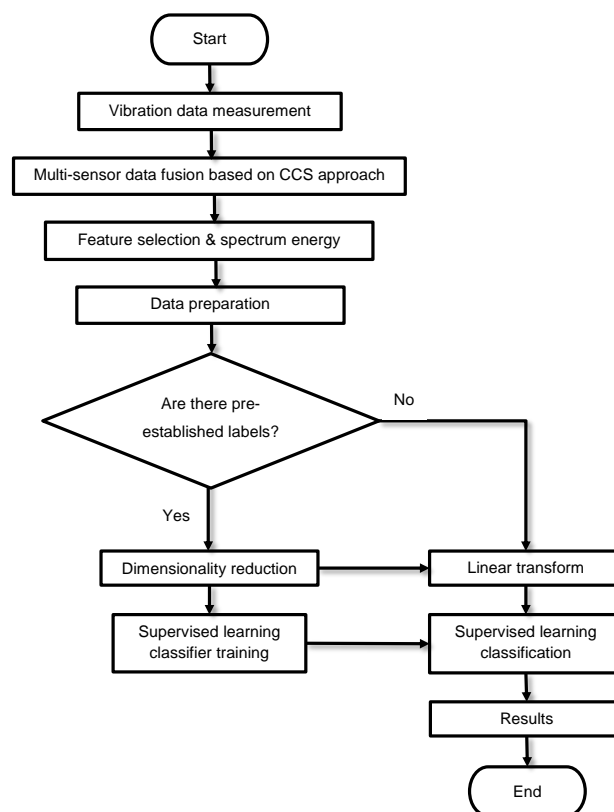

Figure 1. Proposed automated fault classification approach flow chart.

2.1. Coherent Composite Spectrum (CCS)

Full details of the computational approach of CCS have been described in several earlier studies [7,18,26–29]. However, in order for the present study to be independently comprehendible without necessarily consulting other articles, an abridged form of CCS [26–28] computation is again provided here. If the number of bearings on a typical industrial rotating machine is *b*, each of which has a fitted vibration sensor, the measured time-series signals can be divided into n_s equal-length

segments. The power spectral density (PSD) of the signal x_p from the pth bearing at the frequency f_k can be calculated as [26–28]:

$$S_{x_p x_p}(f_k) = \frac{\sum_{r=1}^{n_s}\left(X_p^r(f_k)X_p^{r^*}(f_k)\right)}{n_s} \tag{1}$$

where $X_p^r(f_k)$ is the discrete Fourier transform (DFT) of the rth segment of the signal x_p, and $X_p^{r^*}(f_k)$ is its complex conjugate, for $p = 1, 2, \cdots, b$. Similarly, the cross-power spectral density (CSD) for the signals x_p and x_{p+1} can be calculated as:

$$S_{x_p x_{p+1}}(f_k) = \frac{\sum_{r=1}^{n_s}\left(X_p^r(f_k)X_{p+1}^{r^*}(f_k)\right)}{n_s} \tag{2}$$

The coherence of the signals x_p and x_{p+1} for background noise suppression can be calculated as:

$$\gamma_{p(p+1)}^2(f_k) = \frac{\left|S_{x_p x_{p+1}}(f_k)\right|^2}{S_{x_p x_p}(f_k)S_{x_{p+1} x_{p+1}}(f_k)} \tag{3}$$

The coherent CSD of the signals from the pth and $(p+1)$th bearings can then be defined as:

$$S_{x_p \gamma_{p(p+1)}^2 x_{p+1}}^r(f_k) = \left[X_p^r(f_k)\gamma_{p(p+1)}^2(f_k)X_{p+1}^{r^*}(f_k)\right] \tag{4}$$

Therefore, each of the rth segments from each signal can be fused into a single component, $X_{CCS}^r(f_k)$, thus:

$$X_{CCS}^r(f_k) = \sqrt{\left(S_{x_1 \gamma_{12}^2 x_2}^r(f_k)S_{x_2 \gamma_{23}^2 x_3}^r(f_k)\cdots S_{x_{(b-1)} \gamma_{(b-1)b}^2 x_b}^r(f_k)\right)^{\frac{1}{(b-1)}}} \tag{5}$$

The CCS for the entire machine can then be calculated as:

$$S_{CCS}(f_k) = \frac{\sum_{r=1}^{n_s} X_{CCS}^r(f_k)X_{CCS}^{r^*}(f_k)}{n_s} \tag{6}$$

The $S_{CCS}(f_k)$ is a sequence of complex numbers that allows for the computation of the single-sided amplitude spectrum of the CCS generated from Equation (6):

$$A_{CCS}(f_k) = \left|\frac{2}{N}S_{CCS}(f_k)\right|, \quad (k = 1, 2, \cdots, N/2) \tag{7}$$

2.2. Feature Extraction

Earlier CCS based diagnoses have only applied maximum amplitude values of each harmonic during classification. In the current study however, we explore the use of an entirely new feature—spectrum energy (SE). Just as root mean square (rms) gives a representation of the energy content of the time waveform, SE provides information about the energy content of the measured vibration signal in the frequency domain. Considering that the CCS represents several time waveforms becomes evident [30]. SE is particularly useful because it is a universal indicator capable of showing changing trends in vibration data either due to the dynamic characteristics of the machine of interest or emergence of incipient faults. For a typical $A_{CCS}(f_k)$ computed as per Equation (7) at a frequency

f_k, where $f_k = (k-1)df$, $k = 1, 2, \cdots, N/2$, N is the number of data points and df is the frequency resolution, the SE between the selected harmonics at intervals of df can be defined as:

$$A_{SE}(f_k) = \sum_{i=k-l}^{k+l} A_{CCS}(f_i) \times df, \quad (k = 1, 2, \cdots, N/2; \; l = 10, 15, \; 20) \tag{8}$$

2.3. Dimensionality Reduction

The complexity of industrial rotating machines makes typical vibration data non-linear and highly dimensional. This dimensionality often makes diagnosis difficult, owing to the interference of redundant data points with core information that actually indicates variability. In this study, PCA was used to reduce data dimensionality, while still retaining the data points that aid the differentiation of various machine states [31–33]. This is particularly useful because it helps rationalise the amount of data that needs to be analysed, thereby reducing the amount of time required to implement the necessary repair/replace decisions. The implementation of PCA in this study was performed in 2 stages. The former stage involved centralising the datasets while the latter stage involved performing the singular value decomposition (SVD) on the aforementioned centralised data.

2.3.1. Data Centralisation and Standardisation

For a real matrix $A \in R^{m \times n}$, where m is the number of samples and n is the number of features (dimensions), PCA requires the centralisation of each column. In the matrix, $A \in R^{m \times n}$, a_{ij} represents typical elements of the matrix, while x_{ij} on the other hand is a corresponding element of the centralised matrix X.

To centralise the matrix A, the element x_{ij} of the centralised matrix X is defined as:

$$x_{ij} = a_{ij} - \overline{A}_j \tag{9}$$

where $\overline{A}_j$ is the sample mean of the elements of the jth column of matrix A, which can be computed as:

$$\overline{A}_j = \frac{1}{m} \sum_{i=1}^{m} a_{ij} \tag{10}$$

It is vital to note that the centralised data is of a column-wise zero mean form, which is essential for the subsequent computations.

To standardise the matrix A, the element x_{ij} of the standardised matrix X is defined as:

$$x_{ij} = \frac{\left(a_{ij} - \overline{A}_j\right)}{S_j} \tag{11}$$

where S_j is the sample standard deviation of the jth column of A, which is mathematically represented as:

$$S_j = \sqrt{\frac{\sum_{i=1}^{m}\left(a_{ij} - \overline{A}_j\right)^2}{m-1}} \tag{12}$$

The standardised data is column-wise mean zero with standard deviation that equals unity, while still retaining the shape properties of the original data.

2.3.2. Singular Value Decomposition (SVD)

SVD here was used to perform PCA on the centralised/standardised data, where the SVD of matrix X is defined as:

$$X = U\Sigma V^T \tag{13}$$

where the matrix $\boldsymbol{\Sigma} \in R^{m \times n}$ is a rectangular diagonal matrix of positive numbers σ_i, called the singular values of $\boldsymbol{X}$. The columns of the matrix $\boldsymbol{U} \in R^{m \times m}$ are orthogonal unit vectors, and referred to as the left singular vectors of $\boldsymbol{X}$. Similarly, the columns of the matrix $\boldsymbol{V} \in R^{n \times n}$ are orthogonal unit vectors and are referred to as the right singular vectors of $\boldsymbol{X}$.

The covariance matrix $\boldsymbol{C} \in R^{n \times n}$ can thus be computed as:

$$C = \frac{1}{m-1} X^T X = \frac{1}{m-1} \left(U\Sigma V^T \right)^T U\Sigma V^T = \frac{1}{m-1} V\Sigma^T \Sigma V^T = \frac{1}{m-1} V\hat{\Sigma}^2 V^T \tag{14}$$

where $\boldsymbol{U}^T \boldsymbol{U} = \boldsymbol{E}$, and $\hat{\boldsymbol{\Sigma}}^2$ is defined as:

$$\hat{\boldsymbol{\Sigma}}^2 = \boldsymbol{\Sigma}^T \boldsymbol{\Sigma} = \begin{bmatrix} \sigma_1^2 & & & \\ & \sigma_2^2 & & \\ & & \ddots & \\ & & & \sigma_n^2 \end{bmatrix} \tag{15}$$

By comparing this to the factorisation of the eigenvectors of the covariance matrix $\boldsymbol{C}$, it can be seen that the right singular vectors $\boldsymbol{V}$ of $\boldsymbol{X}$ are in fact equivalent to the eigenvectors of $\boldsymbol{C}$. A relationship between the eigenvalues λ_i of $\boldsymbol{C}$ and the singular values σ_i of $\boldsymbol{X}$ can also be derived thus:

$$\lambda_i = \frac{\sigma_i^2}{m-1} \tag{16}$$

The SVD therefore enables the calculation of the score matrix (result) $\boldsymbol{T} \in R^{m \times n}$ for a PCA, which can be mathematically represented as:

$$T = XV = U\Sigma V^T V = U\Sigma \tag{17}$$

By considering only the first L largest singular values and their corresponding singular vectors, the truncated score matrix $\boldsymbol{T}_L \in R^{m \times L}$ can be defined as:

$$T_L = U_L \Sigma_L = XV_L \tag{18}$$

where $\boldsymbol{U}_L \in R^{m \times L}$, $\boldsymbol{\Sigma}_L \in R^{L \times L}$, and $\boldsymbol{V}_L \in R^{n \times L}$.

2.4. Supervised Learning

Supervised learning is a specific machine learning category, whereby an algorithm can either learn a pattern or build a model (function) using labelled training data, and subsequently infer new instances based on such earlier learned patterns or models [33–39]. Solving a specific supervised learning problem requires the following various steps including; data type determination, training dataset collection, input features determination, learning algorithm determination, adjustment of the learning algorithm parameters and learning accuracy evaluation. Considering that this study represents the first application of machine learning to CCS-based faults diagnosis, the authors found it useful to explore a wide range of machine learning algorithms [40,41], so as to compare their performance under each of the considered scenarios and then select the most appropriate with regards to ease of deployment and accuracy. Based on this premise, the five different classes of supervised learning algorithms considered for this study are k-Nearest Neighbours (k-NN), Naïve Bayes classifier, linear support vector machine (SVM), Gaussian SVM, artificial neural networks (ANN) and K-fold cross-validation. The justification for selecting these particular learning techniques is mainly their reasonably straightforward computational approach and verifiable success with non-linear datasets including vibration data.

2.4.1. k-Nearest Neighbours (k-NN)

k-NN is an instance-based learning algorithm that assumes that instances in a dataset are in the vicinity of other instances, in feature space, with similar properties [40]. The classification of an object is determined by a "majority vote" of its neighbours, and each object is assigned to the most common class of its k nearest neighbours, where k is minimally small integer. In the case k = 1, the class of the object is solely determined by its nearest neighbour. If we define a set of training data with labels, $T = \{(x_1, y_1), (x_2, y_2), \cdots, (x_N, y_N)\}$, where x_i is the feature vector of each instance and $y_i = c_1, c_2, \cdots, c_K, i = 1, 2, \cdots, N$ is the corresponding label. For a test sample (x, y), the k instances nearest to x are considered by the k-NN algorithm based on a given distance measure. The set of k nearest neighbours of x is written as $N_k(x)$. The label of test sample x can thus be decided by the following decision function:

$$y = \arg\max_{c_j} \sum_{x_i \in N_k(x)} I(y_i = c_j), \quad i = 1, 2, \cdots, N; \; j = 1, 2, \cdots, K \tag{19}$$

where I is the indicator function.

2.4.2. Naïve Bayes Classifier

The Naïve Bayes classifier is a probabilistic classifier based on Bayes' Theorem, under the assumption of conditional independence between features [40,42]. For the same training set $T = \{(x_1, y_1), (x_2, y_2), \cdots, (x_N, y_N)\}$ with labels $y_i = c_1, c_2, \cdots, c_K, i = 1, 2, \cdots, N$, the number of the possible values for x^l $(l = 1, 2, \cdots, n)$ is given by S_l and the number of the possible values for Y is given by K. Firstly, the assumption of conditional independence implies that the joint probability distribution $P(X, Y)$ of the input and output can be calculated by the Naive Bayes classifier, using the conditional probability distribution displayed in Equation (20):

$$P(X = x | Y = c_j) = P(X^{(1)} = x^{(1)}, \cdots, X^{(n)} = x^{(n)} | Y = c_j) = \prod_{l=1}^{n} P(X^{(l)} = x^{(l)}), j = 1, 2, \cdots, K \tag{20}$$

Secondly, based on the learned model and by applying Bayes' Theorem, the output label y with the maximum posterior probability can be computed given any input x as:

$$P(Y = c_j | X = x) = \frac{P(X = x | Y = c_j) P(Y = c_j)}{\sum_j P(X = x | Y = c_j) P(Y = c_j)} \tag{21}$$

$$y = \arg\max_{c_j} \sum_{x_i \in N_k(x)} P(Y = c_j) \prod_l P(X^{(l)} = x^{(l)} | Y = c_j) \tag{22}$$

2.4.3. Support Vector Machine (SVM)

SVM constructs a hyperplane or a set of hyperplanes, in high or infinite-dimensional space, which can be used for classification, regression or other tasks [34,39]. The generalisation error of this classifier can be best reduced when the classification boundary is far away from the nearest training data points. The SVM can be viewed as a constrained quadratic optimisation problem, which can be solved using the method of structural risk minimization [42]. The SVM constructs an optimal separation hyperplane $f(x) = 0$ between datasets, given by:

$$y = f(x) = W^T x + b = \sum_{i=1}^{N} W_i x_i + b \tag{23}$$

where W is an N-dimensional vector and b is a scalar.

2.4.4. Artificial Neural Network (ANN)

ANNs are mathematical or computational models that imitate the structure and function of a biological neural network such an animal's central nervous system, particularly the brain. The networks can be used to estimate or approximate any non-linear function. These networks have the ability to "learn" and summarise data; through experimental application to known data, which implies that ANNs are able to reasonably act as automatic recognition systems by comparing local situations determined by the complexity of learning and solving practical problems under different scenarios. The simplest and most widespread form of an ANN consists of an input layer, a hidden layer, and an output layer. Generally speaking, an ANN is composed of multiple layers of "neurons". Each layer of neurons has an input and output, where the input is the output from the previous layer. Each layer, $Layer(i)$, is composed of N_i (where N_i signifies N on the ith layer), and each layer of neurons on N_i is composed of network neurons [34,36,40]. The collateral neuron takes the output of the corresponding neuron on N_{i-1} as its input. The connections between neurons are called synapses and each synapse is assigned a weight, which determines the contribution of the previous neuron on the subsequent ones [34,36,40]. The output y can thus be written as:

$$y = f\left(W^T x\right) = f\left(\sum_{i=1}^{N} W_i x_i + b\right) \tag{24}$$

where f is the activation function, W are the weights and b is the scalar bias term. The weights W are assigned through an iterative training process. The transfer function adopted here is the sigmoid symmetric transfer function. Since the ANN type is backward propagation, scaled conjugate gradient (SCG) was used as learning algorithm as well as for overfitting avoidance.

2.4.5. K-Fold Cross-Validation

Cross-validation is a model validation technique that enables a machine learning algorithm's prediction accuracy to be evaluated in practice [41]. In order to perform K-fold cross-validation, the original dataset is split randomly into K subsets of roughly equal size [41]. Of these K subsets, a single subset is retained to be used as validation data, while the other $K - 1$ subsets are used as training data. Cross-validation is then repeated K times, with each subset being used exactly once for validation. These K results are then averaged to obtain a single estimation. The most common value used is $K = 10$.

3. Experimental Organisation and Data Source

In order to foster a substantial acceptability of any newly proposed technique, an experimental validation is often required, and this is commonly achieved through the aid of an experimental test rig. A representative test rig may be described as that which has been adequately set-up to correctly simulate the investigated faults in the research. Typical industrial rotating machines are complex and multi-components (e.g., rotors, bearings, gears, couplings, electric motors, blades, etc.) structures that often possess multiple failure modes [43]. As it would be unrealistic to examine all faults classes at a single instance, this study only considers some of the most common rotor-related (i.e., low frequency) faults. The remainder of the section presents further details of various components that comprise the test rig as well as the experimental simulation of the studied rotating machine conditions.

3.1. The Rig

The experimental rig is primarily made up of two mild steel shafts which are connected by a rigid coupling. The longer shaft has a length of 1000 mm and a diameter of 20 mm, while the shorter has a length of 500 mm and a diameter of 20 mm. The longer shaft end of the entire rotor assembly was then coupled to a 0.75 kW electric motor via a flexible coupling. In addition to connecting the

rig components, the flexible coupling on the motor end of the rig assembly also serves the function of preventing the transmission of motor-related faults signals. In order to ensure rig balance and acceptable rotor deflections, three similarly machined bright mild steel discs with dimensions of 125 mm × 20 mm × 15 mm (i.e., outer diameter × inner diameter × thickness) were mounted at equal distances on each rotor. The two balance discs mounted on the 1000 mm shaft were located at 300 mm and 190 mm from the flexible coupling and bearing 2, respectively. The third balance disc was placed at mid-span of the 500 mm shaft (i.e., 210 mm from bearings 3 and 4, respectively). The shaft assembly was then supported by 4 SKF-type flange-mounted antifriction ball bearings. Each bearing was secured to its pedestal by four 6 mm-thick bright mild steel threaded bars. Vibration data from each bearing is collected via a 100 mV/g accelerometer mounted at 45 degrees. The rationale behind the diagonally mounted accelerometer is sensor reduction as well as the expectation that both vertical and horizontal responses can be reasonably represented. Figure 2 shows the experimental rig assembly and its core components.
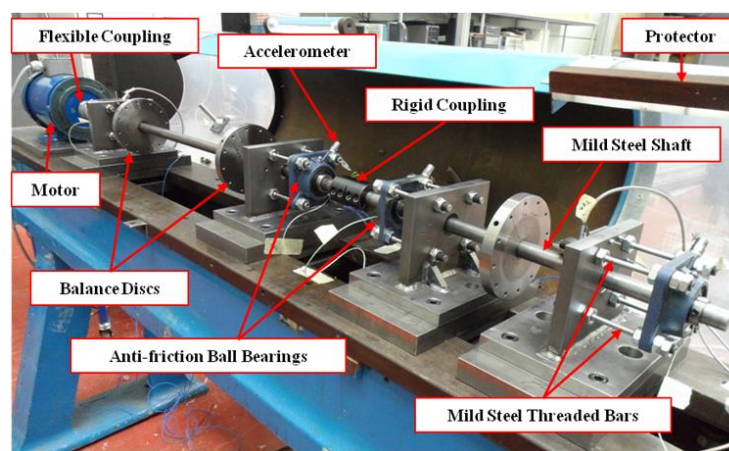

Figure 2. Experimental rig.

3.2. Dynamic Characteristics

Experimental modal analysis is a widely acknowledged design testing and qualification tool across various engineering disciplines [44]. Adequate knowledge about the modal properties of a structure immensely paves way for design upgrades, faults diagnosis and remaining useful life enhancement [45,46]. Based on this premise, the impact-response method of experimental modal analysis was also used here to establish an understanding of the dynamic behaviours of the rig as well as validate the origins of subsequent faults diagnostic [47–50] features (especially spectral peaks). During the experimental modal testing, the complete rig assembly was excited by an ICP-PCB type instrumented hammer and the corresponding vibration responses were measured using the accelerometers. The first few natural frequencies (by appearance) of the experimental rig assembly were determined to be 47 Hz, 55.54 Hz, 57.98 Hz and 127 Hz. The use of threaded bars to connect bearings to their pedestal provided comparable flexibilities in both vertical and horizontal directions. Hence the natural frequencies in both directions were very similar.

3.3. Simulation of Faults

A total of five (i.e., baseline relatively healthy and four faults cases as shown in Figure 3) commonly encountered rotor-related cases were experimentally simulated, so as to cover a reasonably wide range of practical operating conditions of typical industrial rotating machines. It was impossible to achieve a perfect alignment while setting up the baseline case, as often is the situation in real-life scenarios. A summarised description of the mode of experimental simulation for each case, the location of the fault on the rig assembly and the severities are provided in Table 1 for simplification. Considering

that several modern-day industrial rotating machines operate at multiple speeds, vibration data were collected at 3 distinct speeds; 1200 rpm (20 Hz), 1800 rpm (30 Hz) and 2400 rpm (40 Hz) for each case. This approach provides the opportunity to extensively understand the dynamics of the studied class of machines under different operating conditions. Under each experimental scenario (i.e., under each case and each machine speed), vibration data were obtained at a sampling rate of 10 kHz over a timespan of approximately 120 seconds.
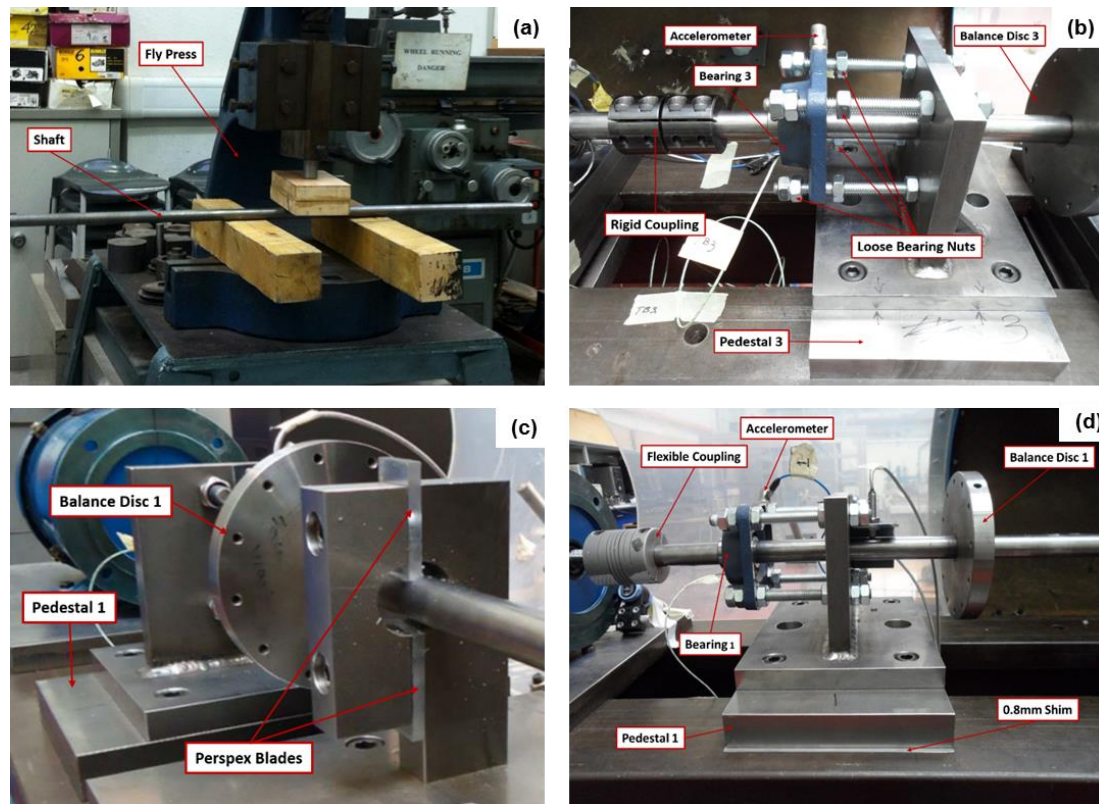
Figure 3. Faults simulation (**a**) rotor bow (**b**) bearing looseness (**c**) rotor rub (**d**) rotor misalignment.

Table 1. Description of experimentally simulated cases.

Case	Description	Abbreviation	Severity and Location
1	Healthy with residual misalignment	Healthy	The reference case likely contains some residual misalignment at the couplings, due to manufacturing and assembly imperfections
2	Rotor bow	Bow	3.4 mm run-out was created at the centre of the 1000 mm shaft.
3	Rotor misalignment	Mlign	0.8 mm mild steel shim beneath RHS of bearing 1 foundation
4	Bearing looseness	Loose	Loosening some of the bearing 3 threaded bar nuts
5	Rotor rub	Rub	Partial rub using two Perspex blades (TDC and BDC of 1000 mm shaft), 275 mm from bearing 1

4. Data Arrangement and Signal Processing Parameters

In this study, vibration data were collected under a total of 15 experimental scenarios. A scenario here represents a specific machine condition (e.g., rotor misalignment) at a specific machine speed (e.g., 40 Hz). Prior to applying any of the tools within the proposed hybrid framework, it was crucial to ensure that the measured vibration data under individual scenarios exhibit comparable signal processing characteristics such as number of data points, frequency resolution, sampling frequency, etc.

4.1. Data Arrangement for CCS Data Fusion

Data preparation for CCS commences with averaging. In the present work, an "average" refers to single complete CCS calculation, which is then converted into a sample with a number of features, for training, validating, or classification in the subsequent stage for machine learning classification. A two-stage overlap method was used for generating enough averages from the raw data (the higher the number of averages, the higher the similarities between the reconstructed and original signals). As shown in Figure 4, the initial stage of data preparation involves splitting the raw data into segments of 20,000 data points with an overlap of 80%. The next stage is to calculate the ith CCS average of the $[2(i-1)+1]$ to $[2(i-1)+10]$th segments from each bearing pedestal using Hanning window. The sampling rate used is $F_S = 10$ kHz and the frequency resolution is $df = 0.5$ Hz. The number of raw data points, as well as the number of averages generated for each case is listed in Table 2.
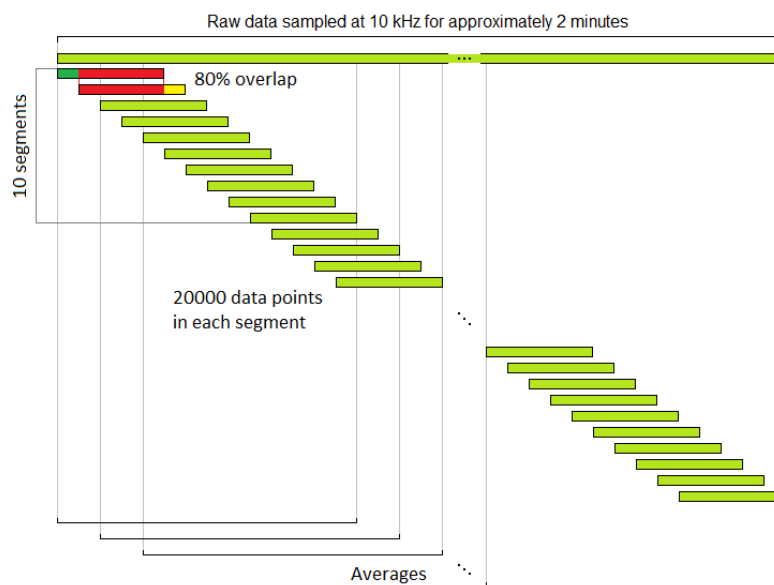

Figure 4. Two-stage overlap method for the preparation of datasets.

Table 2. The number of raw data points and averages in each case.

Speed	20Hz		30Hz		40Hz	
Case Code	Raw Data	Averages	Raw Data	Averages	Raw Data	Averages
Healthy	1,272,000	153	1,279,000	153	1,110,000	132
Bow	1,266,000	152	1,271,000	152	1,266,000	152
Loose	1,058,000	126	1,248,000	150	1,261,000	151
Mlign	996,000	118	927,000	109	928,000	110
Rub	1,259,000	151	1,271,000	152	994,000	118
Total	-	700	-	716	-	663

After the data preparation, Equations (1)–(7) were then use to generate the typical CCS shown in Figure 5. Each spectrum represents a specific case (i.e., Healthy, Bow, Loose, Mlign and Rub) at 2400 rpm (40 Hz). It is useful to reiterate that each CCS is a fusion of all vibration data from all 4 bearings, thereby providing a complete dynamics of the entire machine. During the faults diagnosis, only the first five harmonic components were considered, as it was adjudged that these would contain sufficient information to distinguish faults at these frequencies. Based on the types and magnitudes of the harmonics present alone, the Bow case exhibits much higher amplitudes at all harmonics than the other cases. In the Rub case however, the patterns of harmonics amplitudes can be described as the

reverse of the Bow. The higher harmonics (especially 3X, 4X and 5X) had the highest amplitudes as opposed to Bow which had a far superior 1X amplitude. The Loose and Bow cases also displayed similar 1X and 4X harmonic patterns. Unlike in the other cases whereby harmonic patterns sometimes appeared similar but amplitudes may differ and vice versa, Healthy and Mlign cases were immensely similar and almost undistinguishable on all counts. This phenomenon was anticipated because of the low severity of misalignment in Mlign case coupled with the inherent residual misalignment in the Healthy case. In order to enhance the ability to distinguish faults at all speeds, SE feature was applied here based on Equation (8). The computed SEs for the harmonics were then used as inputs to the PCA stage.
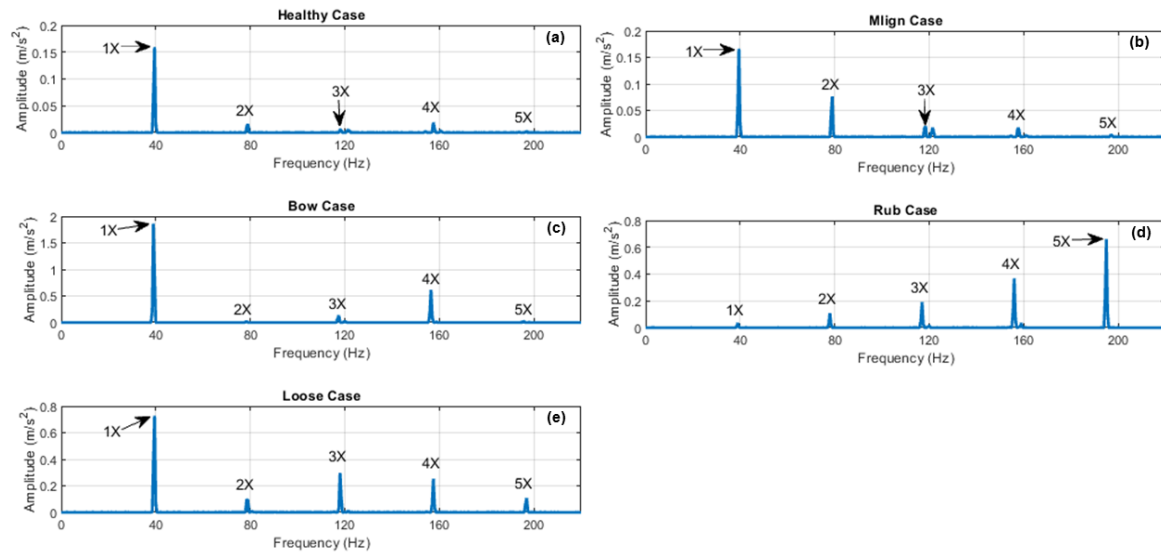

Figure 5. Typical CCS for all cases at 40Hz. (**a**) Healthy; (**b**) Mlign; (**c**) Bow; (**d**) Rub; (**e**) Loose.

4.2. Data Arrangement for Dimensionality Reduction

The SE of the CCS allows for the extraction of two features. Firstly the SE of 1X, 2X, 3X, 4X and 5X; and secondly the normalised harmonics is basically their ratios (e.g., 2X/1X, 3X/1X, 4X/1X and 5X/1X). By centralising and standardising each of these features according to Equations (9)–(12), 4 input data types (i.e., centralised SE, centralised ratio, standardised SE and standardised ratio), were generated for the PCA. The PCA converts these original features into the same number of principal components (PCs). The percentage of explained variance of each component indicates how much information the resultant PC holds. Typically, the first few PCs hold the vast majority of information. Based on this premise, the remaining PCs can, therefore, be discarded to reduce dimensionality of the dataset without necessarily compromising diagnosis quality. Table 3 provides a list of the percentages of explained variance for each PC for the different features at all speeds. Additionally, it is apparent that the combination of PCs 1 and 2 account for more than $\frac{3}{4}$ of the explained variance for each feature at all speeds.

Table 3. Percentage of explained variance by each PC.

Speed	Feature Extracted	PC1	PC2	PC3	PC4	PC5	PC 1&2
20 Hz	Centralised SE	98.8046	0.9327	0.1915	0.0697	0.0015	99.7373
	Centralised ratio	83.4228	11.3585	4.6620	0.5567	N/A	94.7813
	Standardised SE	47.3572	35.6890	13.2328	3.3719	0.3492	83.0462
	Standardised ratio	45.8536	26.7937	22.3915	4.9611	N/A	72.6473
30 Hz	Centralised SE	80.2678	14.5710	4.8346	0.3196	0.0069	94.8389
	Centralised ratio	86.4172	10.9302	2.3409	0.3117	N/A	97.3474
	Standardised SE	50.2902	32.7920	14.7593	2.1318	0.0267	83.0822
	Standardised ratio	57.7854	24.8192	15.9871	1.4083	N/A	82.6047
40 Hz	Centralised SE	85.3494	13.2058	0.7743	0.6089	0.0616	98.5552
	Centralised ratio	97.3550	2.3694	0.2433	0.0323	N/A	99.7244
	Standardised SE	55.1189	31.6004	10.7211	1.8511	0.7085	86.7193
	Standardised ratio	97.5927	1.5837	0.6559	0.1677	N/A	99.1764

Figure 6 compares the PCA results of all four features at 40Hz so as to examine the stability of SE with PCA alone, so as to justify the need for more advanced machine learning approaches. Of all the speeds considered, only 40 Hz offered reasonable separation for all cases with SE as a feature. This was perhaps due to the enhanced amplitude of vibration at this speed due to its closeness to the first natural frequency. Despite the reasonably good performance of SE at 40 Hz, the use of PCA alone is limited because it is unable to effectively integrate all speeds into a single map. This is due to disparity in amplitude of vibration at different speeds. Hence multiple charts would be required for each speed.
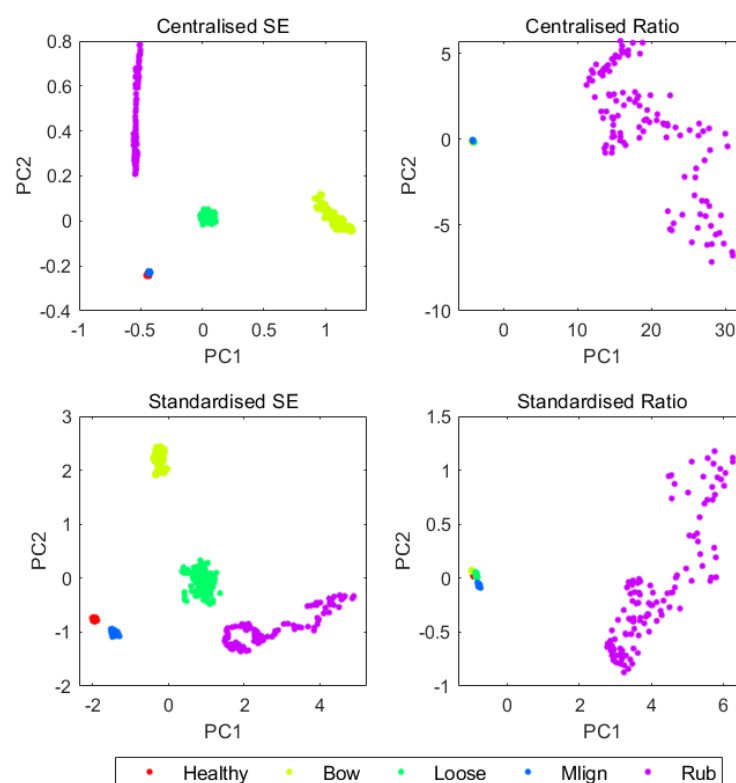
Figure 6. PCA results for four feature extraction methods at 40 Hz.

5. Data Analysis and Discussion of Results

The results will be initially examined in the context of the classification accuracy of the various machine learning techniques with regards to all features at all machine speeds. This comparison would give an indication of the most optimised technique-feature combinations under different scenarios of speeds and case. The remainder of the results analysis entails the assessment of visualisation strength of the techniques.

5.1. Accuracy Comparison

As previously highlighted, earlier combinations of CCS and PCA for rotor-related faults classification yielded encouraging outcomes, all such efforts were wholly based on manual human observations of perceived patterns of new unlabelled samples. The dominance of human intervention was adjudged a potential weak link especially from the viewpoint of repeatability, which could eventually jeopardise the quality diagnosis when dealing with critical industrial rotating machines. Therefore, implementing an approach that is capable of learning historical patterns and then using such knowledge to perform future classifications would highly enhance reliability. By using PCs 1 and 2 with labels as input data for training and validating the machine learning classifiers as well as K-fold cross-validation method (i.e., K = 10), Table 4 displays a comparison of their accuracies for all features.

Table 4. Accuracy of different machine learning classifiers (%).

Speed	Feature Extracted	k-NN k = 5	Naïve Bayes	Linear SVM	Gaussian SVM	ANN 2-10-5	ANN 2-20-5
20 Hz	Centralised SE	98.14	98.29	98.4	98.1	97.7	97.9
	Centralised Ratio	98.71	98.57	98.9	99.3	99.0	99.3
	Standardised SE	93.43	85.43	94.0	94.4	93.3	94.0
	Standardised Ratio	88.57	84.71	88.0	88.7	89.7	90.1
30 Hz	Centralised SE	94.13	94.41	96.1	94.7	94.1	95.9
	Centralised Ratio	96.79	96.51	97.9	98.6	83.5	83.9
	Standardised SE	99.16	98.46	98.9	99.4	98.3	98.6
	Standardised Ratio	100	100	100	100	100	100
40 Hz	Centralised SE	98.19	97.89	98.0	98.5	96.8	98.0
	Centralised Ratio	100	95.17	99.4	99.7	87.9	98.0
	Standardised SE	100	100	100	100	100	100
	Standardised Ratio	100	100	100	100	99.8	96.2

The results show that feature type has a significant effect on classification accuracy. The features which demonstrated the best performance at the PCA stage (i.e., centralised ratio at 20 Hz, standardised ratio at 30 Hz, and standardised SE at 40 Hz) also represent the best features for automatic fault classification for all the tested machine learning classifiers.

5.2. Visualised Decision Rules

In order to further enhance understanding of the classification mechanisms, the 2D decision rules presented in Figures 7–9 were developed. The trained classifiers divide the relevant regions of the PC1-PC2 plane into different sections. Each section of the region corresponds to individual faults. Furthermore, new unlabelled data points will then be assigned to sections that exhibit similar characteristics.

It is important to highlight that, under all three machine speeds examined, instances of severe over-fitting phenomena occurred for both Naïve Bayes and Gaussian SVM classifiers. The sections corresponding to Healthy and Mlign cases are quite minimal which could lead to incorrect classification of future data points. While a similar phenomenon was also evident for linear SVM under the Healthy case at 30 Hz, it demonstrated good outcomes at 20 Hz and 40 Hz speeds. K-NN and the two ANN classifiers displayed very good performance at all speeds. Judging by all-round performance, ANN

offered the best results under all scenarios. The ANN architecture adopted here was based on a two-layered backward propagation (BP) network that applied sigmoid hidden and soft max output neurons. This architecture has the capability to classify vectors arbitrarily, provided adequate neurons exist within its hidden layer. The 2-10-5 and 2-20-5 configurations visible in Table 4 respectively represent the input number of neurons for the hidden layer output whereby the two inputs represent PC1 and PC2. The datasets was divided into 70 (for training), 15 (for validation) and 15 (for testing).
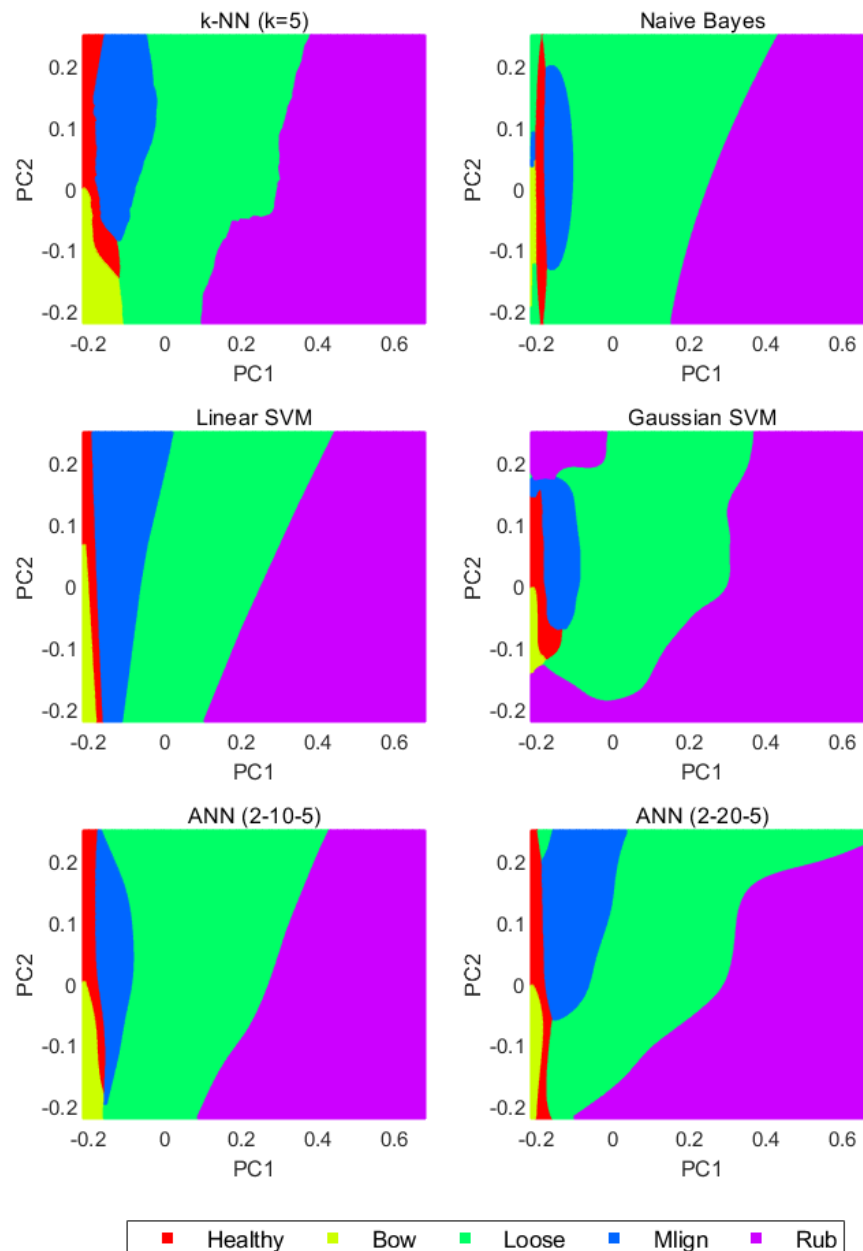
Figure 7. Visualised decision rules for the centralised ratio at 20 Hz.
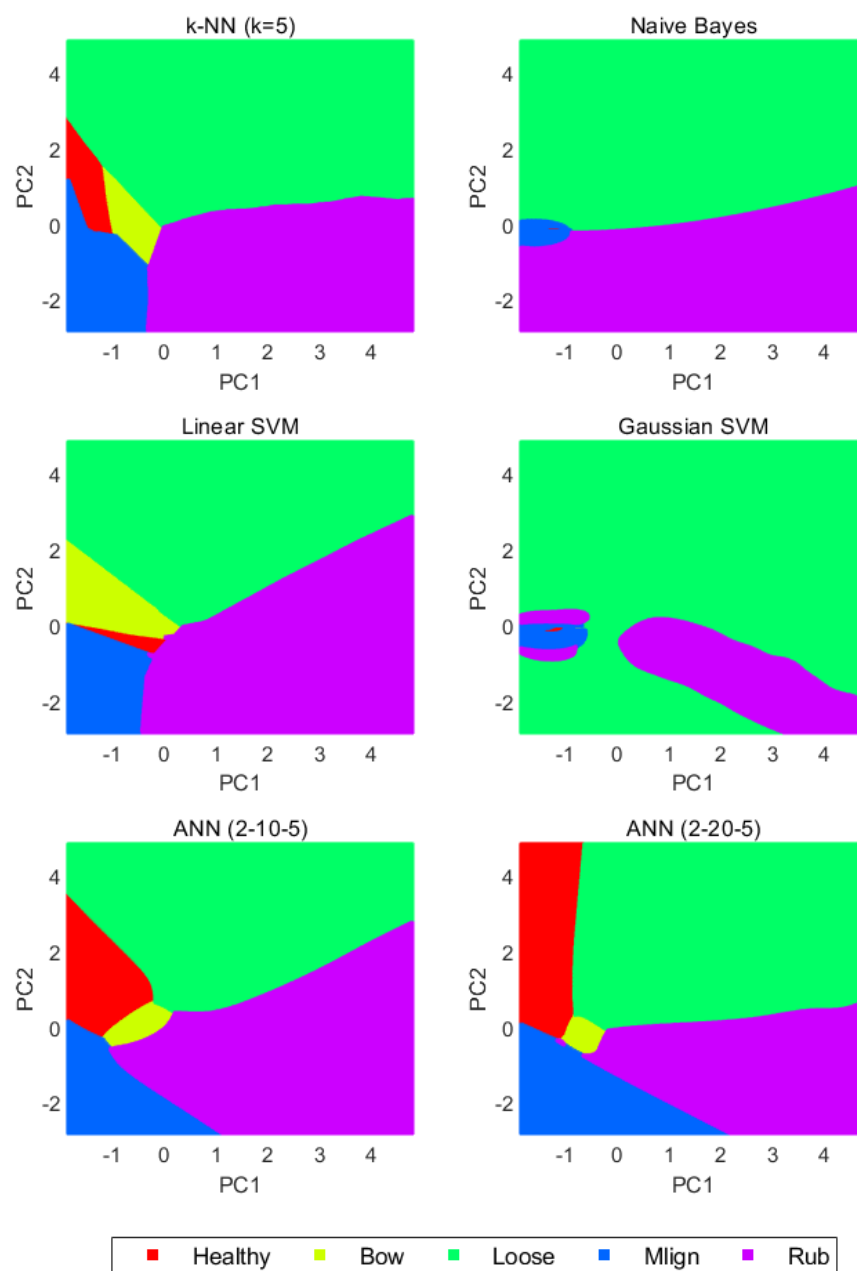
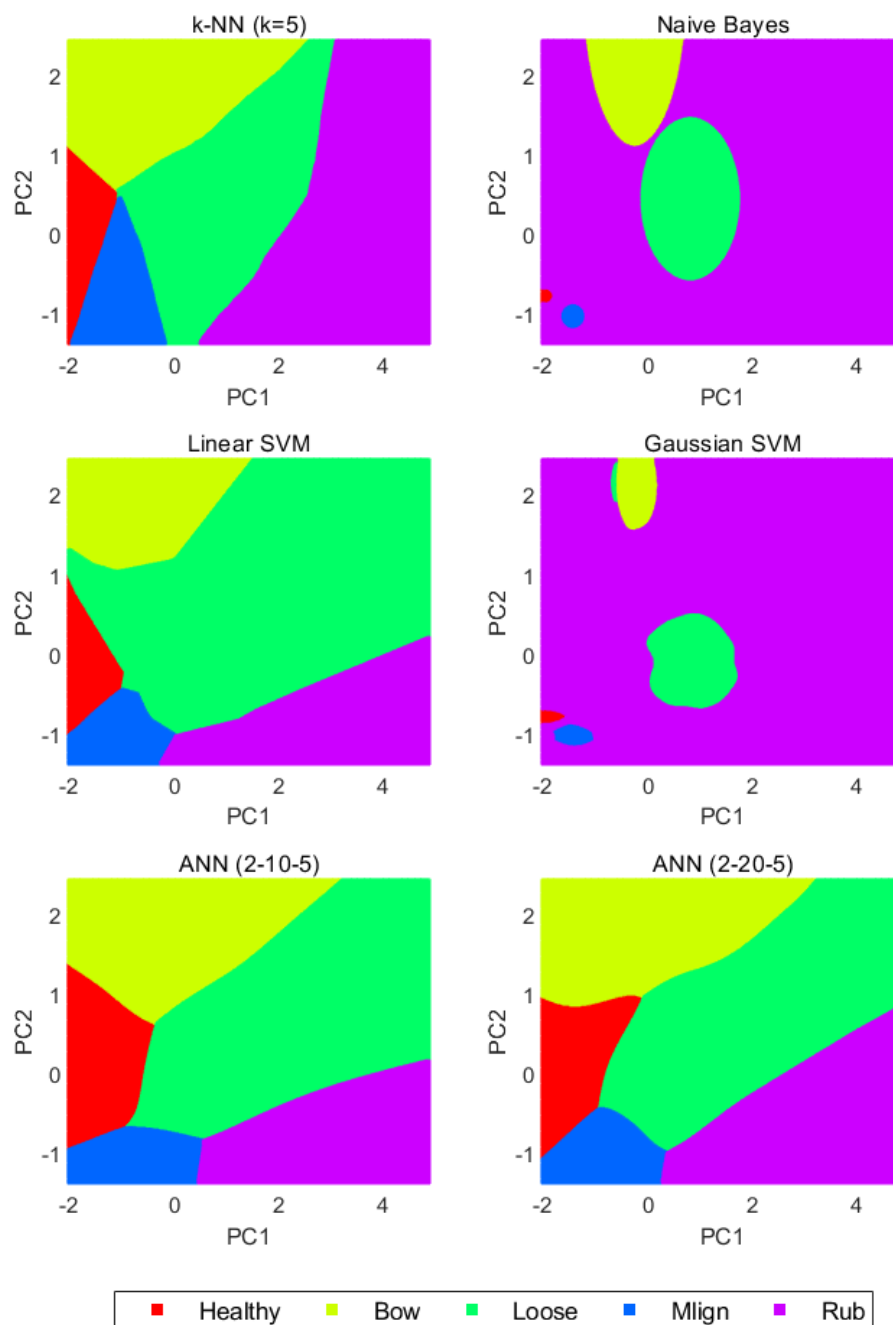**Figure 8.** Visualised decision rules for the standardised ratio at 30 Hz.

Figure 9. Visualised decision rules for the standardised SE at 40 Hz.

6. Conclusions and Future Work

Industrial rotating machines faults detection and classification using vibration-based approaches is widely applied across various industries over several decades. Despite the wealth of knowledge attributable to these classes of approaches, a significant amount of time is still attributable to their field-based implementation, which consequently impacts overall downtime and organisational profit margins. This is perhaps why recent research efforts have immensely focussed on the development of techniques that are capable of rationalising faults diagnosis data through data fusion. Such initiatives include the development of a composite spectrum (CCS) approach that uses a single spectrum to describe entire rotating machine dynamics, irrespective of the number of measurement locations. While initial findings from this approach yielded very encouraging outcomes, there still

exist significant opportunities for improvement especially with regard to minimising or eliminating human intervention/subjectivity so as to enhance repeatability.

Building upon the earlier CCS multi-sensor data fusion technique, the current study presents an automatic hybrid approach that uses spectrum energy (SE) as its feature. The proficiency of the new SE feature with various machine learning classifiers with regards to different machine scenarios of speed and conditions is presented. This multi-features/multi-classifiers/multi-scenarios approach allows for adequate performance comparison. It was observed that classifiers, such as Naïve Bayes and Gaussian support vector machine (SVM), displayed instances of severe overfitting, owing to the minimal nature of the sections corresponding to Healthy and Mlign cases which could lead to incorrect classification of future data points. While this situation was also apparent for linear SVM under the Healthy case at 30 Hz, its performance at other speeds was quite good. While K-NN classifiers displayed reasonably good performance at all speeds, ANN offered an all-round best set of results under all scenarios. The findings recorded in the current study demonstrate immense opportunity for automating data fusion-based faults classification of industrial rotating machines. While the current study only examined rotor-related faults, future studies are planned towards testing the presented approach on other classes of faults especially high frequency faults associated with bearings and gears.

References

1. Yunusa-Kaltungo, A.; Sinha, J. *Condition Monitoring: A Simple and Practical Approach*; LAP Lambert Academic Publishing: Saarbrücken, Germany, 2012.

2. Kostek, R.; Landowski, B.; Muślewski, Ł. Simulation of rolling bearing vibration in diagnostics. *J. Vibroengin.* **2015**, *17*, 4268–4278.

3. Saidi, L.; Ali, J.B.; Fnaiech, F. Application of higher order spectral features and support vector machines for bearing faults classification. *ISA Trans.* **2015**, *54*, 193–206. [CrossRef] [PubMed]

4. Liang, B.; Iwnicki, S.D.; Zhao, Y. Application of power spectrum, cepstrum, higher order spectrum and neural network analyses for induction motor fault diagnosis. *Mech. Syst. Signal Process.* **2013**, *39*, 342–360. [CrossRef]

5. Yunusa-Kaltungo, A.; Sinha, J.K.; Nembhard, A.D. Use of composite higher order spectra for faults diagnosis of rotating machines with different foundation flexibilities. *Measurement* **2015**, *70*, 47–61. [CrossRef]

6. Yunusa-Kaltungo, A.; Sinha, J.K. Faults diagnosis in rotating machines using higher order spectra. In *Proceedings of the ASME Turbo Expo 2014: Turbine Technical Conference and Exposition, Düsseldorf, Germany, 16–20 June 2014*; American Society of Mechanical Engineers Digital Collection; ASME: New York, NY, USA, 2014.

7. Arthur, N.; Penman, J. Induction machine condition monitoring with higher order spectra. *IEEE Trans. Ind. Electron.* **2000**, *47*, 1031–1041. [CrossRef]

8. Chow, T.W.; Tan, H.Z. HOS-based nonparametric and parametric methodologies for machine fault detection. *IEEE Trans. Ind. Electron.* **2000**, *47*, 1051–1059. [CrossRef]

9. McCormick, A.C.; Nandi, A.K. Bispectral and trispectral features for machine condition diagnosis. *IEEE Proc. Vis. Image Signal Process.* **1999**, *146*, 229–234. [CrossRef]

10. Yang, D.M.; Stronach, A.F.; MacConnell, P.; Penman, J. Third-order spectral techniques for the diagnosis of motor bearing condition using artificial neural networks. *Mech. Syst. Signal Process.* **2002**, *16*, 391–411. [CrossRef]

11. Yunusa-Kaltungo, A.; Sinha, J.K. Combined bispectrum and trispectrum for faults diagnosis in rotating machines. *Proc. Inst. Mech. Eng. Part O J. Risk Reliab.* **2014**, *228*, 419–428.

12. Luo, S.; Cheng, J.; Zeng, M.; Yang, Y. An intelligent fault diagnosis model for rotating machinery based on multi-scale higher order singular spectrum analysis and GA-VPMCD. *Measurement* **2016**, *87*, 38–50. [CrossRef]

13. Saidi, L.; Henao, H.; Fnaiech, F.; Capolino, G.A.; Cirrincione, G. Application of higher order spectral analysis for rotor broken bar detection in induction machines. In Proceedings of the 8th IEEE Symposium on Diagnostics for Electrical Machines, Power Electronics & Drives, Bologna, Italy, 5–8 September 2011; pp. 31–38.

14. Yunusa-Kaltungo, A.; Sinha, J.K.; Elbhbah, K. HOS analysis of measured vibration data on rotating machines with different simulated faults. In Proceedings of the Advances in Condition Monitoring of Machinery in Non-Stationary Operations, Lyon, France, 15–17 December 2014; Springer: Berlin/Heidelberg, Germany, 2016; pp. 81–89.

15. Saidi, L.; Ali, J.B.; Fnaiech, F. Bi-spectrum based-EMD applied to the non-stationary vibration signals for bearing faults diagnosis. *ISA Trans.* **2014**, *53*, 1650–1660. [CrossRef] [PubMed]

16. Yunusa-Kaltungo, A.; Sinha, J.K. Generic vibration-based faults identification approach for identical rotating machines installed on different foundations, VIRM 11-Vibrations in Rotating. *Machinery* **2016**, *11*, 499–510.

17. de la Rosa, J.J.; Muñoz, A.M. Higher-order cumulants and spectral kurtosis for early detection of subterranean termites. *Mech. Syst. Signal Process.* **2008**, *22*, 279–294. [CrossRef]

18. Saidi, L.; Fnaiech, F. Bearing defects decision making using higher order spectra features and support vector machines. In Proceedings of the IEEE 14th International Conference on Sciences and Techniques of Automatic Control & Computer Engineering-STA'2013, Sousse, Tunisia, 20–22 December 2013; pp. 419–424.

19. McCormick, A.C. Cyclostationary and Higher-order Statistical Signal Processing Algorithms for Machine Condition Monitoring. Ph.D. Thesis, University of Strathclyde, Glasgow, UK, 1998.

20. Howard, I.M. Higher-order spectral techniques for machine vibration condition monitoring. *Proc. Inst. Mech. Eng. Part G J. Aerosp. Eng.* **1997**, *211*, 211–219. [CrossRef]

21. Collis, W.B.; White, P.R.; Hammond, J.K. Higher-order spectra: The bispectrum and trispectrum. *Mech. Syst. Signal Pr.* **1998**, *12*, 375–394. [CrossRef]

22. Rivola, A.; White, P.R. Use of higher order spectra in condition monitoring: Simulation and experiments. In *Proceedings of the 1999 ASME Design Engineering Technical Conferences, Las Vegas, NV, USA, 12–15 September 1999*; Paper No. DETC99/VIB-8332; ASME: New York, NY, USA, 1999.

23. Yunusa-Kaltungo, A.; Sinha, J.K. A comparison of signal processing tools: Higher order spectra versus higher order coherences. *J. Vib. Eng. Technol.* **2015**, *3*, 461–472.

24. Elbhbah, K.; Sinha, J.K. Vibration-based condition monitoring of rotating machines using a machine composite spectrum. *J. Sound Vib.* **2013**, *332*, 2831–2845. [CrossRef]

25. Sinha, J.K.; Elbhbah, K. A future possibility of vibration based condition monitoring of rotating machines. *Mech. Syst. Signal Process.* **2013**, *34*, 231–240. [CrossRef]

26. Yunusa-Kaltungo, A.; Sinha, J.K.; Elbhbah, K. An improved data fusion technique for faults diagnosis in rotating machines. *Measurement* **2014**, *58*, 27–32. [CrossRef]

27. Yunusa-Kaltungo, A.; Sinha, J.K.; Nembhard, A.D. A novel fault diagnosis technique for enhancing maintenance and reliability of rotating machines. *Struct. Health Monit.* **2015**, *14*, 604–621. [CrossRef]

28. Nembhard, A.D.; Sinha, J.K. Unified Multi-speed analysis (UMA) for the condition monitoring of aero-engines. *Mech. Syst. Signal Process.* **2015**, *64*, 84–99. [CrossRef]

29. Jolliffe, I. *Principal Component Analysis*; Springer: Berlin/Heidelberg, Germany, 2011.

30. Nagbe, K.; Cugliari, J.; Jacques, J. Short-Term Electricity Demand Forecasting Using a Functional State Space Model. *Energies* **2018**, *11*, 1120. [CrossRef]

31. Shen, Y.; Abubakar, M.; Liu, H.; Hussain, F. Power Quality Disturbance Monitoring and Classification Based on Improved PCA and Convolution Neural Network for Wind-Grid Distribution Systems. *Energies* **2019**, *12*, 1280. [CrossRef]

32. Papaioannou, G.P.; Dikaiakos, C.; Dramountanis, A.; Papaioannou, P.G. Analysis and Modeling for Short-to Medium-Term Load Forecasting Using a Hybrid Manifold Learning Principal Component Model and Comparison with Classical Statistical Models (SARIMAX, Exponential Smoothing) and Artificial Intelligence Models (ANN, SVM): The Case of Greek Electricity Market. *Energies* **2016**, *9*, 635.

33. AbuBaker, M. Data Mining Applications in Understanding Electricity Consumers' Behavior: A Case Study of Tulkarm District, Palestine. *Energies* **2019**, *12*, 4287. [CrossRef]

34. Fei, W.; Moses, P. Fault Current Tracing and Identification via Machine Learning Considering Distributed Energy Resources in Distribution Networks. *Energies* **2019**, *12*, 4333. [CrossRef]

35. Zeng, B.; Guo, J.; Zhu, W.; Xiao, Z.; Yuan, F.; Huang, S. A Transformer Fault Diagnosis Model Based On Hybrid Grey Wolf Optimizer and LS-SVM. *Energies* **2019**, *12*, 4170. [CrossRef]

36. Feng, S.; Chen, J.; Tang, Y. Identification of Low Frequency Oscillations Based on Multidimensional Features and ReliefF-mRMR. *Energies* **2019**, *12*, 2762. [CrossRef]

37. Xiao, C.; Liu, Z.; Zhang, T.; Zhang, L. On Fault Prediction for Wind Turbine Pitch System Using Radar Chart and Support Vector Machine Approach. *Energies* **2019**, *12*, 2693. [CrossRef]

38. Liu, R.; Yang, B.; Zio, E.; Chen, X. Artificial intelligence for fault diagnosis of rotating machinery: A review. *Mech. Syst. Signal Process.* **2018**, *108*, 33–47. [CrossRef]

39. Pająk, M.; Muślewski, Ł.; Landowski, B.; Grządziela, A. Fuzzy identification of the reliability state of the mine detecting ship propulsion system. *Pol. Marit. Res.* **2019**, *26*, 55–64. [CrossRef]

40. Russell, S.J.; Norvig, P. *Artificial Intelligence: A Modern Approach*; Pearson Education Limited: Kuala Lumpur, Malaysia, 2016.

41. Yunusa-Kaltungo, A.; Kermani, M.M.; Labib, A. Investigation of critical failures using root cause analysis methods: Case study of ASH cement PLC. *Eng. Fail. Anal.* **2017**, *73*, 25–45. [CrossRef]

42. Reynolds, P.; Pavic, A. Quality assurance procedures for the modal testing of building floor structures. *Exp. Tech.* **2000**, *24*, 36–41. [CrossRef]

43. Ewins, D.J. Basics and state-of-the-art of modal testing. *Sadhana* **2000**, *25*, 207–220. [CrossRef]

44. Ewins, D.J. *Modal Testing: Theory, Practice and Application*, 2nd ed.; Research Studies Press: Hertfordshire, UK, 2000.

45. Wang, J.; Xu, M.; Zhang, C.; Huang, B.; Gu, F. Online Bearing Clearance Monitoring Based on an Accurate Vibration Analysis. *Energies* **2020**, *13*, 389. [CrossRef]

46. Aggarwal, A.; Strangas, E.G. Review of Detection Methods of Static Eccentricity for Interior Permanent Magnet Synchronous Machine. *Energies* **2019**, *12*, 4105. [CrossRef]

47. Li, X.; Elasha, F.; Shanbr, S.; Mba, D. Remaining Useful Life Prediction of Rolling Element Bearings Using Supervised Machine Learning. *Energies* **2019**, *12*, 2705. [CrossRef]

48. Salomon, C.P.; Ferreira, C.; Sant'Ana, W.C.; Lambert-Torres, G.; Borges da Silva, L.E.; Bonaldi, E.L.; de Oliveira, L.E.L.; Torres, B.S. A Study of Fault Diagnosis Based on Electrical Signature Analysis for Synchronous Generators Predictive Maintenance in Bulk Electric Systems. *Energies* **2019**, *12*, 1506. [CrossRef]

49. Shi, R.; Zhang, X. VSG-Based Dynamic Frequency Support Control for Autonomous PV–Diesel Microgrids. *Energies* **2018**, *11*, 1814. [CrossRef]

50. Caruso, M.; Di Tommaso, A.O.; Marignetti, F.; Miceli, R.; Ricco Galluzzo, G. A General Mathematical Formulation for Winding Layout Arrangement of Electrical Machines. *Energies* **2018**, *11*, 446. [CrossRef]