Knowledge-Based Systems 276 (2023) 110744

Contents lists available at ScienceDirect

Knowledge-Based Systems

journal homepage: www.elsevier.com/locate/knosys

On extraction, ranking and selection of data-driven and physics-informed features for bearing fault diagnostics

T. Haj Mohamad [a], A. Abbasi [b,*], K. Kappaganthu [c], C. Nataraj [b]

[a] Palo Alto Research Center (PARC, a Xerox Co.), 3333 Coyote Hill Road, Palo Alto, CA 94304, USA
[b] Villanova Center for Analytics of Dynamic Systems, 800 Lancaster Ave, Villanova, PA 19085, USA
[c] Engineering Manager, Google, CA, USA

ARTICLE INFO

Article history:
Received 10 December 2022
Received in revised form 30 May 2023
Accepted 19 June 2023
Available online 25 June 2023

Keywords:
Condition-based maintenance
Diagnostics
Rigid rotor
Rolling element bearings
Residual analysis
Statistical techniques
Cross-sample entropy technique
Proper orthogonal decomposition technique
Smooth orthogonal decomposition technique

ABSTRACT

Many traditional bearing fault detection techniques rely on pattern recognition using black box machine learning models, which lack generalizability to out of sample cases and are generally only effective for an *a priori* defined operating condition set. Therefore, the developed models are required to be retrained each time the operating condition changes because of the dependence of the dynamic response on them. The lack of effective techniques that can be applied and adapted to different operation domains of the system is a well-established need, which motivates the development of hybrid approaches which integrate both physics and data based techniques. This paper discusses the development of a hybrid physics-informed method that integrates information from physics and data for fault diagnostics of a rotor-bearing system. A physics-based model is used to simulate the dynamics of the system in a defect free condition and a machine fault simulator is used to generate experimental data. Physics-informed features are extracted by residual-based and cross-sample entropy techniques, while the data-based features are obtained using Proper and Smooth Orthogonal Decomposition techniques. Subsequently, these two collections of extracted features are fused into a hybrid feature set. Then, the features are ranked using the Shapley values ranking technique. Comparing the accuracy of classifications upon using the features of physics-informed, data-based and hybrid sets shows improved capability of the proposed framework to generalize diagnostics for various operating conditions.

1. Introduction

Optimum maintenance strategies decrease downtime and operating costs, and enhance reliability, availability, maintainability and safety of mechanical systems. Condition-Based Maintenance (CBM) can detect abnormality and plan maintenance actions efficiently by scheduling it only when required. Diagnostics and Prognostics are two fundamental concepts in CBM. Diagnostics is concerned with detection of abnormal conditions (fault detection), determining defective components (fault isolation), and clarifying the nature of the abnormality (fault identification). Prognostics focuses on abnormality progression prediction and future states estimation based on evaluating the deviation between current and expected normal states [1,2]. Industrial rotating machinery like turbines, generators, motors, pumps, compressors and gearboxes have a high risk of failure due to their severe working conditions. Hence, early and accurate diagnostics is highly important in these systems. The strong nonlinearity in such systems makes development of accurate fault diagnostics approaches a challenging problem [3–5].

Notable advances in this line were achieved by taking advantage of machine learning techniques. Their ability in mapping the information from the measurement space or feature space to fault space in systems with highly nonlinear dynamics could address this problem. However, in practice, it is not easy to implement machine learning techniques since purely experimental data does not contain insightful information of the system dynamics and specifically, the health condition. In addition, generality of this mapping is highly dependent on the data used, which can undermine its performance when dealing with out of sample scenarios [1,6,7]. In general, improving the measurement and feature extraction techniques can significantly improve the performance of machine learning based diagnostics approaches. A wide range of prior research has focused on developing data-driven feature extraction techniques and addressing domain dependency problem [8–12]. Notable feature extraction methods worth mentioning are *Proper Orthogonal Decomposition* (POD) and a recent method called *Smooth Orthogonal Decomposition* (SOD). POD, also known as *Principal Component Analysis* (PCA), singular value decomposition, or Karhunen–Loève expansion [13–15], is

* Corresponding author.
E-mail address: aabbasi@villanova.edu (A. Abbasi).

https://doi.org/10.1016/j.knosys.2023.110744
0950-7051/

a statistical method used for extracting features from multi-dimensional data containing information on system dynamics. SOD can be viewed as an extension to POD. SOD was at first developed by Chelidze and colleagues for damage identification and vibration analysis [16,17]. This method was used for other applications as well such as model reduction of PDEs [18], nonlinear control systems [19], and diagnostics of systems with variable operating conditions [20].

Another notable research focus in diagnostic techniques for engineering systems pertains to the challenges posed by limited training sample size, skewed data distribution, and domain shift. To address these challenges, researchers have proposed various methods such as transfer learning [21,22], one-class tensor hyperdisk [23], physics-informed gated recurrent graph attention unit network [24], and clustering graph convolutional network with multiple adversarial learning [25].

Also, deep learning-based methods have gained notable attention due to their adaptive capabilities and end-to-end convenience [26]. Yan et al. [27] developed a hybrid robust convolutional autoencoder for unsupervised anomaly detection of machine tools in the presence of noise. This method demonstrated a high capability for effectively fusing multi-sensor information and improving robustness to noise effects.

Although all these techniques could improve the performance of data based diagnostics, the lack of a mechanistic view in mapping from feature space to fault space still undermines their efficacy. Integrating physics could profoundly address this problem from different aspects such as domain modification, data generation, error minimization, outlier detection and mechanistic mapping. In other words, the principal idea here is integrating the underlying physics of dynamic response with faulty parameters into machine learning capitalizing on its ability to discover complex patterns. As a result, many studies have attempted to develop machine learning techniques that could leverage mathematical laws governed by underlying physics [28–32]. In one of the first attempts, the physics-based modeling ability to capture the underlying mechanisms of a rotating system was used to extract robust, interpretable, and informative features for diagnostics [8]. Also, the idea of using mechanistic view of physics for handling limited number of data samples and modifying the hyper parameter space was studied by Raissi and Karniadakis [33]. They obtained the solution of a partial differential equation (as the underlying physics) using two deep neural networks. The first neural network was designed to predict a general solution by considering all possible solutions while the second one was responsible for the spatiotemporal solution. In this architecture the first network provides prior knowledge that is used in the second network. The idea of using physics-informed recurrent neural networks was investigated by Yu et al. [30]. In a more applied study Chao et al. [32] developed a physics-informed deep learning approach for diagnostics.

Although integrating physics into data based diagnostic approaches could potentially improve their capability, feature ranking and selection are still of high importance. In other words, data-driven and physics-informed features can provide a wide range of feature candidates that should be ranked accurately and selected optimally for a specific diagnostic objective. Without such a ranking process the relative importance of features over each other is not clear. In addition, accurate feature ranking is required for optimum selection of a subset of features from the initial feature set since it can result in dimensionality reduction without compromising on accuracy [1].

The current study proposes a comprehensive hybrid framework for diagnostics of a rotating system, which includes feature extraction, ranking, and selection. This framework takes advantage of both experimental data and physics-based model to create a hybrid feature set. The data-based features are obtained using the POD and SOD feature extraction techniques, utilizing the experimental data generated by a fault simulator machine. The physics-informed features are extracted by statistical and cross-sample entropy techniques, and by considering the residual between the simulation of a physics-based model and experimental data. Adopting a hybrid approach amplifies the information content of inputs of the diagnostic mechanism. Additionally, the SHapley Additive exPlanations (SHAP) method is applied to the hybrid feature set, increasing the explainability of the diagnostic mechanism. The results emphasize the importance of hybridizing the model with experimental data to improve the effectiveness of diagnostics, enhance interpretability of the method, and enable generalization over a range of operational conditions.

2. Rotor-bearing system modeling

The rotor-bearing system includes a massless shaft and a rigid disk and is modeled by a four degree of freedom rotor. The rotor is supported by nonlinear rolling element bearings which are modeled using Hertzian contact forces. The bearing stiffness is assumed to be implicit in the bearing force. The rotor-bearing system and the defect free rolling element bearing schematic diagrams are shown in Fig. 1.

The rotor-bearing system has four degrees of freedom $q = [V\ W\ B\ \Gamma]^T$. V and W denote the displacement degrees of freedom in y and z directions, respectively. B and Γ denote the corresponding angular degrees of freedom and Ω stands for the rotating speed of the rotor. A rigid rotor with mass m, inertia I_D and polar moment of inertia I_p has the following forces acting on it: the bearing force (Q_b) and the unbalance force (Q_u). Using Lagrange equations, the equation of motion for the rotor-bearing system can be derived and are given by Eq. (1). The details of mathematical modeling can be found in [34,35].

$$M\ddot{q} + (D - \Omega G)\dot{q} = Q_b + Q_u \tag{1}$$

where M, D and G are the mass, damping and gyroscopic matrices given by Eq. (2).

$$M = \begin{bmatrix} m & 0 & 0 & 0 \\ 0 & m & 0 & 0 \\ 0 & 0 & I_D & 0 \\ 0 & 0 & 0 & I_D \end{bmatrix}, D = \begin{bmatrix} c & 0 & 0 & 0 \\ 0 & c & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$
$$G = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -I_p \\ 0 & 0 & I_p & 0 \end{bmatrix} \tag{2}$$

The unbalance force (Q_u) can be written as follows.

$$Q_u = m\Omega^2 \begin{bmatrix} \rho_{uy} \\ \rho_{uz} \\ 0 \\ 0 \end{bmatrix} + m\Omega^2 \begin{bmatrix} -\rho_{uz} \\ \rho_{uy} \\ 0 \\ 0 \end{bmatrix} \tag{3}$$

where $\rho_{uy} = \rho_u \cos(\phi_u)$, $\rho_{uz} = \rho_u \sin(\phi_u)$; ρ_u and ϕ_u denote the unbalance magnitude and phase of the rotor, respectively.

To determine the load on the rotor from the bearing, the net bearing forces are transferred from the bearing location to the rotor center of mass as shown in Eq. (4).

$$Q_b = \sum_{i=1}^{N_b} A_{bi}^T Q_{bi} \tag{4}$$

where N_b denotes the number of bearings and A_{bi} denotes the transformation matrix of the bearing at a distance a_i from the center of mass as shown in Eq. (5).

$$A_{bi} = \begin{bmatrix} 1 & 0 & 0 & -a_i \\ 0 & 1 & a_i & 0 \end{bmatrix} \tag{5}$$

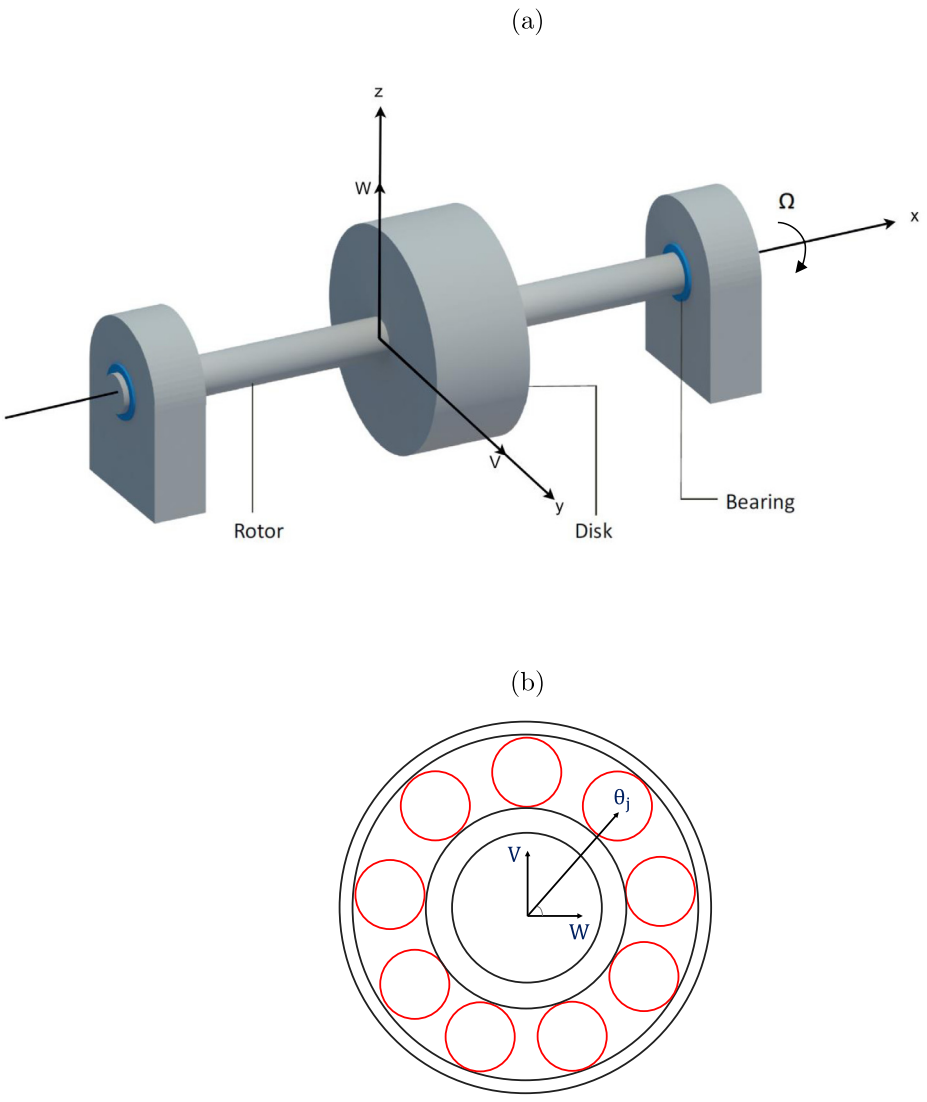


Fig. 1. (a) The schematic diagram of the rotor-bearing system, (b) rolling element bearing.

Also, Q_{bi} is the bearing force vector that can be written as $Q_{bi} = \left[F_{yi}\ F_{zi}\right]^T$, where F_{yi}, F_{zi} are the y and z components of the force exerted by the ith bearing on the rotor. These forces are the resultants of the nonlinear force exerted by each rolling element. The force exerted by the jth rolling element in the load zone is obtained from Hertzian contact stress [36,37] and is given by: $Q_j = -K_p\delta_j^n$, where K_p is the effective stiffness of the bearing; n is a constant dependent on the type of the bearing ($n = 1.5$ for ball bearings) and δ_j is the deflection of the jth rolling element. The effective stiffness of the bearing can be computed from the geometry of the bearing [36] and is given by Eq. (6).

$$K_p = \left[\left(\frac{1}{K_i}\right)^{1/n} + \left(\frac{1}{K_o}\right)^{1/n}\right]^{-n} \tag{6}$$

where K_i and K_o correspond to the inner and outer race stiffness, respectively. For steel balls and races, they are given by Eq. (7) [36].

$$K_{i,o} = 2.15 \times 10^5 \left(\sum \rho_{i,o}\right)^{-1/2} \left(\delta_{i,o}^*\right)^{-3/2} \tag{7}$$

where $\sum \rho$ is the sum of the curvatures and $\delta_{i,o}^*$ is a function of ρ that can be determined by interpolation from standard references such as [36,38]. Thus, if θ_j denotes the angular position of jth rolling element, the force exerted by the ith bearing on the shaft is given by Eq. (8).

$$Q_{bi} = \begin{bmatrix} \sum_{j=1}^{Nb} K_p\gamma_j\delta_j^n \cos(\theta_j) \\ \sum_{j=1}^{Nb} K_p\gamma_j\delta_j^n \sin(\theta_j) \end{bmatrix} \tag{8}$$

where,

$$\gamma_j = \begin{cases} 0 & \delta_j < 0 \\ 1 & \delta_j > 0 \end{cases} \tag{9}$$

and,

$$\delta_j = v\cos(\theta_j) + w\sin(\theta_j) \tag{10}$$

More details on calculating the effective stiffness of the bearing can be found in [39–41].

It should be noted that rolling element bearings are one of the principal sources of nonlinearity in a rotating system. Nonlinear restoring forces due to contact of curved surfaces, radial clearance between races, and defects are major contributing factors to the nonlinearity and complexity. Accordingly, the diagnostics problem and identification/isolation of faults under this condition are not straightforward. Traditional diagnostic methods generally do not consider the effects of nonlinearity in diagnostics reasoning,

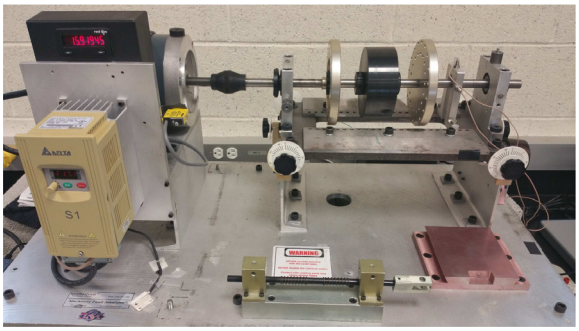


Fig. 2. Machine fault simulator set-up.

which results in lack of general applicability and ineffective prediction for complex engineering systems. That brings about safety concerns in addition to high maintenance costs. Therefore, incorporating the physics-based knowledge about nonlinear aspects of a dynamic system in diagnostics reasoning is critically important to ensure a safe and efficient operation [20,41–44].

3. Response simulation and unbalance identification

Experimental data is collected using a Machine Fault Simulator (MFS) shown in Fig. 2 [45]. This machine is a test rig with multiple components that can be handily assembled and disassembled to study a wide range of machinery defects. It basically consists of a motor-driven shaft mounted on two bearings. Shafts and bearings with different sizes and conditions can be used. GE/Bently Nevada 7200 series proximity probe sensors are used to measure the vertical and horizontal vibration displacement of the shaft close to the bearing housing. The data collected from the transducers are processed through the corresponding conditioning units and are digitized using a National Instruments NI USB 6363 data acquisition system.

Three bearing conditions are considered for experimental data collection including Defect Free (DF) bearings, bearings with Inner Race Defects (IRD), and bearings with Outer Race Defects (ORD). For each bearing condition, the data were collected at 19 different speeds ranging from 300 rpm to 3000 rpm with alternating increments of 120 and 180 rpm (e.g., 300, 420, 600, 720, 900, . . ., 3000), at a sampling rate of 10 kHz, and for a time length of 5 seconds. At each rotating speed, 10 sets of data were collected for a total of 570 sampled signals.

To simulate the DF response, the unbalance parameters, ρ_u and ϕ_u, are first estimated using the collected data. To achieve that, an optimization problem is defined using the residual or the difference between the model response and the data over all the operating speeds for one rotation as follows.

$$\min_{\rho_u,\ \phi_u} f(\rho_u,\ \phi_u) = \sum_{i=1}^{N_s} \frac{1}{N_i} [R_i(\rho_u,\ \phi_u)]^2 \tag{11}$$

where N_s is the number of rotating speeds and N_i is the number of data points in the signal at a particular rotating speed. The residual of the signal R_i is given by: $R_i = V_{e_i} - V_{m_i}(\rho_u,\ \phi_u)$, where V_{e_i} and V_{m_i} correspond to the experimental measurements and model simulation, respectively. The Least Squares method is used to estimate the unbalance parameters, which are then used to simulate the system response. Fig. 3(a) shows the comparison between the simulated DF model and a real system with a DF, an IRD and an ORD at 3000 rpm. The model and the DF system show similar patterns, whereas the IRD and ORD systems show different patterns from the model. This discrepancy in behavior carries over to a difference between the model and the data, which is evident in the residual of the signals as can be seen in Fig. 3(b). The residual has lower values when the system is DF, whereas high peaks appear for both IRD and ORD. The residuals of the systems in this figure are obtained at one operating condition. In general, they are complex and dependent on the rotational speed.

4. Feature extraction and ranking

An overview of the hybrid framework is shown in Fig. 4. First, the data-based features are extracted by applying proper and smooth orthogonal decomposition techniques on pure experimental data. Second, the experimental data along with the DF model are leveraged to generate a physics-informed feature set using residual and cross-sample entropy based features. The hybrid feature set is then formed using the data-based feature set and the physics informed feature set. The hybrid set is then ranked and the performance is calculated for multiple machine learning models that are built using the ranked hybrid feature set. The subset with the highest performance is then selected as an optimal feature set. This approach is discussed next in detail.

4.1. Data-based feature extraction

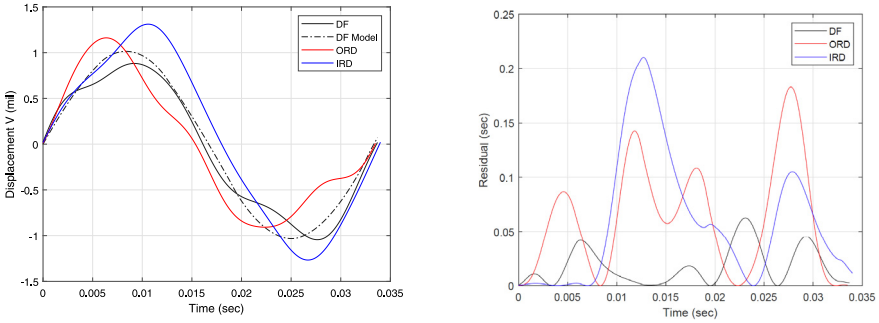
Smooth and proper orthogonal decomposition techniques are implemented to extract data-driven features from the experimental data. POD, also known as *Principal Component Analysis* (PCA), singular value decomposition, or Karhunen–Loève expansion [13,14], is a statistical method used for extracting features from multi-dimensional data containing information about a system. PCA takes a cloud of n-dimensional data and finds the principal direction in which data has maximum variance, with the condition that each direction is orthogonal to all other directions. SOD is another method for extracting features from data [16,17] and can extract information directly from the subspace on which the dynamical system evolves. SOD can be considered as an extension of POD, which alters the principal directions so that the evolution of data shows the maximum smoothness and condition of basis orthogonality is satisfied [46]. Hence, both these methods can address the domain dependency limitations of common diagnostic approaches. Their extraction procedure is discussed in detail.

The first feature set for the SOD technique consists of the angles lying between the *Smooth Orthogonal Modes* (SOMs) and the x-axis. *Smooth Orthogonal Values* (SOVs), which correspond to the degree of smoothness of the coordinates, are used as the second set of features. Finally, the components of SOMs are selected as the final feature set.

For the POD technique, the angles lying between the *Proper Orthogonal Modes* (POMs) and the x-axis are selected as the first feature set. Additionally, *proper orthogonal values* POVs, which correspond to the variances of the projected data on the POMs, are used as the second set of features. The final set of POD composes the components of POMs. Table 1 summarizes all the SOD-POD based features.

4.2. Physics-informed feature extraction

In this section, we utilize both the physics-based model and data to enhance the information content of the input domain of the diagnostic method by extracting physics-informed features. This hybridization aims to leverage the advantages of both methods and overcome their mutual limitations. At the implementation level, the correlation between the physics-based model and collected data is considered.



(a) (b)

Fig. 3. (a) Time series comparison for one rotation between the model and systems with DF, IRD and ORD, (b) Residue of DF, IRD and ORD signals.

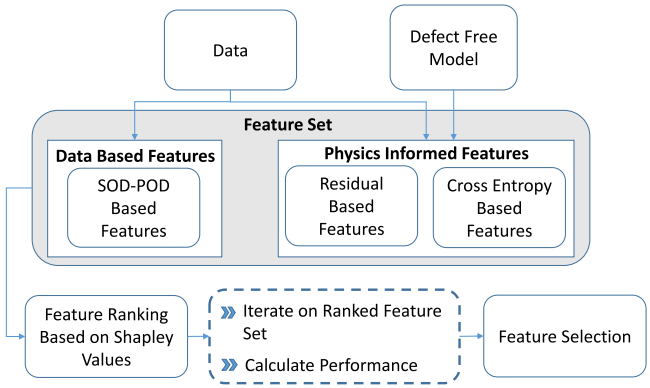


Fig. 4. Summary of data generation and feature selection steps.

Table 1
Smooth and proper orthogonal decomposition-based features.

Feat name	Label	Feat name	Label
Angles with SOMs	SOD θ_1 SOD θ_2	Angles with POMs	POD θ_1 POD θ_2
Smooth orthogonal value	SOV_1 SOV_2	Proper orthogonal value	POV_1 POV_2
Componentsof SOMs	SOM_{11} SOM_{12} SOM_{21} SOM_{22}	POMs components	POM_{11} POM_{12} POM_{21} POM_{22}

The correlation between the DF model simulation and the system response is quantified by the residual, which forms the basis for extracting residual-based features. These features aim to extract information from deviations between the physics-based model and experimental data, which have been proven to be informative about health variations [47].

Table 2 lists the residual-based features along with their labels that will be used in the next sections. The listed features are standard features, except the ENA4 and NVR2 features, which are proposed here as residual-based features for bearing diagnostics. ENA4 was originally developed by NASA [48–50] for gear diagnostics in order to detect a defect and when the system continues to react to this defect as it grows. ENA4 is the kurtosis of the residual signal normalized by the square of the variance of the residual signal of the DF system as defined by Eq. (12).

$$\text{ENA4} = \frac{\frac{1}{n}\sum_{i=1}^{N_i}(R_i - \bar{R})^4}{(\widetilde{M}_2)^2} \tag{12}$$

where $\widetilde{M}_2$ is the residual variance of the DF system. Additionally, NVR2 is a new feature proposed in this paper and is based on the variance of the residual signal normalized by the residual variance of the DF system as defined by Eq. (13).

Table 2
Residual-based statistical features.

Feature name	Label
Mean square error	MSE
Skewness	Skewness
Kurtosis	Kurtosis
ENA4	ENA4
Normalized variance of residual	NVR2

$$\text{NVR2} = \frac{\frac{1}{n}\sum_{i=1}^{N_i}(R_i - \bar{R})^2}{\widetilde{M}_2} \tag{13}$$

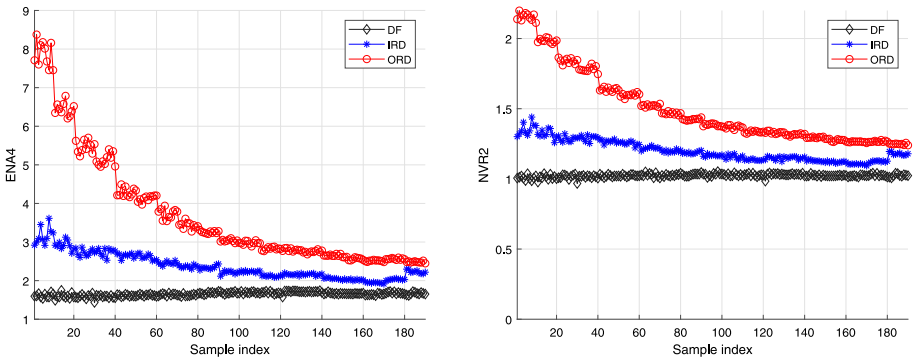
Fig. 5 shows ENR4 and NVR2 extracted at rotating speeds ranging from 300 rpm to 3000 rpm. Note that each set of 10 samples corresponds to a specific speed; for example, the first 10 samples correspond to 300 rpm, and so on. One observes that the ENA4 and NVR2 distinguish between all the bearing conditions at all speeds. The discrimination is the highest at lower speeds and features for the faulty conditions decrease with increasing speed. For both features, the faulty conditions have a higher feature value than the DF condition. In particular, the ORD condition has higher feature values than the IRD condition. In summary, NVR2 for DF condition maintains a relatively constant value equal to 1, and hence, ENA4 and NVR2 provide significant information about the bearing health status at various rotating speeds.

The second set of physics-informed features is extracted using the cross-sample entropy method. This method relies on the concept of cross-sample entropy, which is a measure of the non-linearity of a signal. The main objective is to compare different parts of the signal and measure their similarity using a tolerance parameter [51].

To use the cross-sample entropy method, the first step is to select a tolerance parameter r, which determines the maximum distance allowed between two points in the signal for them to be considered similar. The second step is to select a range of template lengths m_e, which are used to compare different parts of the signal [52,53]. In this method, the tolerance parameter (r) is used to find the number of matches at various template lengths (m_e). These matches are used to determine the probability of matches for various template lengths. Cross-sample entropy $CrsEn(.)$ is the negative natural logarithm of the conditional probability such that if two signals are within a distance of r for m_e points, then they are also within a distance of r for $m_e + 1$ points.

For completeness, the calculation of cross-sample entropy is provided as follows.

Consider signals $\bar{x} = [x(1)\ x(2)\ \ldots\ x(N_i - 1)]^T$ and $\bar{y} = [y(1)\ y(2)\ \ldots\ y(N_i - 1)]^T$ of length N_i. If m_e is an integer less than



(a) (b)

Fig. 5. ENA4 (a) and NVR2 (b) features for various bearing conditions at various rotating speeds ranging from 300 rpm to 3000 rpm.

N_i, then $(N_i - m + 1)$ sub-signals can be formed from each of the original signals. These are labeled $\bar{x}^i_{m_e}$, $\bar{y}^i_{m_e}$ and are given by the following equations.

$$\bar{x}^i_{m_e} = [x(i)\ x(i+1)\ x(i+2)\ \ldots\ x(i+m_e-1)]^T \tag{14}$$

$$\bar{y}^i_{m_e} = [y(i)\ y(i+1)\ y(i+2)\ \ldots\ y(i+m_e-1)]^T \tag{15}$$

If r is the acceptable tolerance, then two sub-signals $\bar{x}^i_{m_e}$ and $\bar{y}^j_{m_e}$ are considered to match if the difference $d[\bar{x}^i_{m_e},\ \bar{y}^j_{m_e}]$ given by the following equation, is less than r.

$$d[\bar{x}^i_{m_e},\ \bar{y}^j_{m_e}] = max(|\bar{x}^i_{m_e} - \bar{y}^j_{m_e}|) \tag{16}$$

Let $B_i(m_e)$ be $[N_i - m_e + 1]^{-1}$ times the number of sub-signals of length m_e that match $\bar{x}^i_{m_e}$, and $A_i(m_e)$ be $[N_i - m_e + 1]^{-1}$ times the number of length $(m_e + 1)$ that match $\bar{x}^i_{m_e+1}$. Then, $A(m_e)$ and $B(m_e)$ are defined as follows.

$$A(m_e) = \frac{1}{N_i - m_e} \sum_{i=1}^{N_i - m_e} A_i(m_e) \tag{17}$$

$$B(m_e) = \frac{1}{N_i - m_e} \sum_{i=1}^{N_i - m_e} B_i(m_e) \tag{18}$$

For a finite length time series, the cross-sample entropy is written as follows.

$$CrsEn(x, y, m_e, r) = -\ln A/B \tag{19}$$

The cross-sample entropy is evaluated between the simulated model and the system response for DF, IRD and ORD conditions. In calculating the cross-sample entropy, a choice of parameters such as tolerance (r) and sample length (m_e) needs to be made. Fig. 6 illustrates how the cross-sample entropy is evaluated for DF, ORD, and IRD at different template lengths (up to $m_e = 5$) and a tolerance of $r = 0.1$. It can be seen that the DF cross-sample entropy is lower than that of the defective conditions at various template lengths, which indicates higher similarity of the DF data to the model.

In order to ensure high performance of the algorithm, it is necessary to select those values of m_e and r for which the difference between cross-sample entropies is the highest, which will provide robustness against any noise and model uncertainty. The difference d_{crs} in cross-sample entropy between DF and defective is chosen as the selection criterion. The magnitude of the difference in cross-sample entropies can be obtained using Eq. (20) as:

$$d_{crs} = CrsEn(S_t, L_m) - CrsEn(S_t, S_m) \tag{20}$$

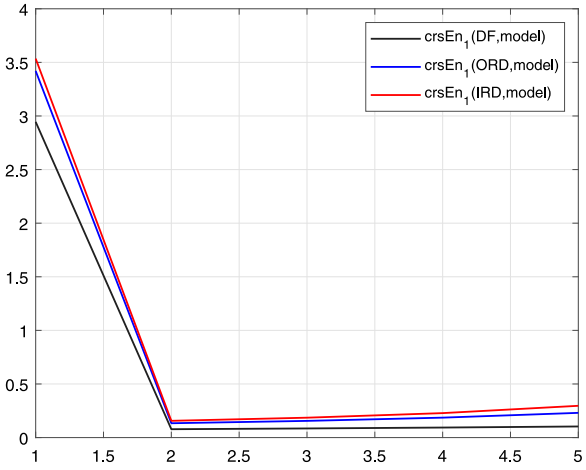


Fig. 6. Comparison of cross-sample entropy at template length (up to $m_e = 5$) and a tolerance $r = 0.1$.

Table 3
Cross-sample-based features.

Feature name	Label
Crs Samp En at $m_e = 5$ and $r = 0.05$	$crsEn_1(5, 0.05)$
Crs Samp En at $m_e = 5$ and $r = 0.05$	$crsEn_2(5, 0.05)$
Crs Samp En at $m_e = 5$ and $r = 0.1$	$crsEn_1(5, 0.1)$
Crs Samp En at $m_e = 5$ and $r = 0.1$	$crsEn_2(5, 0.1)$
Crs Samp En at $m_e = 7$ and $r = 0.06$	$crsEn_1(7, 0.06)$
Crs Samp En at $m_e = 7$ and $r = 0.06$	$crsEn_2(7, 0.06)$
Crs Samp En at $m_e = 10$ and $r = 0.1$	$crsEn_1(10, 0.1)$
Crs Samp En at $m_e = 10$ and $r = 0.1$	$crsEn_2(10, 0.1)$

This magnitude is calculated as m_e varies from 1 to 10 and r ranges from 0.01 to 0.2. S_t represents the test signal, S_m represents the DF simulated signal and L_m represents the defective simulated signal. The variation of d_{crs} over these ranges of m_e and r is shown in Fig. 7 for ORD condition. For the ORD, the difference is found to be greatest at $m_e = 5$ and $r = 0.05$ as shown in Fig. 7, which also corresponds to a high difference value for IRD (when the calculation were done separately). Using the cross-sample entropy between experimental data for various bearing conditions and the DF model of the system, cross-sample-based features that measure the similarity of the data to the DF models are extracted. Table 3 lists the cross-sample-based features that were considered, which are the first three features for each cross-sample entropy using ($m_e = 5$ and $r = 0.05$), ($m_e = 5$ and $r = 0.1$), ($m_e = 7$ and $r = 0.06$) and ($m_e = 10$ and $r = 0.1$), respectively.

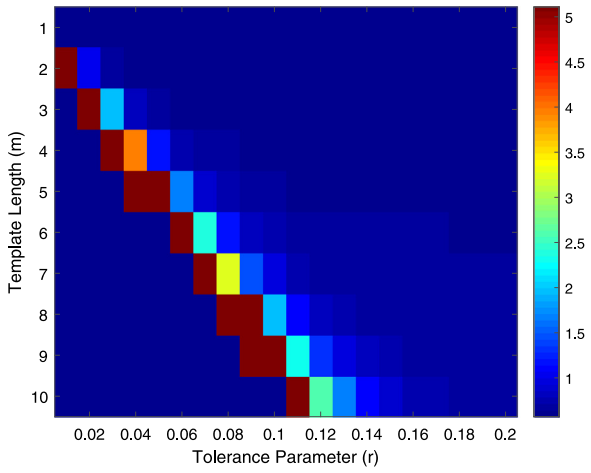


Fig. 7. Cross-sample entropy difference for ORD.

4.3. Feature ranking and selection using SHapley Additive exPlanations

Feature selection is the process of automatically selecting a subset of features that is most useful and most relevant to create the predictive model for the problem under investigation. Feature selection methods can be used to identify and remove irrelevant and redundant features. Moreover, feature selection helps to reduce overfitting and reduces computational costs. In this section, an optimal hybrid feature is achieved upon applying a feature selection technique to the two generated feature sets. In the current study, SHapley Additive exPlanations is implemented, which is a unified feature selection approach that is known as SHAP for short and was introduced by Lundberg et al. in [54].

A primary concept in SHapley Additive exPlanations is the calculation of Shapley values. Shapley values are a solution concept in cooperative game theory introduced by Shapley [55] and are based on the idea that the outcome of each possible combination of players should be considered to determine the importance of a single player. In our case, this corresponds to each possible combination of the features.

The Shapley value for a given input is a weighted average of the marginal contributions of that input to all possible subsets; its generalized formula is:

$$\phi_{\{i\}} = \sum_{S \subseteq F \setminus \{i\}} \frac{|S|!(|F| - |S| - 1)!}{|F|!} \rho_i$$

where the marginal contribution ρ can be computed using:

$$\rho_i = f_{S\cup\{i\}}(x_{S\cup\{i\}}) - f_S(x_S) \tag{21}$$

where, F is a set of all features and S is a subset. x_S represents the values of the input features in the set S, $f_S(x_S)$ is the prediction of the trained model without feature $\{i\}$ and $f_{S\cup\{i\}}(x_{S\cup\{i\}})$ is the prediction of another trained model with the feature $\{i\}$ present. SHAP values are used as a measure of feature importance. The SHAP value is calculated for each feature and across all observations in order to rank the features based on their average marginal contributions. The global feature importance values are calculated for the entire dataset by computing the average of the SHAP values for all the test data ($mean(|\text{SHAP}value|)$). By applying sampling approximation to Eq. (21), the need to retrain the model 2_F^N times is eliminated [54].

In the following section, the ability of SHAP values to reveal the importance of features is used for their ranking and then for obtaining a hybrid feature set.

4.4. Feature ranking and selection results

To examine the significance of hybridizing physics-based models and data in the diagnostics of the rotating system, we collected experimental and simulated data from three rotor speed configurations. The collected data will be used to evaluate the capability of interpolation and extrapolation/generalization. The speed configurations are as follows.

- Case (1): Shifted speed domain, where the training speeds range from 300 to 3000 rpm with increments of 300 rpm and the test speeds range from 420 to 2820 rpm with increments of 300 rpm.
- Case (2): Bounded speed domain, where the model is trained on two lower, middle and upper speeds, i.e., the training speed set includes 300, 420, 1200, 1320, 2820, 3000 rpm and the testing speeds are the remaining speeds.
- Case (3): Independent speed domain, where the training speed set includes all the speeds between 2100 and 3000, while the test speeds range from 300 to 1920 rpm.

A graphical representation of the considered speed domains is illustrated in Fig. 8. Configuration of speed samples from Case (1) to Case (3) is designed to increase the number of out of training samples. This step-wise reasoning provides for a comprehensive assessment of the interpolation and extrapolation/generalization capability of the proposed diagnostics algorithm.

At the first step, the ranking process is implemented on two feature sets: *data-based features* and *physics-informed features*. Then, these two feature sets are combined and go through the ranking process to create the third feature set: *hybrid features*.

For applying Shapley values ranking technique, TreeSHAP is implemented and random forests are selected as the prediction model. The main idea behind TreeSHAP is to compute the marginal contribution of each feature to the final predictions by passing through the decision tree and calculating the difference in predictions when a specific feature is included versus when it is excluded. This approach allows for determining the importance of each feature in outputs of the model and provides insights into the marginal contribution of each feature on model predictions. Random forests are an ensemble of decision trees based on a combination of multiple decision trees to improve predictive accuracy and generalization capability [56].

Multiple random forests with various tree numbers (up to 100 trees) are generated to best predict the bearing condition using grid search. Gini is selected as the impurity function, and square root of number of features considered at each split is selected for the Shapley values ranking.

4.4.1. Data-based feature ranking

The data-based feature set represented by SOD and POD-based features for the shifted speed domain case is ranked and shown in Fig. 9. The ranking figure displays the ranked features on the x-axis and the mean of the Shapley values on the y-axis for each condition, i.e., DF, IRD and ORD. The second component of the second smooth orthogonal mode shows a significantly higher importance measure compared with other features. This indicates that by comparing POD and SOD-based features, the greatest contributor to the prediction of DF, IRD and ORD is SOD-based features. The second ranked feature is the second angle of the proper orthogonal mode with the positive x-axis. The first component of the second smooth orthogonal mode is then ranked third. One observes that the importance of the ranked features decreases by increasing the feature rank. For the bounded speed case, the first three ranked features are similar to the ones shown in the shifted speed case; however, the rank of the first component of the second smooth mode goes up to the second ranked feature. The independent speed case shows a similar feature importance pattern except that the proper orthogonal value is ranked third.

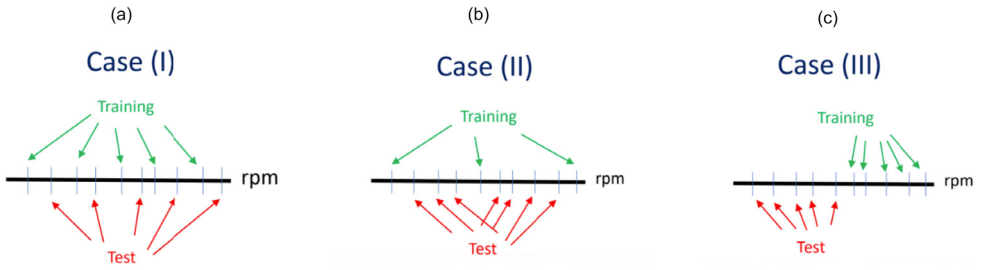


Fig. 8. (a) Shifted speed domain, (b) Bounded speed domain, (c) Independent speed domain.

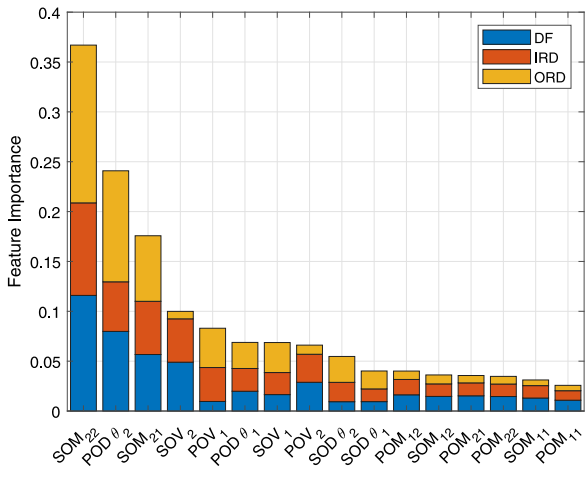


Fig. 9. Data-based features ranking for the shifted speed domain case.

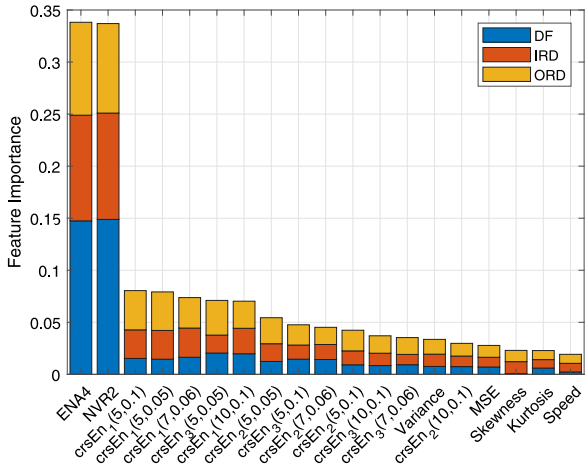


Fig. 10. Physics-informed feature ranking for the shifted speed domain case.

4.4.2. Physics-informed feature ranking

The ranked features for physics-informed features that consist of residual-based and cross-sample entropy-based features are shown in Fig. 10 for the shifted speed domain case. As can be seen, the most important two features are ENA4 and NVR2, which is also true for the other speed cases. *This is a notable generalization of the ENA4 applicability to various speed cases and other systems, i.e., bearing diagnostics.* Cross-sample entropy features then follow in the ranking.

4.4.3. Hybrid feature ranking

The same feature ranking technique is applied to the hybrid feature set obtained by combining both the physics-informed feature set and the data-based feature set. Fig. 11 presents this ranking results for the shifted speed domain case. The physics-informed features, NVR2 and ENA4, are ranked first for all speed cases with significant difference in importance compared to other extracted features. This indicates that NVR2 and ENA4 provide the most useful information about the bearing health conditions. The second smooth orthogonal mode and the POD angle follow next.

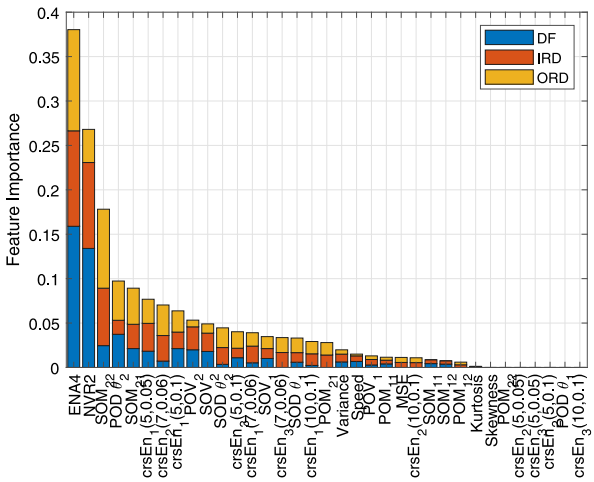


Fig. 11. Hybrid features ranking for the Shifted speed domain case.

5. Performance evaluation

In order to compare the performance of available feature sets, multiple random forest models were trained using the ranked physics-informed, data-based and hybrid features. The performance, presented by means of accuracy, is evaluated and shown in Fig. 12 for the three speed cases. For example, in Fig. 12, to evaluate the performance for physics-informed features at index 3, ENA4, NVR2 and $crsEn_1(5, 0.1)$ are used as the feature set.

For shifted and bounded speed cases, the best performance among all approaches is achieved with 100% accuracy when the first three features from the hybrid approach are used, which include two physics-informed features and one data-based feature. For the shifted speed case, the physics-informed feature set achieves 100% using all available features, whereas the data-based set achieves 82% using the first eight features. For the bounded speed case, the physics-informed feature set achieves an accuracy of 94% while the data-based feature set achieves 64% accuracy. For the independent speed case, the hybrid feature set achieves the highest accuracy of 83.6% with the first eight features, the physics-informed feature set achieves 84.4% accuracy using the first eleven features and the data-based achieve 44% accuracy using the first ten features. This quantitative analysis supports advantages of including physics-informed features in the hybrid feature set.

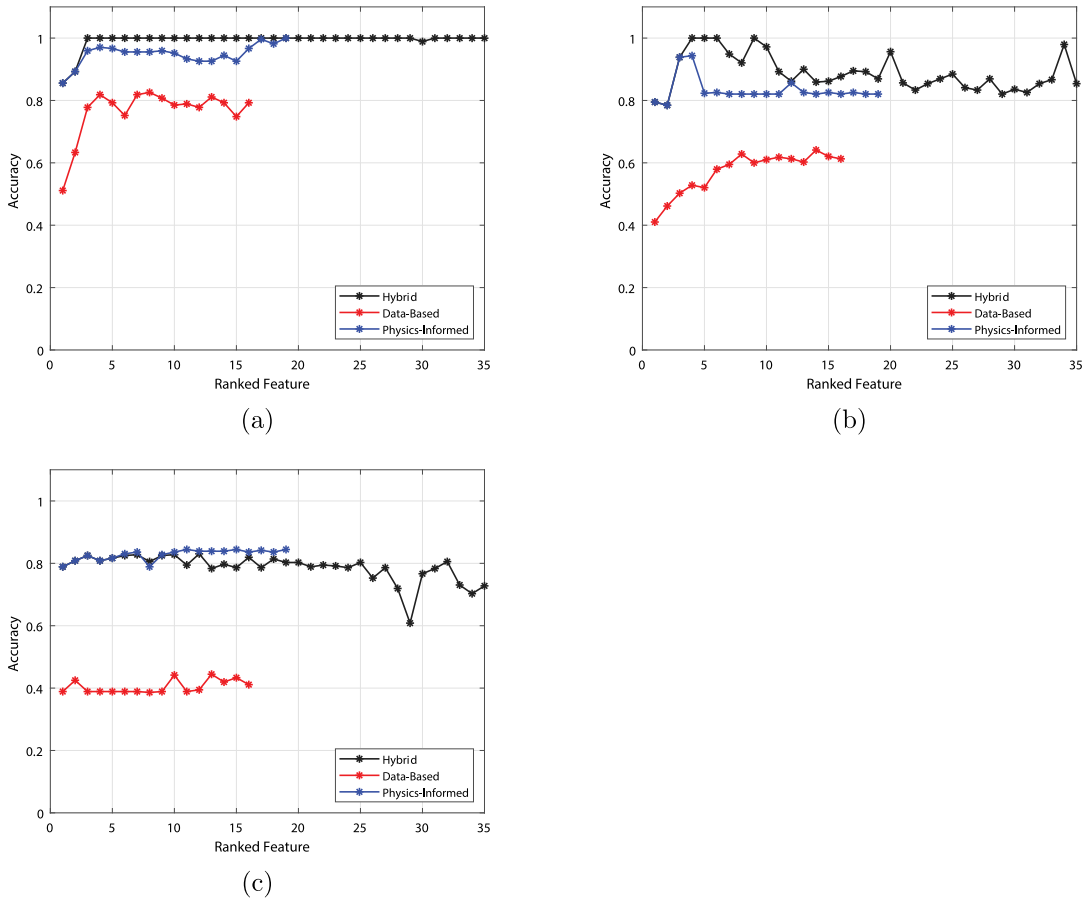


Fig. 12. The classification performance for various feature sets (a) Case 1, (b) Case 2, (c) Case 3.

It can be observed that the physics-informed feature set outperforms the data-based feature set for various speed cases. Additionally, the hybrid feature set outperforms both the physics-informed and data-based feature sets for the shifted and bounded speed cases, whereas the physics-informed feature set has a slightly better performance for the independent speed case. Also, the hybrid feature set is the fastest in reaching higher accuracy compared with the other sets.

6. Conclusion

The importance of reducing downtime and enhancing reliability, availability, maintainability, and safety of mechanical systems has led to emergence of many maintenance approaches. Condition-based maintenance seeks to propose maintenance actions based on condition monitoring data without interrupting the operation of the system and only when it is needed. The complexity of mapping the measurement space to the fault space is a challenging problem in diagnostics and particularly for nonlinear systems. The ability of machine learning techniques in creating such a mapping could be highly effective in resolving this problem. On the other hand, their high dependency on the availability and quality of data affects negatively their performance in practical applications. Hence, many processing techniques, such as scrubbing, fusion, transformation, feature extraction, feature ranking, and feature selection focus on resolving this problem. Most often in feature extraction techniques, informative characteristics of a signal are achieved by features that are highly indicative of faults. These features can be even more informative by taking advantage of the ability of physics in capturing mathematical abstraction between the parameter and fault spaces.

The concept of systematically integrating data-based and physics-based approaches has been explored in this study. In addition to improving the feature extraction, ranking and selection are significant for achieving an optimal feature set with lowest dimension and highest information. The proposed framework has been applied to the challenging, but important problem of fault classification of a nonlinear rotor-bearing system with a defective ball-bearing. The classes that have been considered are defect free, inner race defect and outer race defect [57–59]. Features are extracted using data and physics and then ranked to select the best feature set for fault classification under variable speed configurations (shifted, bounded and independent).

It can be observed that the physics-informed feature set outperforms the data-based feature set for various speed cases. This confirms the importance of taking advantage of the mechanistic approach in extracting features. The hybrid feature set outperforms both the physics-informed and data-based feature sets for the shifted and bounded speed cases. This establishes high generalizability of the hybrid approach across the operating conditions. Additionally, the hybrid feature set requires the least number of features in reaching a higher accuracy compared with the other feature sets. The result can be attributed to the importance of feature extraction and ranking in effective diagnostics.

For shifted and bounded speed cases, the best performance among all approaches is achieved with 100% accuracy when the first three features from the hybrid approach are used, which include two physics-informed features and one data-based feature. For the shifted speed case, the physics-informed feature set achieves 100% using all available features, whereas the data-based feature set achieves 82% using the first eight features.

For the bounded speed case, the physics-informed feature set achieves an accuracy of 94% while the data-based feature set achieves 64% accuracy. For the independent speed case, the hybrid feature set achieves the highest accuracy of 83.6% with the first eight features, the physics-informed feature set achieves 84.4% accuracy using the first eleven features and the data-based achieve 44% accuracy using the first ten features.

These quantitative results emphasize the generality of the proposed framework to obtain an optimal feature set for different operating speeds and demonstrates the significance of integrating information extracted from physics in order to develop a domain adaptable approach. The presented framework for feature extraction, ranking and selection for fault diagnostics can be generalized to many other applications which we are exploring. It can be seen that although integrating physics in the feature set can extend the input data domain, accurate ranking and selection are still highly important and the Shapley technique can accurately achieve this goal.

CRediT authorship contribution statement

T. Haj Mohamad: Conceptualization, Data curation, Formal analysis, Methodology, Software, Writing – original draft. **A. Abbasi:** Conceptualization, Data curation, Investigation, Software, Validation, Visualization, Writing – original draft, Writing – review & editing. **K. Kappaganthu:** Supervision, Writing – review & editing. **C. Nataraj:** Conceptualization, Funding acquisition, Investigation, Methodology, Project administration, Resources, Supervision.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

This project was supported by research grants from the Office of Naval Research, USA (Grant No. N00014-19-1-2070 and N00014-22-1-2480). We are grateful to ONR and Capt. Lynn Petersen, the program manager for recognizing the importance of this work.

References

[1] A.K.S. Jardine, D. Lin, D. Banjevic, A review on machinery diagnostics and prognostics implementing condition-based maintenance, Mech. Syst. Signal Process. 20 (7) (2006) 1483–1510.
[2] R.B. Randall, Vibration-Based Condition Monitoring: Industrial, Aerospace and Automotive Applications, John Wiley & Sons, 2011.
[3] R.B. Randall, J. Antoni, Rolling element bearing diagnostics -a tutorial, Mech. Syst. Signal Process. 25 (2) (2011) 485–520.
[4] Y. Lei, J. Lin, Z. He, M.J. Zuo, A review on empirical mode decomposition in fault diagnosis of rotating machinery, Mech. Syst. Signal Process. 35 (1-2) (2013) 108–126.
[5] W.A. Smith, R.B. Randall, Rolling element bearing diagnostics using the case western reserve university data: A benchmark study, Mech. Syst. Signal Process. 64 (2015) 100–131.
[6] M. He, D. He, Deep learning based approach for bearing fault diagnosis, IEEE Trans. Ind. Appl. 53 (3) (2017) 3057–3065.
[7] S. Zhang, S. Zhang, B. Wang, T.G. Habetler, Deep learning algorithms for bearing fault diagnostics, a comprehensive review, IEEE Access 8 (2020) 29857–29881.
[8] K. Kappaganthu, C. Nataraj, Feature selection for fault detection in rolling element bearings using mutual information, ASME J. Vib. Acoust. 133 (6) (2011).
[9] K. Kappaganthu, C. Nataraj, Feature selection for fault detection in rolling element bearings using mutual information, J. Vib. Acoust. 133 (6) (2011).
[10] M.D. Prieto, G. Cirrincione, A.G. Espinosa, J.A. Ortega, H. Henao, Bearing fault detection by a novel condition-monitoring scheme based on statistical-time features and neural networks, IEEE Trans. Ind. Electron. 60 (8) (2013) 3398–3407.
[11] R. Bajric, N. Zuber, S. Isic, Recent advances in vibration signal processing techniques for gear fault detection-a review, Appl. Mech. Mater. 430 (2013) 78–83.
[12] H. Gao, L. Liang, X. Chen, G. Xu, Feature extraction and recognition for rolling element bearing fault utilizing short-time Fourier transform and non-negative matrix factorization, Chin. J. Mech. Eng. 28 (1) (2015) 96–105.
[13] T.R. Smith, J. Moehlis, P. Holmes, Low-dimensional modelling of turbulence using the proper orthogonal decomposition: a tutorial, Nonlinear Dynam. 41 (1-3) (2005) 275–307.
[14] I. Georgiou, Advanced proper orthogonal decomposition tools: using reduced order models to identify normal modes of vibration and slow invariant manifolds in the dynamics of planar nonlinear rods, Nonlinear Dynam. 41 (1-3) (2005) 69–110.
[15] Z. Liu, T.H. Mohamad, S. Ilbeigi, C. Nataraj, Advances in Nonlinear Dynamics, Springer, 2022, pp. 451–461.
[16] D. Chelidze, Identifying robust subspaces for dynamically consistent reduced-order models, in: Conference Proceedings of the Society for Experimental Mechanics Series, Vol. 2, Springer, New York LLC, 2014, pp. 123–130.
[17] D. Chelidze, W. Zhou, Smooth orthogonal decomposition-based vibration mode identification, J. Sound Vib. 292 (3-5) (2006) 461–473.
[18] J.A. Atwell, J.T. Borggaard, B.B. King, Reduced order controllers for Burgers' equation with a nonlinear observer, Appl. Math. Comput. Sci. 11 (6) (2001) 1311–1330.
[19] S. Ilbeigi, D. Chelidze, A new approach to model reduction of nonlinear control systems using smooth orthogonal decomposition, Internat. J. Robust Nonlinear Control 28 (15) (2018) 4367–4381.
[20] T.H. Mohamad, S. Ilbeigi, C. Nataraj, Proper and smooth orthogonal decompositions for detection of inner race defects in rolling element bearings with variable rotational speeds, in: Nonlinear Dynamics of Structures, Systems and Devices, Springer, 2020, pp. 493–501.
[21] S. Simani, S. Farsoni, P. Castaldi, Supervisory control and data acquisition for fault diagnosis of wind turbines via deep transfer learning, Energies 16 (9) (2023) 3644.
[22] Y. Liu, G. Yang, S. Qiao, M. Liu, L. Qu, N. Han, T. Wu, G. Yuan, Y. Peng, Imbalanced data classification: Using transfer learning and active sampling, Eng. Appl. Artif. Intell. 117 (2023) 105621.
[23] Z. He, Y. Zeng, H. Shao, H. Hu, X. Xu, Novel motor fault detection scheme based on one-class tensor hyperdisk, Knowl.-Based Syst. 262 (2023) 110259.
[24] W. Wu, C. Song, J. Zhao, Z. Xu, Physics-informed gated recurrent graph attention unit network for anomaly detection in industrial cyber-physical systems, Inform. Sci. 629 (2023) 618–633.
[25] H. Wen, W. Guo, X. Li, A novel deep clustering network using multi-representation autoencoder and adversarial learning for large cross-domain fault diagnosis of rolling bearings, Expert Syst. Appl. 225 (2023) 120066.
[26] J. Cen, Z. Yang, X. Liu, J. Xiong, H. Chen, A review of data-driven machinery fault diagnosis using machine learning algorithms, J. Vib. Eng. Technol. (2022).
[27] S. Yan, H. Shao, Y. Xiao, B. Liu, J. Wan, Hybrid robust convolutional autoencoder for unsupervised anomaly detection of machine tools under noises, Robot. Comput.-Integr. Manuf. 79 (2023) 102441.
[28] A.P. Singh, S. Medida, K. Duraisamy, Machine-learning-augmented predictive modeling of turbulent separated flows over airfoils, AIAA J. 55 (7) (2017) 2215–2227.
[29] M. Raissi, G.E. Karniadakis, Hidden physics models: Machine learning of nonlinear partial differential equations, J. Comput. Phys. 357 (2018) 125–141.
[30] Y. Yu, H. Yao, Y. Liu, Physics-based learning for aircraft dynamics simulation, in: PHM Society Conference, Vol. 10, 2018.
[31] J.N. Kani, A.H. Elsheikh, Reduced-order modeling of subsurface multi-phase flow models using deep residual recurrent neural networks, Transp. Porous Media 126 (3) (2019) 713–741.
[32] A. Dourado, F.A. Viana, Physics-informed neural networks for missing physics estimation in cumulative damage models: A case study in corrosion fatigue, J. Comput. Inf. Sci. Eng. 20 (6) (2020).
[33] M. Raissi, Deep hidden physics models: Deep learning of nonlinear partial differential equations, J. Mach. Learn. Res. 19 (1) (2018) 932–955.
[34] K. Kappaganthu, An Integrative Approach for Machinery Diagnostics, (Ph.D. thesis), 2010.

[35] A. Abbasi, F. Nazari, C. Nataraj, Adaptive modeling of vibrations and structural fatigue for analyzing crack propagation in a rotating system, J. Sound Vib. (2022).
[36] T.A. Harris, Rolling Bearing Analysis, fourth ed., Wiley-Interscience, 2002.
[37] R. Pandiyarajan, M. Starvin, K. Ganesh, Contact stress distribution of large diameter ball bearing using hertzian elliptical contact theory, Procedia Eng. 38 (2012) 264–269.
[38] S.P. Harsha, K. Sandeep, R. Prakash, Non-linear dynamic behaviors of rolling element bearings due to surface waviness, J. Sound Vib. 272 (3–5) (2004) 557–580.
[39] C. Nataraj, R.G. Pietrusko, Dynamic response of rigid rotors supported on rolling element bearings with an outer raceway defect, in: Proceedings of the ASME International Design Engineering Technical Conferences and Computers and Information in Engineering Conference - DETC2005, Vol. 1 B, New York, NY,10016-5990, United States, 2005, pp. 1249–1261.
[40] K. Kappaganthu, C. Nataraj, B. Samanta, Feature selection for bearing fault detection based on mutual information, in: K. Gupta (Ed.), IUTAM Symposium on Emerging Trends in Rotor Dynamics, Springer Netherlands, Dordrecht, 2011, pp. 523–533.
[41] K. Kappaganthu, C. Nataraj, K. Kappaganthu, C. Nataraj, Nonlinear modeling and analysis of a rolling element bearing with a clearance, Commun. Nonlinear Sci. Numer. Simul. 16 (10) (2011) 4134–4145.
[42] C. Nataraj, S.P. Harsha, Nonlinear Vibration Analysis of an Unbalanced Rotor on Rolling Element Bearing Due To Cage Run-Out, in: Transaction of Nonlinear Science and Complexity (NSC), Nonlinear Science and Complexity, World Scientific, 2006.
[43] K. Kappaganthu, C. Nataraj, Mutual information based feature selection from data driven and model based techniques for fault detection in rolling element bearings, Int. J. Mech. Sci. (2011) 941–953.
[44] T.H. Mohamad, C.A.K. Kwuimy, C. Nataraj, Discrimination of multiple faults in bearings using density-based orthogonal functions of the time response, in: ASME 2017 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, American Society of Mechanical Engineers, Cleveland, OH, 2017.
[45] T.H. Mohamad, M. Samadani, C. Nataraj, Rolling element bearing diagnostics using extended phase space topology, J. Vib. Acoust. 140 (6) (2018).
[46] Z. Liu, T.H. Mohamad, S. Ilbeigi, C. Nataraj, Early detection of cracks in a gear-train system using proper and smooth orthogonal decompositions, in: Advances in Nonlinear Dynamics, Springer International Publishing, 2022.
[47] X. Dai, Z. Gao, From model, signal to knowledge: A data-driven perspective of fault detection and diagnosis, IEEE Trans. Ind. Inform. 9 (4) (2013) 2226–2238.
[48] J.J. Zakrajsek, D.P. Townsend, H.J. Decker, An Analysis of Gear Fault Detection Methods As Applied To Pitting Fatigue Failure Data, Technical Report, National Aeronautics and Space Administration (NASA), 1993.
[49] K. McClintic, M. Lebold, K. Maynard, C. Byington, R. Campbell, Residual and difference feature analysis with transitional gearbox data, in: Proceedings of the 54th Meeting of the Society for Machinery Failure Prevention Technology, 2000, 1–4.
[50] T. Sreenuch, A. Tsourdos, I.K. Jennions, Distributed embedded condition monitoring systems based on osa-cbm standard, Comput. Stand. Interfaces 35 (2) (2013) 238–246.
[51] J.S. Richman, J.R. Moorman, Physiological time-series analysis using approximate entropy and sample entropy, Am. J. Physiol.-Heart and Circ. Physiol. 278 (6) (2000) H2039–H2049.
[52] Y. Zhang, Y. Li, L. Kong, Q. Niu, Y. Bai, Improved dbscan spindle bearing condition monitoring method based on kurtosis and sample entropy, Machines 10 (5) (2022) 363.
[53] H. Donajkowski, S. Leyasi, G. Mellos, C.R. Farrar, A. Scheinker, J.-S. Pei, N.A. Lieven, Comparison of Complexity Measures for Structural Health Monitoring, Springer, 2020, pp. 27–39.
[54] S.M. Lundberg, S.-I. Lee, A unified approach to interpreting model predictions, Adv. Neural Inf. Process. Syst. 30 (2017).
[55] R. Serrano, Lloyd shapley's matching and game theory, Scand. J. Econ. 115 (3) (2013).
[56] M. Kamp, I. Koprinska, A. Bibal, T. Bouadi, B. Frénay, L. Galárraga, J. Oramas, L. Adilova, Y. Krishnamurthy, B. Kang, et al., Machine learning and principles and practice of knowledge discovery in databases: International workshops of ECML pkdd 2021, virtual event, september 13-17, 2021, in: Proceedings, Part II, Springer Nature, 2022.
[57] A. Morhain, D. Mba, Bearing defect diagnosis and acoustic emission, Proc. Inst. Mech. Eng. 217 (4) (2003) 257–272.
[58] J.R. Stack, T.G. Habetler, R.G. Harley, Fault-signature modeling and detection of inner-race bearing faults, IEEE Trans. Ind. Appl. 42 (1) (2006) 61–68.
[59] R. Yang, Y. Jin, L. Hou, Y. Chen, Study for ball bearing outer race characteristic defect frequency based on nonlinear dynamics analysis, Nonlinear Dynam. 90 (2) (2017) 781–796.