Check for updates

Role of artificial intelligence in rotor fault diagnosis: a comprehensive review

Aneesh G. Nath[1] (ORCID) **· Sandeep S. Udmale[2] · Sanjay Kumar Singh[1]**

Abstract

Artificial intelligence (AI)-based rotor fault diagnosis (RFD) poses a variety of challenges to the prognostics and health management (PHM) of the Industry 4.0 revolution. Rotor faults have drawn more attention from the AI research community in terms of utilizing fault-specific characteristics in its feature engineering, compared to any other rotating machinery faults. While the rotor faults, specifically structural rotor faults (SRF), have proven to be the root cause of most of the rotating machinery issues, the research in this field largely revolves around bearing and gear faults. Within this scenario, this paper is the first of its kind to attempt to review and define the role of AI in RFD and provides an all-encompassing review of rotor faults for the researchers and academics. In addition, this study is unique in three ways: (i) it emphasizes the use of fault-specific characteristic features with AI, (ii) it is grounded in fault-wise analysis rather than component-wise analysis with appropriate fault categorization, and (iii) it portrays the current research and analysis in accordance with different phases of an AI-based RFD framework. Finally, the section on future research directions is aimed at bridging the gap between a laboratory-based solution and a real-world industrial solution for RFD.

✉ Aneesh G. Nath
    aneeshgnath.rs.cse18@itbhu.ac.in

    Sandeep S. Udmale
    ssudmale@it.vjti.ac.in

    Sanjay Kumar Singh
    sks.cse@itbhu.ac.in

1   Department of Computer Science and Engineering, Indian Institute of Technology (BHU), Varanasi, Uttar Pradesh 221005, India

2   Department of Computer Engineering and IT, Veermata Jijabai Technological Institute (VJTI), Mumbai, Maharashtra 400019, India

1 Introduction

Industry 4.0 is bolstered by the benefits of self-configuration, self-optimization, early awareness, decision making, and predictive maintenance capabilities (Qin et al. 2016). In order to accelerate the transition from the conventional manufacturing systems to one that caters to the contemporary industrial revolution, the cyber-physical system (CPS) (Mosterman and Zander 2016) has been integrated into Industry 4.0. In CPS, the physical machine working environment is virtually modeled by incorporating sensors in the machine components. These technical advancements significantly contributed to the emergence of the discipline known as prognostics and health management (PHM) as an indispensable arm of Industry 4.0 (Lee et al. 2018). Failure detection and predictive maintenance (FDPM) is a vital component of PHM in that it ensures benefits such as optimum cost, safety, availability, and reliability by preventing the chances of catastrophic failures, serious accidents, and unexpected shutdowns of the whole system. Rotating machinery constitutes approximately 40% of all machinery and is more prone to both deterioration and failures in the mechanical and electromechanical systems (Chen 2009). At the component level, the rotating machinery is mainly divided into three parts: bearings, gears, and rotors (Chen et al. 2018). While there is a plethora of literature related to bearing and gear fault identification, comparatively fewer works address the faults affecting the rotor system (Devendiran and Manivannan 2016; Wang et al. 2017; Zhang et al. 2019b).

In the early days, the rotor fault diagnosis (RFD) research was categorized into two main streams: (i) model-based diagnosis and (ii) signal-processing-based diagnosis. In terms of the former, the physical quantities and features of the monitored system, such as stiffness, mass, and damping-matrices, are modeled using a mathematical model coupled with the facts of physics (Pennacchi et al. 2006). The modeling could be performed either at the system-level or at the component-level. However, researchers have encountered certain shortcomings with this method. For one, it has proven to be unfeasible as the system becomes more complicated with more sophisticated machinery and advanced technologies. Moreover, this approach is insufficient for updating real-time processes with newly read data. More research studies on the model-based identification of rotor faults can be found in the literature review from Bachschmid et al. (2002). Meanwhile, techniques based on signal processing (Ricci and Pennacchi 2011) rely on extracting fault-specific characteristic features from the acquired signal and coupling them with the statistical model of the suspected fault. Fault diagnosis using pure signal-processing-based methods also has certain limitations, much like the model-based methods. Nonetheless, the characteristic features identified by the signal processing methods remained the bedrock of fault diagnosis and the developed decision-making algorithms for some time (Li et al. 2015a; Rai and Mohanty 2007; Cui et al. 2017; Chen et al. 2018).

More recently, the motto of 'right information and data at the right time for decision making' became the driving force of AI-based data-driven FDPM (Oztemel and Gursev 2020; Angelopoulos et al. 2020) within the context of Industry 4.0. Under the umbrella of AI, there are two main ideologies, namely, shallow learning (SL)—which is popularly referred to as machine learning (ML)—and deep learning (DL). The availability of low-cost sensors and big data has enhanced the data-driven decision-making philosophy. The widespread use of ML and the rapid emergence of DL algorithms have ensured their popularity in the field of rotating machinery fault diagnosis (Liu et al. 2018; Zhao et al. 2019). Moreover, the robust and adaptable nature of these algorithms (Liu et al. 2018) has ensured that they are highly regarded among the research community. Unfortunately, the majority

of the research related to AI-based fault diagnosis of rotating machinery only focuses on the bearing or gear faults (Wei et al. 2019; Chen et al. 2018; Liu et al. 2018), with little attention paid to RFD. As such, the literature appears somewhat fragmentary, falling short of providing opportunities for exploiting the fault-specific characteristics of RFD and utilizing them in the feature engineering phase of AI to produce significant research improvements within the domain. Furthermore, the analysis of rotor faults is paramount since they not only have an unmediated, often disastrous impact on the performance and structural attributes of the affected equipment but may also cause secondary faults in the surrounding equipment such as the bearings and gears (Xue et al. 2013).

Within this context, this paper endeavors to study and review the recent developments in ML and DL approaches in terms of RFD, as well as to gain a thorough understanding of the characteristics and features of rotor faults.

The significance of this review on RFD can be summarized as follows:

1. Most importantly, this review presents an attempt to instill within the relevant researchers, a research practice aimed at utilizing the fault characteristics and the prior knowledge on the failure in automated diagnosis, rather than merely adopting data-driven AI approaches.
2. The majority of the studies on rotating machinery fault diagnosis focus on component-wise analysis rather than fault-wise analysis, thus leaving loopholes in utilizing the fault characteristics and the associated features (Xia et al. 2012). In view of this, this study pursues an appropriate fault-wise categorization throughout the review process to characterize the error as per the suspected fault.
3. The existing studies on rotor faults make little reference to the fault characteristics, the correlation between the faults, and the fault simulation process. This work thus attempts to provide a comprehensive reference point for RFD researchers, elaborating on rotor faults, including in terms of the practical fault simulation in testbeds.
4. While structural rotor faults (SRF) are a common and straightforward fault in rotating machinery, compared to bearing or gear faults, they receive little attention from the research community. This paper thus ventures to draw more AI research attention toward SRF and the attendant properties.
5. Rotor faults, specifically SRFs, are characterized by spectral changes in the rotating frequency and the harmonic frequencies in the vibration spectrum (refer Table 1), and the research that utilizes these fault-specific distinctive frequency components (DFC) with AI is thus emphasized in this literature review.Axl: Axial, Rad: Radial, Assoc.: Associated, Brg. Flts: Bearing Faults, C: Cause of, E: Effect
6. The studies on RFD promote the development of AI-based tools and techniques to ensure the 'integrity of the rotating machinery equipment,' which is an inevitable element of industrial safety and reliability.
7. The review traces the trajectory of fault diagnosis from traditional ML to advanced DL data-driven approaches along with the attendant advantages and disadvantages as well as a comparative analysis.

1.1 Rotating machinery fault categorization

As noted above, the appropriate categorization of rotating machinery faults is crucial to going beyond component-wise analysis and enhancing the fault-characteristics-based analysis for RFD. An example of fault categorization is shown in Fig. 1. Here, the rotor faults

Table 1 Description of rotor faults

Fault Category (Fault No.)			Vibration Characteristics			Assoc. Faults	Cause & Effect
		Plane	Symptomatic Frequency		Phase		
Strucural Rotor Faults (I)	Misalignment (1)	Angular (I1a)	Axl	2x peaks are higher than 1x peak. Stronger 3x axial and torsional responses at 1/3rd of critical speed.	A phase shift of 180° in the axial direction.	IIa (C), Brg. Flts(E)	Cause: Component expansion and cold alignment due to thermal distortions. Improper coupling alignment, Uneven foundation, imparted forces from other components, bent shaft, mass unbalance, etc. Effect: Early fatigue failure, excessive heat and friction, and component damage.
		Parallel (I1b)	Rad	2x peaks are higher than 1x peak. Stronger 1x axial and torsional responses.	A phase shift of 180° in the radial direction.		
		Parallel & Angular (I1c)	Axl & Rad	2x peaks are higher than 1x peaks. Presence of 3x harmonics.	A phase shift of 180° in radial and axial directions.		
	Unbalance (2)	Static (I2a)	Rad	Larger 1x and harmonics with less than 15% of the 1x.	A phase shift of 0° in Radial direction	IIa (C), I3a (C), IId (C)	Cause: Rotor mass eccentricity due to assembly errors, manufacturing defects, debris deposit, new fittings without counterbalancing etc. Presence of other faults like bow shaft, corrosion, falling of damaged rotor part, etc. Effect: Higher dynamic load and early fatigue failure, wobbling movement.
		Couple (I2b)	Rad	Larger 1x and harmonics with less than 15% of the 1x.	A phase shift of 180° in Radial direction		
		Dynamic (I2c)	Rad	Larger 1x and harmonics with less than 15% of the 1x.	0° to 180° phase shift.		
	Looseness (3)	Component (I3a)	Axl & Rad	Dominating 2x. Presence of 1x to 10x harmonics, subharmonics (x/2,x/3, etc.), multiples of subharmonics (2x/3, 4x/3, etc.).	Unstable phase reading.	I1 (E), I2 (E)	Cause: Improper fittings, tear and wear, thermal expansion. Effect: Damage or detach of components, secondary faults like misalignment and unbalance.
		Structural (I3b)	Rad	Higher 1x and/or 2x. Harmonics frequency amplitude above 50% of 1x.	180° phase shift between base plate and foundation.	I2 (C)	Cause: Loosened bolts or bedplates, deteriorated concrete foundation, loose or distorted machine mountings. Effect: Component looseness, unbalance, and misalignment.
Shaft Faults (II)		Bent Shaft (IIa)	Axl & Rad	Higher 1x when bent at the middle of shaft. 2x generated when bent close to the coupling.	A phase shift of 180° in the axial & 0° in the radial directions.	I2a (E)	Cause: 1) Permanent: Large unbalance force, uneven shrink fittings, residual stress or collision. 2) Temporary: Friction and thermal distortion, high length to width ratio. Effect: Higher dynamic load and early fatigue failure, unbalance.
		Shaft Crack (IIb)	Rad	Dominating 2x. Presence of 1x and certain subharmonics (1/2x,1/x, etc.).	Directly proportional to the depth of the crack.	IIa (E), IIc (E)	Cause: Severe thermal and mechanical stresses or manufacturing defects. Effect: Reduction in bending stiffness, shaft bow.
		Rub Impact Fault (IIc)	Rad	Higher 1x with 2x, 3x and other harmonics. Presence of 1/2X, 1/3X, etc.	Inconsistent phase.	I1 (C), I2 (C), IIa (E)	Cause: Excessive unbalance, misalignment, self-excited instability, resonance, thermal distortions. Effect: Excessive heat, thermal bow.
		Shaft Corrosion & Wear (IId)		Not directly evident.		IIb (E).	Cause: Electrochemical reaction due to environmental factors. Effect: Metal loss, crack.
Broken Rotor Bar (III)			Rad	Case 1: Higher 1x with, up to 4x. Pole-pass sidebands with 2x line frequency. Case2: Higher 1x with 2x line frequency sidebands.	Inconsistent phase.	IIa (E), I2 (E)	Cause: Thermal, magnetic, dynamic, mechanical, and environmental stresses. Effects: Uneven current flow, thermal bow and unbalance.

Axl: Axial, Rad: Radial, Assoc.: Associated, Brg. Flts: Bearing Faults, C: Cause of, E: Effect

are separated from the bearing and gear faults in the first step before the rotor faults are categorized from a 'cause of vibration' perspective. According to this, SRF or 1x fault is the primary cause of vibration, followed by the shaft-related faults—the secondary cause of vibration—with these two factors constituting the first two categories. Meanwhile, the broken rotor bar (BRB) fault has also been included in the fault category list. The most common and critical rotor faults, such as unbalance (UB), misalignment (MA), and looseness (LS), can be categorized under SRF (Chen 2009). A bent shaft (BS), a shaft crack (SC), a rub impact fault (RIF), and corrosion and wear (Cr&Wr) are the faults that affect the shafts. They are frequently associated with SRF and are considered to fall within the shaft fault category. In a practical scenario, it cannot be assumed that a single fault occurs in the rotor at any given time. Hence, we incorporated compound fault (CF) in the fault grouping process to represent the combination of multiple faults. Meanwhile, with the exception of rotor faults, any faults highlighted in the existing literature are considered to fall under the 'other faults'(OF) category in the remainder of this paper. The rotor unbalance induced by the uneven distribution of mass in the rotor. It causes the inertia axis of the rotor to become
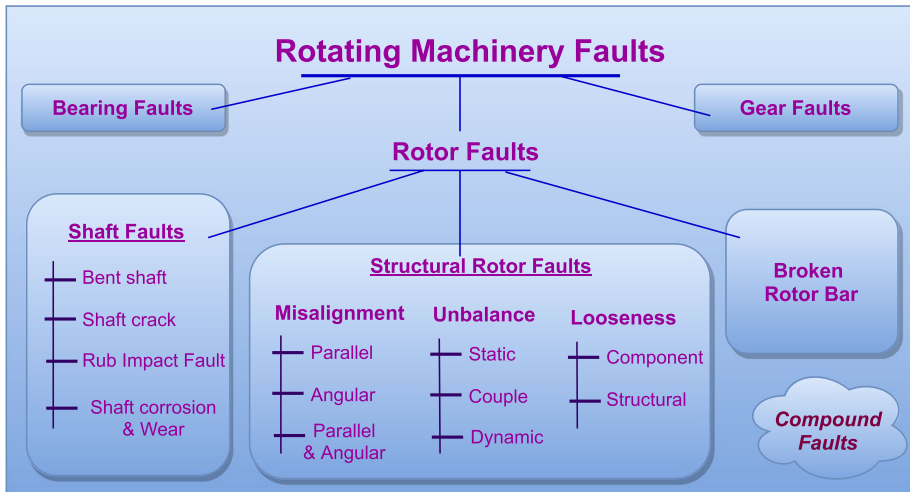
Fig. 1 Categorization of rotating machinery faults

misaligned with the geometric axis, which induces vibration in the rotor (MacCamhaoil 2016). The improper alignment of the couplings, shafts, and bearings, the thermal distortion of the bearing-housing supports, and the asymmetry in the applied load are the causes of the misalignment. This results in the bearings carrying a load higher than they are specifically designed for (Bognatz 1995). Looseness can be caused by improper assembly or the long-term running of machinery (Wu et al. 2010), and the existing literature is replete with research on bearing pedestal looseness. Bearing-related looseness can create an effect similar to that of an unbalance, while component-based looseness may result in secondary damage and detachments. It is noteworthy that around 40% of rotor-related problems can be attributed to unbalance, 30% to misalignment, 20% to resonance, and the remaining 10% to other reasons (Fahy and Thompson 2016). The contact between the rotor and the stationary parts of the machine under tighter clearances results in the rub fault. Shaft cracks are developed by severe thermal and mechanical stresses, which means the shaft cannot withstand the forces generated during normal operation (Patel and Darpe 2009b). A broken rotor bar fault (MA Cruz 2000) commonly occurs in induction motor (IM) rotors, triggering an uneven current flow through the rotor and creating both thermal and bending issues with the rotor. The electrochemical reaction due to environmental factors exacerbates the corrosion-based faults on the shaft's surface (Bonnett 2000). Section 3 provides a detailed description of every fault considered in this review.

The remainder of the paper is organized as follows. Section 2 provides an overview of the state-of-the-art techniques, methods, and algorithms used in different phases of an ideal data-driven AI-based RFD framework. Section 3 then examines the significant rotor faults along with their vibrational frequency and phase characteristics, the causes and effects, and the associated faults, while an explanation of the practical aspects and a rotor testbed-based implementation are presented at the end of the section. The research frontiers of ML- and DL-based RFD are analyzed thoroughly in Sects. 4 and 5, respectively. Here, the aim is to highlight the works that demonstrate the fault-specific feature processing, the governing data-related issues, and other RFD challenges. Section 6 then summarises the trends and challenges of each phase and provides recommendations for an insightful selection of

the techniques and methods for RFD problems. Future research directions are provided in Sect. 7 with the main aim of bridging the gap between laboratory solutions and industry-acceptable solutions. Finally, Sect. 8 presents the concluding remarks.

2 AI-based RFD framework

An ideal AI-based framework for RFD comprises three main phases: data acquisition, feature processing, and classification, as illustrated in Fig. 2. The ML part follows the above-stated phases sequentially, while DL excludes the feature processing phase due to its inbuilt feature learning capability. Nonetheless, there exist certain DL models such as belief generative models that prefer signal processed data to raw data. The methods and techniques adopted in the different phases of the RFD process under this framework are demonstrated in this section.

2.1 Data acquisition

The data acquisition phase of RFD heavily depends on the data source, the type of acquired signal, and the way the data is saved or presented. Hence, we divided this phase into three divisions: data source, sensing method, and input representation. The commonly available datasets for rotating machinery fault diagnosis -such as the case western reserve university (CWRU) dataset, the Paderborn university dataset, the PRONOSTIA dataset, the intelligent maintenance systems (IMS) dataset -have been devised for bearing-fault diagnosis and are thus not particularly relevant to RFD, especially to SRF. The SCADA (supervisory control and data acquisition) (Zhang and Wang 2014) for wind turbines (WT) was
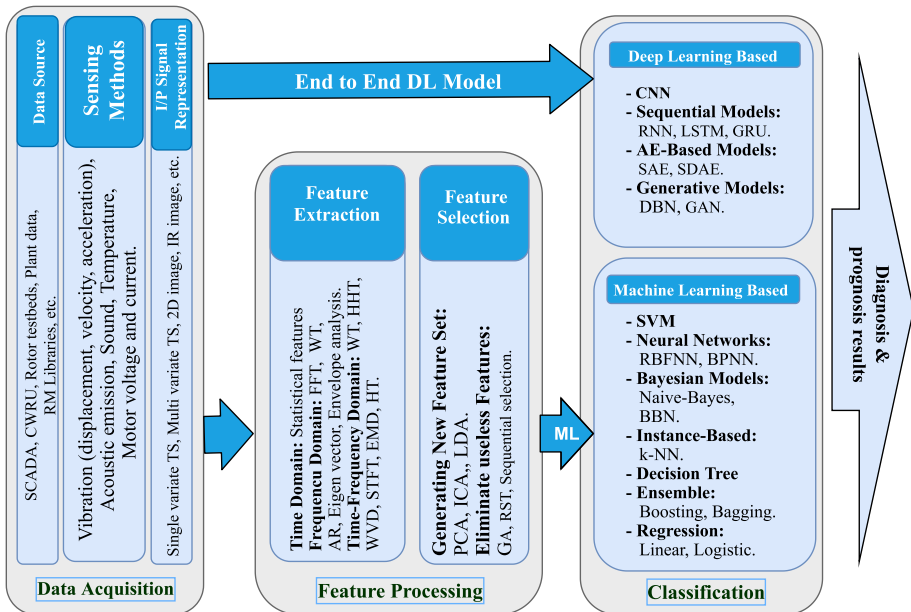

Fig. 2 AI-based RFD framework

1. VFD-L Speed Controller　2. Voltage Regulator　3. Signal conditioner　4. Three-Phase Induction Motor　5. Coupling　6. Disc with Notch
7. Tachometer　　8. General purpose accelerometer　　9. Piezoelectric Accelerometer　　10. Proximity Sensor　　11. Rotating Disk　　12. Shaft
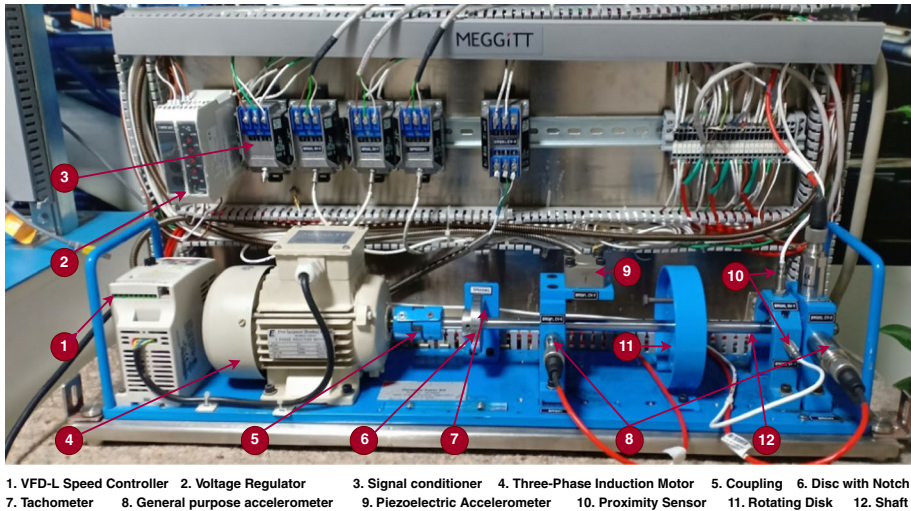
Fig. 3 A vibration data acquisition testbed (Courtesy : Meggitt India Pvt. Ltd.)

established to be a useful data source that addresses certain fault conditions of the turbine rotor. Induction motor is another acceptable data source from which voltage and current data can be acquired and monitored for broken rotor bar fault diagnosis, while it is mostly unsuitable for detecting other rotor faults. The data collected from the real-world industrial scenario is highly imbalanced, making it challenging to produce generalized solutions for RFD. The sensible and ubiquitous method adopted by the research community for collecting rotor fault data is the rotor testbeds (RTB) method. This requires employing dedicated processing software/hardware along with a computer to ensure the data collection process is smooth. A typical testbed setup with simulated faulty rotor conditions is described in the rotor fault overview section (Sect. 3). Another, but the not-so-popular option is the 'rotating machinery library' (RML) method demonstrated by Ishibashi et al. (2017), in which faulty rotor data is artificially created by applying rotor dynamics theory using efficient programming languages.

Rotor-related faults affect the nature and behavior of the vibration (Hariharan and Srinivasan 2009), and vibration has thus become the most widely used signal sensing method for RFD. The subharmonic and superharmonic frequency components of non-linear and complicated vibration motion characterize different rotor faults in distinct terms, which makes it an option of preference for RFD researchers. A rotor testbed used for acquiring the vibration data with the provisions for rotor fault simulation is shown in Fig. 3. Furthermore, the vibration-related studies have proved that SRF is the primary cause of vibration, while other shaft faults, such as rub and crack, are secondary phenomena causing the rotor to exhibit excessive vibration (Patel and Darpe 2009c). While vibration signals carry the system's dynamic information, the signals are sensitive to noise as well as the sensor mount positions. In such a scenario, alternative sensing methods such as voltage and current (V&C) (Fang and Ma 2006) have been used for data acquisition, especially in terms of induction motor rotor fault analysis. Meanwhile, a number of researchers have focused on dealing with a variety of other sensing methods, including acoustic emission (AE) (Li et al. 2004a), sound (Saimurugan and Ramachandran 2014), and temperature (Langarica et al. 2019b). Here, certain works used combined signals, which are formed either by simply

combining signals or by applying a fusion of two or more types of signal (Li et al. 2004a; Timusk et al. 2008).

As signals from various sensors are applied to a wide variety of processing models undergoing enhancement on an almost daily basis, it is crucial to know how the input is being represented. For example, vibration signals are usually represented in a univariate or multivariate time-series. However, when applying them to a convolutional neural network (CNN) model, the preferable representation is a 2D image format. The methods used to accomplish such transformations can affect the performance of the whole framework, often to a great extent.

2.2 Feature processing

The next phase announces feature processing in which the information buried in a raw signal is extracted by suppressing the noise, identifying the fault specific features, and presenting only the necessary features to the succeeding phase. Hence, feature processing accounts for a variety of signal processing techniques, and it is performed in two steps, known as feature extraction and feature selection. The methods followed for feature extraction are primarily categorized into the time-domain (TD), frequency-domain (FD), and time-frequency-domain (TFD) techniques. Here, the statistical features (Vyas and Satishkumar 2001), such as mean, variance, root mean square, skewness, or kurtosis, are commonly operated in the TD feature extraction process. In the FD, fast Fourier transform (FFT) (Tajik et al. 2015a), discrete Fourier transform (DFT), power spectrum analysis, autoregressive (AR) model (Yuan and Chu 2006), eigenvector, envelope analysis, and Welch's method are the frequently applied approaches for the feature extraction. The most popular methods—which include short-time Fourier transform (STFT) (Walker et al. 2014), empirical mode decomposition (EMD), and wavelet packet decomposition (WPD) (Bin et al. 2012)—fall under the TFD analysis, which also includes methods such as Hilbert-Huang transform(HHT), Hilbert transform (HT) (Konar et al. 2015), Wigner-Ville distribution (WVD) (Li et al. 2011), and wavelet transform (WT) (Younus and Yang 2012). Due to their exceptional performance, EMD and HHT have been the popular choices among these methods, particularly with non-linear and non-stationary signal processing. Here, it should be noted that rather than depend on these methods directly to extract feature parameters, it is better to utilize them in fault-specific symptom parameter extraction. Xue et al. (2013) made one such attempt in utilizing DFCs (1x to 5x of rotation frequency) to identify the SRFs.

The next operation is the feature selection, in which the prominent features are selected either by generating new features or by eliminating non-relevant features from the existing feature-set. The popular dimensionality reduction techniques, which include principal component analysis (PCA) (Uddin et al. 2014), linear discriminant analysis (LDA) (Tajik et al. 2015b), and independent component analysis (ICA) (Saimurugan and Ramachandran 2014), fall within the first category, while rough set theory (RST) (Konar et al. 2015), genetic algorithm (GA) (Konar et al. 2015) and sequential selection (SS) (Guyon and Elisseeff 2003) are used in the second category methods. There are no significant observations in terms of feature selection within the context of RFD, as feature extraction plays a key role in utilizing the fault-specific characteristics.

2.3 Classification

In the final phase of the AI-based framework, the classification/prediction is performed by employing ML- or DL-based models. Each of these models is further divided into sub-categories based on their underlying principle, architecture, type of input, or design philosophy, etc. It is evident from the framework that ML-based models utilize the features supplied from the previous phase, while DL-based models accept the raw data as input. The classical shallow learning ML algorithms necessitate the feature processing phase, which demands extensive domain expertise and time. The most widely applied ML classifiers for RFD are the support vector machine (SVM) model (Fengqi and Meng 2006), the artificial neural network (ANN) model (Roemer et al. 1995), and a number of their variants. Researchers have experimented with different kernels in SVM, while various learning algorithms and activation functions have been manipulated for ANN in conjunction with all the popular feature extraction methods. Under the instance-based category, the k-Nearest Neighbor (k-NN) (Chen et al. 2011) algorithm is popularly used in RFD. Probability-based Bayesian methods - including naïve–Bayes (NB) (Yusuf et al. 2013) and the Bayesian belief network (BBN) (Xu 2012)—are also demonstrated in relation to RFD, while the non-parametric type -which includes decision tree (DT) (Nguyen et al. 2008), random forest (RF) algorithms (Quiroz et al. 2018)—as well as simple classifiers such as logistic regression (LR) (Quiroz et al. 2018) and linear discriminant analysis (LDA), (Glowacz 2018) also feature within the literature. A number of researchers have also introduced the AdaBoost (AB) (Martin-Diaz et al. 2018) algorithm and other ensemble classification algorithms (Niu et al. 2008) by combining the hypotheses of already proven models. There are individual attempts (Niu et al. 2007) made to perform the comparative analysis of the above algorithms using the same dataset and signal pre-processing methods.

The conventional SL algorithms are constrained in their ability to learn the non-linear relation of features. The need for extensive computation, the required time (especially for the feature processing), and the demand for specialized expertise in the domain, are other limiting factors. By using multiple-layer deep architectures, DL methods could imbibe advanced levels of representation of input data as they go deeper, which enables them to identify more complex features on their own. CNNs (Janssens et al. 2016) have attracted more attention than any other DL model in RFD since they have proven to be as effective with temporal data as with images, leaving the user with the responsibility of providing appropriate input representation. The autoencoder (AE)-based models, such as stacked AE (SAE) (Lei et al. 2016a) and stacked denoise AE (SDAE) (Zhao et al. 2018), also find their place in RFD. Deep belief networks (DBN) (Oh et al. 2016), which is a generative hybrid graphical model composed of multiple restricted Boltzmann machines (RBM) or AEs, are also widely used in this context. The sequential DL models, well known for their ability to deal with the temporal data, such as a recurrent neural network (RNN) and the variants, known as long short-term memory (LSTM) (Lei et al. 2019) and gated recurrent unit (GRU) (Liao et al. 2019), have also been explored in relation to RFD. Meanwhile, the deep generative model known as the generative adversarial network (GAN) (Lee et al. 2017) has performed its role as a data upsampler in certain RFD works. The details of commonly used ML and DL algorithms are provided in Tables 2 and 4, respectively, while a detailed review of the literature related to these models is provided in Sects. 4 and 5.

3 Theoretical background of rotor faults

As noted above, an array of generic fault categories affecting the rotor, including SRFs, shaft faults, and broken rotor bar faults, are grouped together under the category of rotor faults in this paper. As outlined in the introduction, these three categories are divided into subcategories based on the nature and cause of the faults (Bate 1987; Mais 2002; Alsalaet 2012).

3.1 Structural rotor faults

These faults are also known as 1x faults and are considered as the primary cause of the abnormal vibration in the rotor. The three faults that fall within this category are misalignment, unbalance, and looseness.

3.1.1 Misalignment

The scenario where bearings, shafts, and couplings are not properly aligned along their centrelines is termed misalignment (Alsalaet 2012). As machines operate, the heating and cooling results in component expansion and cold alignment, respectively, the net result of which is the misalignment of the machinery components. Continuous operation with an uneven foundation, a shift in the foundation, or the improper alignment introduced by imparted forces from other components can also lead to misalignment (Mais 2002). In addition, there is a high chance of installation misalignment to the system if the couplings are not properly set up. Meanwhile, faults such as a bent shaft and improper bearing seats also cause a certain type of misalignment with effects and symptoms similar to those of normal misalignment (Alsalaet 2012). Both misalignment and unbalance result in the bearings having to bear a higher dynamic load than they are specifically designed for and ultimately leads to failure due to early fatigue.

In terms of the couplings, misalignment creates excessive heat and friction, and thereby damages the component, while in terms of the horizontal shafts, any misalignment causes excessive vibration on both the vertical and axial planes. Finally, any misalignment in the overhung horizontal shaft is characterized by excessive vibration in the horizontal and axial planes. Axial vibration is a salient component of both these scenarios. In the case of the vertically aligned shaft, excessive horizontal and axial vibration indicates misalignment (Mais 2002).

There are three types of misalignments called parallel, angular, and parallel and angular misalignments (Nakhaeinejad and Ganeriwala 2009).

(i) Parallel misalignment: In parallel misalignment, the centerlines of shafts joined by the coupling are parallel, but they will be at an offset. It is characterized by strong radial vibration.

(ii) Angular misalignment: When a bending force is induced on the shaft by the joint at the coupling, such misalignment is called angular misalignment. In such cases, joining shafts' centerlines are crossed at an angle between them, which causes strong vibration in the axial direction.

(iii) Parallel and Angular misalignment: It is the combination of both parallel and angular misalignments, and hence produces vibration in both axial and radial directions.

Fault characteristics analysis: Misalignment can be analyzed by comparing the ratio between 1x (unbalance indicator) and 2x (misalignment indicator) components. In normal misalignment scenarios, 2x and 1x components are present in the radial vibration spectrum, with 2x being the predominant component with a range up to 150% of 1x (Mais 2002). The severe misalignment conditions are characterized by the harmonics 3x to 8x or even a full high-frequency harmonics series, and its vibration amplitude at 2x running speed will be greater than 150% of that at 1x. Patel and Darpe (2009c) presented certain observations to identify the type of misalignment as well as to distinguish between misalignment and crack faults in the rotor. In order to uniquely identify misalignments, they suggested investigating the presence of strong negative higher harmonic frequency components compared to that of rotor crack. They also observed that stronger 1x axial as well as torsional responses and weak higher harmonics represent parallel misalignment. They added that a strong 3x harmonic component in its axial and torsional vibration response at 1/3rd critical speed indicates angular misalignments. Phase information is also used as a defining feature of misalignments. Across the coupling or machine, a phase shift of 180° in the axial position shows angular misalignment, and the same in radial position indicates parallel misalignment, whereas that in both axial and radial position represents combined misalignment (Alsalaet 2012). The vibration waveforms follow a periodic pattern having one or two cycles in each revolution of the shaft.

3.1.2 Unbalance

The faulty state of a machine that occurs when the centreline of the mass of the rotor (inertia axis) and the center of rotation (geometric axis) are non-coinciding is known as unbalance (MacCamhaoil 2016). Rotor mass eccentricity created by the uneven build-up of debris on the rotor, assembly errors (unidentical blades of the wind turbine, windings of the generator rotor, etc.), and the addition of new fittings to the rotor without appropriate counterbalancing, are the major causes of an unbalanced rotor (Mais 2002).

In addition to these aspects, a number of other rotor faults can also lead to unbalance. For example, a bent shaft or loose parts can change the balance, while factors such as corrosion, abrasion, and the falling of damaged rotor parts also lead to unbalance. The critical components of the machine, which include the gears, bearings, and couplings, may be destroyed by an unbalance fault. In fact, unbalance particularly affects the bearings in that, as noted above, they are forced to carry a higher dynamic load than they are specifically designed for, which ultimately leads to failure due to early fatigue. During operation, rotating structures will experience a wobbling movement that is characteristic of the vibration resulting from the unbalance.

Unbalance produces a radial vibration; that is, it is part vertical and part horizontal. Since the machine is generally more flexible in the horizontal plane, excessive vibration is a good indicator of unbalance. Under ideal conditions, axial measurements should indicate weak vibration as most forces are generated perpendicular to the shaft (Mais 2002). In terms of vertically placed shafts, the unbalance is due to the mass effects of radial plane vibration, and the dominant frequency component will be 1x. A phase shift of 90° occurs as the sensor moves from the horizontal to the vertical position, and no radial phase shift across the machine or coupling occurs (MacCamhaoil 2016).

In general, there are three types of unbalance:

(i) Static unbalance: The unbalance observed when at rest is known as static unbalance or force balance, wherein only one force is involved in affecting the balance. In this case, the inertial axis of a rotor is displaced and lies parallel to the axis of rotation. This issue is widespread in disk-shaped rotors since there is a high chance of uneven mass distribution that creates a parallel shift between the axes. In simple systems, the chance of static unbalance is greater than that of couple unbalance.

(ii) Couple unbalance: Two equal forces or weights placed 180° apart, which make the rotor appear balanced statically, are known as couple unbalance. This issue is not observable when the rotor is at rest and is a typical phenomenon in elongated cylindrical type rotors. A couple unbalance causes the so-called 'wobble effect,' which produces a 180° out-of-phase reading from opposite ends of the shaft. Complicated systems with multiple locations in the rotor with unbalanced weights or systems with more than one coupling are prone to this type of fault.

(iii) Dynamic unbalance: The unbalance condition that occurs in real systems is known as dynamic unbalance. This issue presents a combination of static unbalance and couple unbalance. It has been observed that dynamic unbalance is present in virtually every rotor and that it can be addressed only by applying weights on at least two planes.

Fault characteristics analysis: The vibration wave is sinusoidal and occurs at a frequency of 'one per revolution' (1x), i.e., a single frequency vibration with the same amplitude in all radial directions. Other than the severe faulty states, vibration generally contains 1x only, without any harmonics of it. The 1x with high amplitude and its harmonics with less than 15% of the 1x is an indication of an unbalanced state (Al-Bedoor 2001). In this case, up to the first critical speed of the machine, the amplitude increases with speed, and the phase from the vertical and horizontal measurements differ by 90°. In dynamic unbalance, there will be 180° phase shift in the radial direction, while static unbalance shows a phase shift of 0°. For unbalance due to the bent shaft, the phase shift of 180° happens in the axial direction with no phase shift in the radial direction (Mais 2002).

3.1.3 Looseness

Looseness is categorized into two types depending on whether the part affected by looseness is a mechanical or a structural component. When the mechanical components are fitted incorrectly, this results in component looseness, while the relative movement between the surfaces of the fundamental structures results in structural looseness. This looseness causes excessive horizontal, vertical, and structural vibrations in the horizontal and overhung horizontal shafts, where excessive horizontal and structural vibrations result in the issue extending to the vertical shafts. The looseness issue accounts for more vertical vibrations than horizontal vibrations (Ma et al. 2011).

(i) Component looseness: Rotating components and/or non-rotating connections that constrain the shaft to its rotating axis (such as the bearing base [pedestal], bearing mounts, and bearing caps) can experience looseness due to improper fittings, wear and tear, and thermal expansion. This is known as component looseness, which results in the components becoming damaged or detached from the assembly, thus inducing secondary damage. For example, component looseness can lead to relatively small residual misalignment, causing increased vibrations that affect both the radial and axial planes. If the loosened components are rotor mounted, this may result in unbalance.

(ii) Structural looseness: While the fundamental structures are not supposed to move freely, loosened bolts or bedplates, a deteriorated concrete foundation, and loose or distorted machine mountings can create a slight movement between the surfaces of these structures. This is known as structural looseness and is often created between one vibrating component, generally the foot of the machine and one stationary component, that is, the foundation. Both structural looseness and soft foot looseness can lead to vibration in the radial plane.

Fault characteristics analysis: In the case of component looseness, the initial stage vibration signature contains mostly 1x and 2x components, but with escalated deterioration, the fractional harmonics with increased amplitude starts to appear. Generally, looseness is characterized by several running speed frequency harmonics (1x–10x) with subharmonics (x/2,x/3, etc.) and their integer multiples (2x/3, 4x/3, etc.) of magnitudes greater than 20% of the 1x amplitude. Structural looseness creates 1x and/or 2x radial components with predominant vertical amplitude, subject to the type of issue. For rigidly connected machines with no belts or couplings, the radial 2x signifies looseness. The waveform generated is periodic with one or two cycles per revolution. A phase difference of 180° exists between the foundation and vibrating components in case of structural looseness (Mais 2002).

3.2 Shaft faults

In addition to SRF, faults such as bending, cracking, rub impacts, and corrosion and wear, which affect the shaft directly, can occur and are known as shaft faults. These occur either as a consequence of SRF or due to other external reasons. The majority of shaft faults are known to be the secondary cause of abnormal vibration. The most important shaft faults are outlined below.

3.2.1 Bent shaft

A bent shaft is the most common rotating machinery fault and often develops as a result of thermal distortion, creep, or a large unbalance force (Darpe et al. 2006), while rotor rub is another cause of thermal rotor bowing. As a result of gravity, a rotor can experience cold bow while in a resting position, especially in shafts with a high length to width ratio. The rotor bow on a rotating machine will result in the re-emergence of the static unbalance faulty condition. High torque pressure and improper handling can also be the reason behind a bent shaft. Bent shafts induce effects similar to those of misalignment, causing the shaft to bear a higher dynamic load than they are specifically designed for, ultimately resulting in failure due to early fatigue (Mais 2002).

Fault characteristics analysis: If the bent is close to the middle of the shaft length, then 1x will be dominating, and 2x will be dominant for the bents near to the couplings. The most affected plane is axial, though vertical and horizontal planes will also give out 1x and 2x peaks (Lees et al. 2009). The 2x amplitude can vary from 30% of the 1x amplitude to 100–200% of the 1x amplitude. The spectrum of bent shaft is almost similar to that of misalignment; hence, the phase can be used as a good differentiating indicator. In bent shafts, the radial phase measurements will be in phase, while axial measurements will be 180° out of phase at opposite ends of the component (Mais 2002).

3.2.2 Shaft crack

Severe thermal and mechanical stress or manufacturing defects result in weak spots in the shaft, thus limiting its ability to withstand the forces generated during normal operation. This is known as shaft crack. The causative stress may be cyclical, whereby the initial crack is converted into a fatigue fracture resulting in a sudden breakage of the shaft. The first and foremost effect of shaft crack is the reduction in bending stiffness in the direction of the crack, which results in inducing excessive 2x vibration in the shaft (Darpe et al. 2006). The second effect is the rotor bow, where the bending results in the natural axis shift corresponding to the direction of the crack. This effect generates 1x components, which will progressively add to the already existing residual unbalance frequency component (Lees et al. 2009).

Fault characteristics analysis: The cracked shaft vibration mostly affects the radial plane, and it produces increased 1x vibration along with the 2x and 3x harmonics. In certain scenarios, the presence of subharmonics 1/2x,1/x, etc. are also observed with this fault. Studies revealed that the phase shift is directly proportional to the depth of the crack (Rahman et al. 2013; Darpe et al. 2006).

3.2.3 Rub impact

The contact developed between the rotating components and the stationary components creates a rub impact, the effect of which will remain passive in terms of overall vibration. Excessive unbalance, misalignment, self-excited instability, and resonance can cause stator-rotor rub faults in the system, which results in highly non-linear vibrations (Wang et al. 2016). The static forces working on the rotor in the radial direction and the thermal distortions in the casing have been identified as the cause of stator-rotor rub. The heat induced by the asymmetric friction produced by the rub impact in one-per-revolution fashion can result in a thermal bow in the rotor, which induces both an unbalance effect and a phase difference. As the rub impact becomes stronger, the degree of non-linearity increases, thereby creating higher amplitude harmonics of rotating frequency.

Fault characteristics analysis: The vibration changes in characteristic ways whenever the rub happens between stationary and non-stationary components. So other than the times there is no contact, the waveform seems absolutely normal. The spectrum will be containing higher amplitude subharmonics and superharmonics of the synchronous frequency with a strong rub impact. Chu and Lu (2005) observed the presence of 1/2 fractional harmonic components and 1/3 fractional harmonic components along with the 2x, 3x harmonic components. Meanwhile, the existence of pseudo-resonance and backward whirling components showed by Patel and Darpe (2009a) as the effect of rub. The phase information shows that it can not maintain a consistent phase in rub related vibration motion.

3.2.4 Shaft corrosion and wear

Corrosion and wear fall under the category of nonfracture-type shaft failure. The electrochemical reaction due to environmental factors causes corrosion, and the consequential wear results in the metal being worn away, further increasing the stress and ultimately culminating in fatigue cracks. For the most part, these faults will not result in shaft failure; however, in conjunction with the other faults, they may cause fatigue failures that leave clear evidence. Here, cracks expand in a direction perpendicular to the applied stress on the

area, from where the metal part is removed due to the debris created by oxidation. Meanwhile, pitting presents a corrosion issue that leads to short-term failure, which results in a small amount of material loss from the shaft periphery (Bonnett 2000).

3.3 Broken rotor bar

Broken rotor bar faults are an omnipresent fault in induction motors and are often caused by thermal, magnetic, dynamic, mechanical, and environmental stresses (MA Cruz 2000). The resulting uneven current flow generates localized heating and results in a thermal bow on the rotor, which, in turn, results in an unbalance rotor fault. Though, a higher 1x amplitude characterizes the vibration in the radial direction, stator or rotor current analysis is commonly adopted in broken rotor bar fault diagnoses (Bate 1987).

Fault characteristics analysis: Uneven current flow to induction motor rotor owing to crack or break results in two kinds of vibration changes. In the first case, 1x and harmonics generally up to 4x will accompany with pole-pass sidebands with quite low 2x line frequency Mais (2002). In the second case, the amplitude of rotor bar passing frequency (RBF, which is the number of rotor bars times the running speed) will be higher, with 2x line frequency sidebands (100Hz or 120Hz) (Bate 1987). Due to the motion of vibration, the phase will be consistent in broken rotor bar faults.

Experimental setup for rotor faults

To analyze rotor faults other than broken rotor bar fault, more than 80% of the works depend on rotor testbeds for faulty data generation. Figure 4 presents a diagrammatic representation of a typical rotor testbed-based data acquisition setup along with the fault implementation and the associated frequency responses.

The data acquisition setup contains the RTB, the signal-conditioning unit (SCU), and the monitoring and analyzing unit (MAU) (Xue et al. 2013). The main component of the testbed assembly is an electric motor controlled by a variable frequency drive (VFD, which varies the frequency and voltage supply) connected to a rotating shaft via a flexible coupling. The shaft is supported by bearing housing that is fixed by pedestal bolts to the testbed base. The bearing housing has provisions for mounting sensors in horizontal, vertical, and axial directions. The shaft will be mounted with discs for various purposes. Generally, a disc with one or more notches is closely placed with a tachometer that senses the rotational speed of the shaft. This tachometer reference, along with a sensor waveform, is used to calculate the phase information. Meanwhile, weighing discs, which will likely incorporate grooves, holes, or another eccentric weight connecting provision, will also be mounted on the shafts. The majority of the testbeds are equipped with a metal bush connecting facility to create the rub impact (Lu et al. 2019a). Contact type general-purpose accelerometers and non-contact type proximity sensors are commonly used for the vibration data acquisition. The acquired data is then passed through an SCU, which often performs signal amplification and analog-to-digital conversion (ADC). It is then applied to the MAU, which is generally a computer that carries out the monitoring and analysis of the input signal and any associated tasks with the help of specifically developed software.

The unbalance fault is created by connecting weights (bolts, nuts, washers, etc.) on the weighing disc. Two equal weights are placed 180° apart to create couple unbalance, while
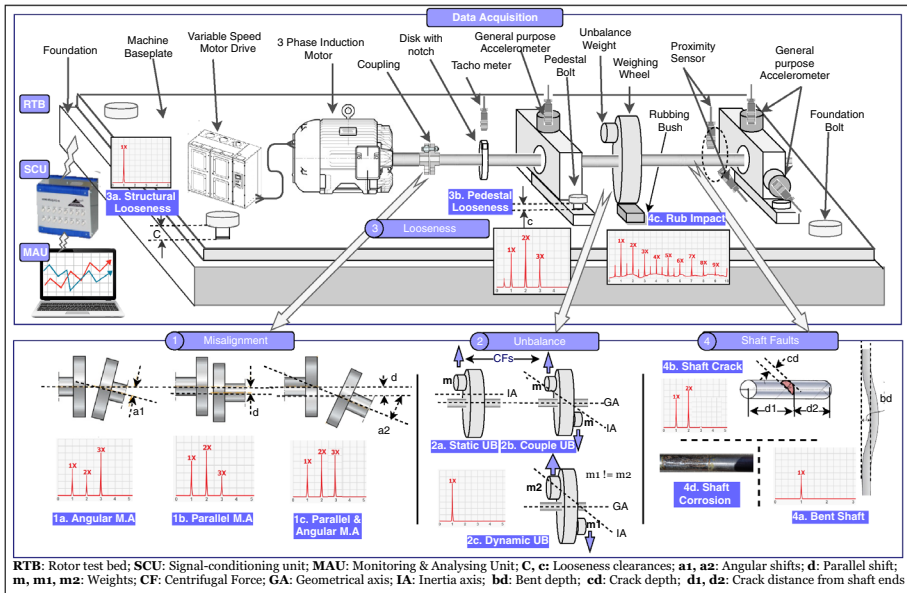
Fig. 4 Rotor fault implementation in a testbed

a one-sided weight is positioned to generate static unbalance. Much like in couple unbalance, if unequal weights are set in place, then dynamic unbalance can be simulated (Walker et al. 2014). The coupling is loosened to induce a parallel shift and/or angular shift to create parallel and/or angular misalignment (Lu et al. 2019a). In terms of loosening faults, the majority of the works deal with pedestal loosening, which is simulated by loosening the pedestal bolt of the bearing housing to create clearance from the testbed base (Xue et al. 2013). Meanwhile, structural looseness can be demonstrated by loosening the bolt connecting the testbed base to the foundation. The stationary bush that is in contact with the shaft or disc periphery will induce rub impact (Walker et al. 2014). The shaft bend and shaft crack provisions are also shown in Fig. 4. Here, the experiment is conducted by setting different fault conditions as described before, and then the testbed is set to different rotational frequencies and load conditions to collect data via repeated trials.

4 Machine learning-based approaches to RFD

RFD research has been successful in employing a variety of SL algorithms before and after the advent of DL methods in the research arena. Irrespective of the fact that the ML methods require a complex feature engineering process that demands sufficient domain expertise and time, it has certain advantages that make it an option of preference for many machine health monitoring applications. The provision for applying domain knowledge and the comparatively fewer data requirements make the ML method significantly more favorable in this regard. One of the earliest methods found in the existing literature involves using ANN for motor fault diagnosis. Thereafter, the research in this domain flourished with several algorithms adopted across the supervised and unsupervised categories. The commonly used ML methods, along with their features, pros and cons, relevance, and compatibility with RFD, are presented in Table 2.

Table 2 Summary on ML classifiers

Classifier	Properties	Pros	Cons	Compactibility and Relevance with SRF
SVM	# Constructs a hyperplane in original input space that maximizes data separation # Solve the problems of binary classification. Multi-class is solved by breaking it into a pair of classes such as one-against-one and one-against-all # Both linear and non-linear kernels are available	# Good Accuracy, and fairly robust against overfitting # Works well in HD space and can model non-linear decision boundaries # Have a wide range of kernels # Can be used for both classification and regression.	# Memory Intensive, expensive in computation and time # Can't scale well to larger datasets # Performance depends on selecting the right kernel	Suitable for RFD with small datasets. Outlier handling ability deals well with sensor data issues. Frequency domain fault specific features will give a good performance. Not suitable for large industrial dataset
Neural Networks: ANN, RBFNN, BPNN	# A non-linear and multivariable model # Network of processing elements called neuron connected by links with weights # Weights are adjusted to reflect the relation between features in training # Can be supervised or unsupervised	# Better for the non-linear data with more input features # Works well with noisy data # Easy to use	# Computationally expensive # Overfitting and local minima selection # Depends on initial parameter selection and parameter tuning need not help always # Need of prior knowledge in feature extraction	Gives provision to apply large feature set of RFD, both in FD.& time FD. Recommended for the RFD industrial problems. It can deal with noisy sensor signals well
Bayesian: NB, BBN	# Naive–Bayes is a generative probability model # Works based on conditional probability and counting # Applicable for classification and regression	# Very simple operation but better performance # Works with less training data without needing distribution information # Faster convergence and scale well with datasets	# Data dependency may affect accuracy # The basic assumption on conditional independence rarely practical in the real world # A well-trained model can beat its performance	RFD can't assure independent feature vector, as fault specific features are highly correlated. Inability to learn from time-series and the problems due to the absence of data are the other limiting factors for using in RFD

Table 2 (continued)

Classifier	Properties	Pros	Cons	Compactibility and Relevance with SRF
Instance-Based: k-NN	# Non-parametric method with interpretability issues # Discovers the unidentified data point using the historical data points by neighborhood information and majority voting # Can be used for classification and regression	# Simple and easy to implement, Needed very less hyperparameter tuning. # Can deal with decision boundaries of any form	# Demands sensible selection of parameter 'k' # Needs feature scaling for smooth performance # Since no trained model, runtime computation increase with an increased sample size	Needs proper signal processing to deal with outliers. Fails to find important fault specific attributes. Hence, difficult to provide good interpretable output for RFD. Feasible with a moderate sample size of RFD
Ensemble: Boosting, Bagging	# By combining the hypothesis of already proven models, it generates a new hypothesis enhanced with the pros and with eliminated cons of different methods	# Provision for combining methods that give the best performance with highly non-linear and non-stationary data.	# The implementation may be time and power-consuming # Complexity and interpretability of the combined methods may not be acceptable	Fault specific features, including DFC, can be derived and utilized by feature fusion and data fusion for RFD. DFC with ensemble methods will give the best result

4.1 Artificial neural network

In AI-based machine health monitoring, the history of the application of ANNs stretches over three decades (Chow et al. 1991), and it has been an indispensable part of the literature on SL of RFD right from its inception. The approaches using ANN in RFD are custom-made according to the variations of ANN architecture and signal processing techniques used in feature processing.

The method of integrating finite element (FE) analysis with the neural networks is adopted in specific works of RFD. In an early attempt, Roemer et al. (1995) proposed FE-based neural networks for classification of unbalance and misalignment rotor faults, which enable real-time machinery diagnostics and component stress prediction. They observed that shaft is affected by low cycle fatigue with the start/stop cycle of the rotor under unbalance condition, while bearings are more affected by high cycle fatigue correlated with rotor speed. Later, Yu and Han (2010) also developed an FE model of the rotor with transverse crack(s), and an ANN is proposed to identify the location and depth of a crack in the rotor. They introduced fracture mechanics theory and the energy principle of Paris for modeling of FE for single and multiple cracks.

Several researchers have tried to find out the best suited ANN architecture and feature extraction method for specific tasks in RFD. The work of Hassan et al. (2003) was one among those, in which they compared different NN models such as perceptrons, linear filters, feed-forward, self-organizing networks, and learning vector quantization (LVQ). They concluded that perceptron and LVQ architectures were superior to others, and could achieve 100% accuracy in unbalance and looseness conditions. Applying the findings, El-Shafei et al. (2007) used the LVQ neural network, in series and in parallel connections separately to a fuzzy inference engine for addressing the SRFs. To overcome the limitation of LVQ in handling compound faults, a feed-forward NN with resilient propagation (RP) algorithm was introduced. They proved the success of NN as a high-resolution configuration diagnostics tool and appropriateness of fuzzy logic as a low-resolution configuration diagnostics tool. They realized the neuro-fuzzy system as an option that can provide a confidence index for fault. Another work from the perspective of performance comparison was done by Tajik et al. (2015b), who applied six different methods for feature extraction and three different NNs (BPNN, RBFNN, PNN) for classification with PCA and LDA dimensionality reduction techniques, for different rotor unbalance conditions. Reddy and Sekhar (2013) compared the performance of ANN by applying statistical features and amplitude in the FD. Applying on unbalance and pedestal looseness conditions, they concluded that statistical features give good results over FD magnitudes.

A few works have registered to incorporate the TD features in ANN, whereas FD and TFD features can be seen in a fair amount of works lately. Vyas and Satishkumar (2001) used statistical features integrated with ANN, in which moments of vibration signals acquired in the TD were utilized as features, thereby reducing the number of inputs produced after high resolution preprocessing techniques. They applied a backpropagation ANN to detect five rotor faults, including looseness, unbalance, and misalignment. Singh and Kumar (2015) used six statistical features extracted from the TD vibration signal and compared the performance of ANN and SVM in terms of single and multiple faults of shaft bow and misalignment, and concluded that SVMs require comparatively less running time. Similarly, in statistical feature extraction, there were some novel attempts from some researchers like Li et al. (2015b). They proposed the feature extraction method by taking the sampled average of some conventional features to improve the performance of ANN

and SVM. Apart from that, an evaluation method is also developed using the decoupling vectors and the thresholds based on simple algebraic computation.

Certain significant contributions have been observed in the preprocessing of input data in FD and TFD for the smooth learning of neural networks in RFD. Bin et al. (2012) conducted FD preprocessing with EMD and WPD to extract five spectral bandwidths energy of vibration signal and successfully identified ten rotor faults, including rotating stall and breathing vibration, loose rotating parts, and fluid-related defects. The ability of WPD in generating a time-frequency spectrum containing the wavelet packet coefficients for each of the depths and nodes, was utilized by Sadeghian et al. (2009) for feature extraction from stator current of induction motor. The extracted coefficients with different frequency resolutions of the same frequency component, together with the slip speed, were applied to the ANN. This work has its significance in the online detection of faults. Chen and Chen (2011) addressed six common faults of the rotor system like unbalance, bow, misalignment, rub, whirl, and whip using a set of individual neural networks based on structured genetic algorithm (sGAINNs). The highlighting feature of this work is that the authors have made a significant contribution in FD preprocessing and in exploiting the effect of frequency harmonics in rotor faults. Unbalance localization was performed by Walker et al. (2014) through the simulation of different types of unbalances, including multiple planes unbalance, and they derived the subsynchronous non-linear features in the FD as an input to ANN. This unbalance localization is particularly suggestive in situations where sensor placement options are limited. The results of unbalance, misalignment, rub, or combinations of these faults were also presented in the same work. Shaft bow, unbalance, and their combined effect was studied by Srinivas et al. (2010a). Daubechies wavelet was used for FD transformation, and they utilized amplitudes of 1x, 2x, 3x, and 4x vibration harmonics in horizontal, vertical, and axial directions resulting in an accuracy of 99.9% using ANN. The same authors continued their study on unbalance and cracked rotors in (Srinivas et al. 2010b). Frequency spectrum analysis on combined faults of unbalance and bearing clearance was another contribution of Srinivas and Holla (2014). As in their previous work (Srinivas et al. 2010a), the same Daubechies wavelet transform and amplitudes of vibration harmonics were utilized in this work too. In one recent work, Pang et al. (2020) used various methods such as ensemble empirical mode decomposition (EEMD), Hu invariant moment feature vector, and morphological image processing with BP neural network. Unbalance, misalignment, oil whirl, and oil whip faults are simulated in both single-span rotor and double-span rotor test rigs achieving an accuracy rate up to 95%.

Moreover, ANN is viewed as a replacement for conventional feature extraction methods. One such approach was proposed with a two-level learning stage by Lei et al. (2016b). In the first stage, an unsupervised two-layer neural network known as sparse filtering was applied to learn features from mechanical vibration signals directly, and then softmax regression was employed to classify the health conditions in the second stage. Without any specific feature extraction technique, an improved accuracy along with the increase in the number of unlabeled data was obtained by this method. The works which consider the fused features to represent the multiple faulty conditions are also available in the ANN literature of RFD. A feature fusion model based on information entropy and the probabilistic neural network was proposed by Jiang et al. (2018). They proved that the accuracy of their fusion method was 10% higher compared to the processes using each of the information entropies separately.

In addition to testbeds, induction motors and wind turbines are the machines where ANNs are employed mainly to identify broken rotor bar faults and unbalances, respectively, often using voltage and current input. Broken rotor bar fault, along with misalignment and

unbalance, was determined by Cabal-Yepez et al. (2015) using statistical features of voltage and current signal. They collected the data from variable speed drives (VSD) rather than the direct connection. A field-programmable gate array-based implementation was developed to offer an online, system-on-chip solution for real-time condition monitoring with ANN to perform classification. The simulation of five imbalance fault conditions of a wind turbine was performed through the TurbSim, FAST, and Simulink by Malik and Mishra (2017) considering the parameters from three-phase stator current and voltage. By using EMD and PCA in preprocessing and ANN in classification, the method achieved a higher degree of accuracy. Cacciola et al. (2016) utilized NN to decide whether the unbalance is present or not on a wind turbine rotor using 1x harmonics amplitude. By using load harmonics analysis, they could determine the severity of unbalance and its location. They additionally classified the root cause of the unbalance by rotor response. Acoustic signal, a supposedly robust input, was applied to ANN by Li et al. (2004b). The signal was converted to normalized power spectra together with the rectified statistic moments to identify unbalance in rotating machinery. They introduced the d-normalization technique to nullify the distance effects of sensors and components.

The ensuing discussion can be summarised as follows. As part of one of the top-rated ML methods, ANNs feature a great deal in RFD, extending to around 22 of the 71 papers falling within the ML category. Here, it is clear that the majority of the researchers have tried different feature processing methods to improve the output, while a fair number of works follow TD approaches, despite the prevailing trend of adopting FD approaches. Similarly, the proportion of works that have modified the ANN structure (23%) and the number of works that consider DFC feature extraction (27%) very high with ANN. Moreover, the percentage of works (73%) that deal with RFD exclusively (without considering bearing or gear faults) demonstrates the fact that ANNs are highly suitable for RFD analysis.

4.2 Support vector machine

As one of the frontrunners in both classification and regression tasks, SVM has been proven to be a decent performer in machine health monitoring over the last two decades. The method is well known for handling data overfitting and for its exceptional performance in terms of accuracy in RFD research.

There are numerous ways in which researchers modify SVMs for adapting to their research problems in RFD. The following are a few approaches that made their way in changing the conventional method of using SVM by upgrading the kernel or incorporating other techniques and data structures with SVM. Widodo and Yang (2007) developed a two-phase kernel algorithm with kernel PCA and ICA. The performance of the classification process using various feature extraction methods and the kernel function were presented in their work. The same authors (Widodo and Yang 2008) later introduced an SVM with a kernel function using the wavelet transform known as wavelet support vector machine (W-SVM) with strong generalization capability. Transient current signal, preprocessed by discrete wavelet transform and PCA, was used as input for the classifier. They proved that introducing non-linear kernel using wavelets improves the SVM performance significantly. They presented another work (Widodo et al. 2007) in which the sequential minimal optimization algorithm carried out training of the SVMs, and various scenarios concerning this were examined using datasets of vibration and stator current signals. They concluded that SVMs achieved high performance in classification using multiclass strategies such as one-against-all (OAA) and one-against-one (OAO). A broken rotor bar fault detection system

based on stationary wavelet packet transform (SWPT) and multiclass wavelet kernel support vector machines (MWSVM), was presented by Keskes et al. (2013). They discussed the classification performance of different multiclass SVMs with various kernel functions. They found that OAA is about five times faster than OAO, and it required less number of singular values. Kang and Kim (2013) proposed a multi-layer SVM (MLSVMs) with RBF kernel, which extracted features from both the TD and the FD. In the TD, it used short-time energy (STE) with the singular value decomposition (SVD) technique to represent various faults. In the FD, it adopted the discrete cosine transform (DCT) with the SVD technique and addressed eight different rotor faults.

By fusing the advantages of the information entropy method and SVM, Fei et al. (2014) proposed an information fusion based method known as process power spectrum entropy and SVM (PPSE-SVM). PPSE feature vector has robust learning and generalization ability, fault tolerance ability, and strong anti-noise interference ability so that the proposed method demonstrated state-of-the-art performance. Zhang et al. (2015) proposed significant improvements for SVM as well as the feature processing methods to bring better performance even in fewer faulty data situations. They applied the fuzzy support vector machine (FSVM) optimized by a multi-population GA. The EMD and WPD were utilized in the preprocessing of the vibration signal. This work succeeded in making full use of rotor dynamics and computational intelligence. Unbalance conditions in wind turbines were identified by Malik and Mishra (2016) using intrinsic mode functions (IMFs) with the EMD method for decomposing stator current input and PCA for dimensionality reduction. The significance of the work was the use of proximal SVM (PSVM), which had been invented by Mangasarian and Wild (2001). By the rendered results, they showed that the method was appropriate for online fault diagnosis applications. Saimurugan and Ramachandran (2014) made a comparison between vibration and sound input signals for an SVM classifier, with and without fast-ICA based preprocessing. They simulated 12 health conditions with bent and/or unbalanced shaft, with different bearing fault conditions. They concluded that in certain circumstances, vibration-based diagnosis performs much better than sound signals in RFD, but both are effective in fast ICA-SVM based classification. Tang et al. (2010) showed that the multiclass SVM trained with chaos particle swarm optimization (CPSO) outperforms ANN in identifying rotor faults like unbalance, misalignment, bending, and bolt looseness. Another work of PSO with SVM done by Duan et al. (2016a) proposed a support vector data description with a binary tree structure for multi-classification problems with unbalanced datasets. The parameters of support vector data description are optimized by PSO, resulting in the application of BT-PSO-SVDD technique for imbalance classification, which was able to outperform ANN, SVM, and fuzzy c-means (FCM).

In data and feature representation, SVMs show their variegation; and specific such works are portrayed in literature. One such far-reaching RFD data representation method that operated with SVM is infrared or thermal images. They have the advantage of being non-contact and non-intrusive, and thereby able to avoid many hurdles commonly experienced with the other sensing methods. Younus and Yang (2012) used the thermal image as an input with a two-stage process for RFD. The decomposed data by a 2-D DWT operation on the input was applied to a Mahalanobis distance and relief algorithm for feature selection. By using SVM and LDA classifiers, they proved the applicability of the method. Uddin et al. (2014) put forward a different approach in which a 2-D gray-scale texture created from TD vibration signals, was applied as input to the feature extractor. The dominant neighborhood structure (DNS) map generated as a result of feature extraction was applied to OAA multiclass support vector machines with RBF kernel. Among the eight rotor health

conditions laid down, this approach gained 100% accuracy. Similarly, Janssens et al. (2015) gave preference specifically to extract features like histogram features, measurement of concentration related to the spatial temperature distribution, Gini coefficient, etc. They identified unbalance and various bearing faults in two different pipeline stages. SVM and random decision forest (RDF) classifier performed the classification task. Another work with a similar kind of data was done by Duan et al. (2016b), in which image segmentation and histogram processing with PCA dimensionality reduction was applied with SVM and NB classifiers. It was also reported that the SVMs work with features extracted by CNN to enhance the fault diagnosis performance (Sun et al. 2017).

Moreover, it is important to highlight some works with SVM, which have given far-reaching attention to the feature extraction process. Li et al. (2015b) developed a method for evaluating feature extraction and made use of SVM to validate the performance in the RFD. Based on the central limit theory, the statistical features were extracted so that it followed a normal distribution and were evaluated with a decoupling technique and compared with thresholds to decide on fault category. An improved version of the distance evaluation technique known as the compensation distance evaluation technique (CDET) was introduced by Fatima et al. (2014) to select the most sensitive features for each rotor fault from the set of statistical features in the TD. A grid search approach was operated to optimize the hyper-parameters of SVM and applied for identifying unbalance as well as bearing faults. It was shown that irrespective of the number of transducers used, a certain level of accuracy could be attained through this method; therefore, it is applicable for online condition monitoring. To extract the fault features around the sub-harmonics and sup-harmonics of the rotational speed, Pang et al. (2018) introduced a novel index known as characteristic frequency band energy entropy (CFBEE). It was applied to the time-frequency spectrum (TFS) derived from raw vibration signals with improved singular spectrum decomposition (ISSD) and HT. This method attained higher accuracy comparing to the EMD-based characteristic frequency band energy entropy method. Martínez-Morales et al. (2018) performed feature data fusion by fusing electric data (stator current) and mechanical data (axial vibration) to identify the misalignment, unbalance, and the bearing faults of three-phase induction motors. OAO multi-classification SVM with RBF kernel performed the fault classification task. Similarly, in a recent work, Gangsar and Tiwari (2019) used vibration and current data for diagnosing mechanical and electrical faults of an induction motor, adopting grid-search methodology with cross-validation to select optimal SVM features for building the best model. Aydin et al. (2007) proposed a method of tuning kernel and penalized parameters of SVM using an artificial immune system and Park's vector approach for extracting features from three-phase motor current. Nguyen et al. (2008) applied GA for optimal feature selection from a set of TD features. SVM and DT were employed for analyzing the classification performance of unbalance, looseness, and bearing faults, concluding that SVM with the selected features works better. Fault specific symptom parameter extraction is the key element of SRF diagnosis. The nature and behavior of DFCs in rotor faults like misalignment, unbalance, and looseness was utilized by Xue et al. (2011) with SVM for fault diagnosis learning. This work was one of the initial attempts to make use of symptom parameters of SRF, addressing varying load conditions. They continued to use DFCs in their next work (Xue et al. 2013) in which they identified the SRF, performing normalization and extraction of DFCs using a multi-band pass filter. The filter range accounted for the change in frequency band under varying load conditions. Lately, Lobato et al. (2019) proved that data augmentation combined with GA for feature optimization could provide better accuracy of 95.19% even with significantly less amount of input data. It was observed that EEMD features with GA could improve the results by

17.31%, and, together with the augmented data, the improvement increased to 20.19%. For the detection of broken rotor bar fault in induction motor, Armaki and Roshanfekr (2010) introduced three new features like harmonic curve area, harmonic crest angle, and harmonic amplitude. These were extracted from power spectral density (PSD) of stator current using FFT, after which the performance of linear, polynomial, and RBF kernels was compared. They concluded that harmonic amplitude is not a useful feature because of its motor load dependency; meanwhile, among the kernels, the RBF kernel provides decent performance. To identify rotor crack under strong noise condition, Li et al. (2011) proposed feature extraction from amplitude and frequency of AE signal based on the pseudo Wigner-Ville distribution (PWVD) with SVM. This method was able to classify cracks of three different depths.

Some researchers demonstrated the dominance of SVM over state-of-the-art ML techniques. Singh and Kumar (2015) compared SVM with ANN in terms of classification accuracy with six statistical features derived from TD and concluded that the time taken to run the model by SVM was remarkably less when compared to ANN technique. In another similar work, Moosavian et al. (2014) compared the performance of the SVM and K-NN on unbalance rotor fault with 29 FD features and found that though KNN is faster than SVM, the latter shows substantially better performance in terms of accuracy. Ruiming and Hongzhong (2006) presented motor current signal analysis (MCSA) with SVM for induction motor rotor faults. FFT was used for FD feature extraction, and the performance comparison was made with majority voting, binary tree, neural network and hybrid matrix. There are some unique contributions to RFD from SVM, out of which more fault specific information or conclusions has been derived. In one such work, Fengqi and Meng (2006) showed that full-spectrum cascade analysis of acceleration signal remains ideal for some particular types of fault like rub impact. The relation with the amplitude of harmonic spectrum rub malfunction intensity was well explored in this work, and significant results were produced. Baccarini et al. (2011) demonstrated the OAA technique for SVM, in which they utilized only one vibration sensor and four SVMs to deliver improved results of classification. They observed that, for the best signal acquisition and analysis, the vertical sensor position is preferred. Another significant contribution in motor RFD was illustrated by Yuan and Chu (2006) in which 14 rotor faults were identified by multiclass SVM. FFT and AR models were used to design a feature space with nine frequency bands, and the PCA assisted in the dimensionality reduction process. In the context of broken rotor bar fault, Kurek and Osowski (2010) proposed an SVM based method that performed online detection of the presence of rotor fault and identification of fault affected rotor bars. Spectral information of the motor current, voltage, and shaft field of one phase was employed as input to SVM. They discovered that sufficient diagnostic information is present in the phase current for SVM based fault diagnosis of induction motor.

The research on RFD using SVM can be summarised as follows. First, it is worth noting that the researchers mainly attempt to change the kernel and incorporate new data structures to SVM to achieve a decent RFD accuracy. SVMs account for the maximum percentage (44%) of the share in ML methods for RFD analysis. One pertinent fact is that in terms of broken rotor bar faults, SVM is the traditional ML model with proven compatibility with voltage and current features. Moreover, as the literature states, the improved performance of SVM heavily depends on FD processing for feature extraction. SVMs succeed in utilizing information entropy fusion and spectrum analysis in certain applications, demonstrating its ability to deal with a host of input representations, while it is even used to evaluate the performance of feature extraction. Due to the excellent potential for handling outliers, SVMs can be used with any form of sensing method. Similarly, a number of the significant

contributions that made use of DFC in RFD involved using SVM for classification. In short, SVMs are the most versatile of the ML models from an RFD perspective.

4.3 k-nearest neighbor

k-NN is an instance-based, non-parametric algorithm renowned for its interpretability and ease of implementation. It is the most popular algorithm used in RFD after SVMs and ANNs.

Chen et al. (2011) proposed a k-NN based fault classification method for unbalance and bearing faults with raw TD vibration signature as input. The significant contribution of this work was the use of the 'maximum cross-correlation sum operator' as a similarity measure, which is shift-invariant and noise-tolerant. The proposed method achieved an error rate of 0.74% over varying operating speeds. Another approach using k-NN was proposed by Biet (2013), which dealt with faults associated with nuclear plant generators using rotor flux measurements and classical electrical measurements. Fisher criterion and the sequential backward selection algorithm were operated for feature selection from the scalar parameters. The k-NN with Euclidean and Mahalanobis distances was applied for classification, and it was able to achieve 85.1% accuracy in rotor faults. This work was a follow up of the previous work (Biet and Bijeire 2011), where they tried a large number of rotor fault combinations in which feature selection was made only from the radial flux density. Recently, Gohari and Eydi (2020) studied the identification of unbalance parameters of a rotating shaft having multi-discs with k-NN and DT. It is concluded that in terms of unbalance locating, KNN presents more accuracy in estimating of unbalance parameters compared to DT.

A more suitable fault specific method for finding broken rotor bar fault using the nearest neighbor (NN) algorithm was demonstrated by Karvelis et al. (2015). Start-up current was taken as input, and its transient analysis applied to the model using wavelet approximation to isolate the characteristic component of the specific fault. The input discretization was performed by maximum entropy partitioning (MEP) and represented with relative information created by an intelligent icon-like approach. Both simulation data, as well as experimental data, were used for testing, and the method produced high-classification accuracy. In another fault specific frequency selection based approach, Glowacz (2018) proposed a method for selecting essential frequency components as features and applied to k-NN along with BPNN as well as words coding based classifiers. It was based on computing the absolute values of differences in the frequency spectrum of acoustic signals, and a threshold was used to determine one or two groups of essential frequency components.

Nguyen and Lee (2008) utilized GA based on a distance criterion for feature extraction. By evaluating the features with k-NN and DT, they proved that selected features improve the accuracy and reduce the running time. The authors continued their work (Nguyen and Lee 2010) with an enhancement of applying weight factors in decision making. They implemented the weight factor for the selection of best features and evaluation of k neighbors. The performance was compared with SVM to prove the dominance of the proposed method. In the works of Son et al. (2009), Timusk et al. (2008), Moosavian et al. (2014), and Yang et al. (2015), k-NN was employed for comparison purpose and proven as one of the best-performing algorithms specifically appreciated for its faster and simpler operation.

The attendant discussion can be summarised as follows. First, while a k-NN is simple and convenient, it is not widely adopted in RFD. In fact, barely 10% of the works utilize k-NN despite its ability to deal with decision boundaries of any form. A number of

researchers developed different similarity measures derived from well-known distance functions, while others operated well-established algorithms such as GA to boost the performance of k-NN. However, the inability to recognize important attributes, the overheads involved in deciding the parameter 'k,' and the interpretability issue due to its non-parametric nature have been identified as the main limitations of this algorithm, which make researchers reluctant to adopt it for RFD analysis.

4.4 Naïve–Bayes

Naïve–Bayes is the most popularly used Bayesian model in RFD, which works on the conditional probability basis. Wang et al. (2012) used envelope features of the motor current extracted using HT and compared it with the conventional motor current features using Naïve–Bayes as well as two other classifiers. Current envelope features, AR coefficients, statistical features, and FD features were applied to mRMR (min redundancy max relevancy rule) for proper feature selection. The results of three different classifiers validated the effectiveness of current envelope features in this work.

Bayesian modeling has a drawback that it is unable to model and learn from the time-series level change of data. As an attempt to mitigate this disadvantage, Yusuf et al. (2013) presented the NB classifier on the fault groups created by the hidden Markov model (HMM). The results showed that the false positives and negatives were reduced with an identification accuracy of 84.55% by the probabilistic belief-based identification capability of Bayes theorem. Duan et al. (2016b) applied image segmentation and histogram processing with PCA dimensionality reduction to the NB classifier. They formulated a region selection criterion named 'dispersion degree' to discriminate fault representative regions. The performance of the algorithm was evaluated using the original image, segmented image analysis, and segmented image analysis with feature reduction methods.

Bayesian Belief Network: In the field of rotating machinery fault diagnosis, incorporating human expertise is essential, especially when there is little one-to-one correspondence between symptoms and faults. Bayesian belief network (BBN) is a Bayesian model that is useful in such situations typically when dealing with compound faults. Xu (2012) developed a BBN with three layers, namely machine running conditions layer, machine faults layer and fault symptoms layer with two topological configurations of causality, and fault symptom. Compared with the traditional Naive Bayesian network, the proposed BBN contained two topological configurations of causality, which concerns the information not only about the fault symptoms but also the machine conditions. The designed system could give reasonable uncertainty inference in rotor faults to cop up with practical knowledge and experience of experts.

Here, the attendant analysis can be summarised as follows. First, as a probability-based classifier, an independence assumption is the core aspect of the NB algorithm. Meanwhile, the higher chance of correlated features in the RFD dataset results in certain restrictions to its application. The inability to model the time-series data and to learn from it, as well as the issues created in the classifier due to the absence of data, are just some of the limiting factors. However, attempts such as HMM to deal with time-series and mRMR for feature extraction have been made to return some ground-breaking results. In the Bayesian family, the importance of BBN cannot be denied as it is capable of dealing with most of the aforementioned problems. In short, the attribute correlation and the probability of missing the RFD sensor data can be given as justifiable reasons for the nominal participation of NB in RFD.

4.5 Fuzzy logic and other methods

The ability of fuzzy logic to mimic human reasoning has been utilized in RFD by certain researchers. El-Shafei et al. (2007) used the LVQ neural network in parallel to a fuzzy inference engine for addressing the SRFs. The whole harmonics and half harmonics up to 6x were utilized in fuzzy decision logic and proved the successes of fuzzy as a low-resolution configuration diagnostics tool for spectral data. They found the neuro-fuzzy system as a good option that provides a confidence index for faults. The fuzzy Sugeno model adaptive neuro-fuzzy inference system (ANFIS) was used by Lei et al. (2008) for rotating machinery fault diagnosis. They employed TD, FD, and TFD features along with an improved distance evaluation technique for feature selection. The results showed the ability of the system by providing high accuracy for finding bearing faults, fault severity and category. Another work by them (Lei et al. 2007), that combined multiple ANFIS with GAs for a more reliable and intelligent RFD system was also reported in the literature. Zhang et al. (2015) introduced a fuzzy support vector machine (FSVM) by integrating fuzzy logic with SVM. Qiu and Rao (2005) summarized a significant analysis of unbalanced non-linear rotor systems involving uncertain parameters and concluded that the fuzzy approach could be considered as more versatile and robust in system design and analysis. Lately, fuzzy neural network technology application in equipment fault diagnosis was proposed by Xu et al. (2020). In this work, network learning weights are transformed into case-based reasoning-based diagnostic guidance operators.

Among the remaining ML algorithms, the RF was used by Yang et al. (2015) with multiple class feature selection (MCFS), and Quiroz et al. (2018) used the same algorithm with TD statistical features. The LDA algorithm was explored for RFD with acoustic signal input by Glowacz (2018), and DT was used by Nguyen and Lee (2008). A vibration image-based diagnosis was proposed by Yan et al. (2019b) in which three features such as histogram of vibration image (HVI), histogram of oriented vibration image (HOVI), and 2-D FFT generated from vibration images were used. The AdaBoost proposed in this work overcome the overfitting problem, and it fuses multiple features. Martin-Diaz et al. (2018) selected AdaBoost for comparative analysis to prove its significance in RFD analysis.

Here, we can summarise the discussion as follows. The decision making based on fuzzy logic has advantages over other methods since the rules derived from the fault state play the key role. As the fuzzy classification system does not have the capacity for pattern recognition as other ML models do, and since certain concerns exist in terms of its implementation, it is not widely accepted in RFD scenarios. However, individual attempts to utilize the harmonics of the rotating frequency as the input to the fuzzy system and the use of ANFIS have returned some significant result improvements. Another way of utilizing fuzzy logic to overcome its limitations is to incorporate it with other ML methods, which has, in fact, been adopted in a number of works related to RFD.

4.6 Classifier fusion

The conventional fault diagnosis method of using a single information source together with a unique decision method exhibits certain shortcomings. Niu et al. (2007) proposed a decision fusion by finding the optimal sequence of classifiers' fusion. It is based on selected decision vectors using the correlation measure of classifiers and sensor fusion method using relativity theory. SVM, LDA, k-NN, improved iterative

scaling (IIS), Gaussian mixture model (GMM), and LVQ classifiers were used, and multi-agent classifiers fusion algorithm was engaged for fusing them. The comparison results with majority voting and Bayesian belief classifiers proved the superiority of multi-agent fusion. The same authors continued the fusion method in (Niu et al. 2008) using multi-level wavelet decomposition with transient current as the input. After the feature extraction using wavelet-based decomposition, the signals were applied to the two decision-level fusion strategies, the multi-agent fusion, and Bayesian belief fusion. The results demonstrate that the proposed framework fusing SVM, ART-KNN, LDA, and k-NN classifiers could achieve an accuracy of 98.5%. In contrast, the best individual classifier could achieve only 90% accuracy.

Santos et al. (2012) combined bagging, AB, general boosting projection (GBPC), and RF classifiers to obtain an ensemble classifier for performing unbalance and misalignment classification of wind turbines under various speed and load conditions using current and vibration signals as inputs. The results of the proposed method compared with C4.5 DTs, k-NN and NB, and found that the AB using J48 DTs as base classifiers achieved the highest accuracy of 96.24%. Tao et al. (2016) proposed a novel classifier ensemble technique known as weighted majority voting with a different confidence level to ensemble NB, RF, and SVM classifiers. Based on the vote through the confidence diversity, they assembled multiple classifiers, and the results were compared with the conventional normal weighted majority voting method. They observed that the performance of the NB classifier was low as compared to the progress made by RF and SVM in terms of classification accuracy. Moreover, the proposed ensemble method gave significantly improved accuracy compared to the conventional weight-based voting method, especially in broken rotor bar fault. The AB ensemble classifier presented by Martin-Diaz et al. (2017a) addressed the issue of imbalance data in rotating machinery fault diagnosis. TD and FD features extracted from stator current were applied to variable filtering through correlation analysis for feature selection. SMOTE sampling technique was followed by a cross-validation technique so that it ensured a rebalancing of data and generalization of the results. AB classifiers with DT (CART) and the NB classifiers were utilized for performance evaluation in the presence of varying sized datasets and imbalance ratios. The results demonstrated the superiority of the AB ensemble classifier in RFD of induction motors.

The fusion of classifiers in RFD literature can be summarized as follows: The works that performed classifier fusion constitute only less than 5% of overall ML literature of RFD. SVM, k-NN, and NB are often found among the fusing classifiers. The different fusing methods like multi-agent fusion, Bayesian belief fusion, decision level fusion, majority voting, and its variants, have been used in most of the works, and their performance comparison has been carried out. The literature proves that classifier fusion in RFD can address multiple drawbacks of individual classifiers, but still, the research in this direction has a long way to go.

4.7 Studies on classifier comparison

There have been certain efforts from researchers to study and compare several ML methods that will help in the selection of the most appropriate classifier according to the requirement. The summary of such classifier comparison studies is given in Table 3. In one such work, Gang et al. (2007) compared the performance of five classifiers with ten TD, three FD, and eight AR features on acceleration, voltage and current, and flux signals of an

Table 3 Summary on ML comparison studies

References	Monitoring technique/Input source	Signal processing	Classifiers (% accuracy or score/5)						Faults
			SVM	ANN	k-NN	NB	RF	Others	
Gang et al. (2007)	V&C, Flux density, vibration/induction Motor	TD, FD, AR	99.4		96.26		100	LDA: 93.54, ART-KNN: 82.3	UB, MA, BS, BRB, OF
Timusk et al. (2008)	Vibration, transient signals/rotor Machines	TD, AR	84		85			K-means: 84, AE: 88, PCA: 92, Combined: 93	UB, BRB, OF
Son et al. (2009)	Vibration, flux, current/induction motor	TD, FD, AR	84.76		100		82.33	LDA: 100	UB, MA, BS, BRB, OF
Wang et al. (2012)	Motor current envelope/induction motor	HT, TD, FD, AR	91.43		96.15	91.43			UB, BS, BRB, OF
Cunha Palácios et al. (2015)	Stator current/NI 6221 Board	TD	85.75	96.5	96.1	69.35		C4.5-DT: 92.75 RIP: 91.55	UB, MA, BRB, OF
Godoy et al. (2016)	Stator current/induction motor	TD	92	93.7	94.9			Fuzzy : 85.88	BRB
Martin-Diaz et al. (2017b)	1-Ø Stator current/Induction Motor	TD, FD	3		3	5			BRB
Martin-Diaz et al. (2018)	1-Ø stator current/induction motor	TD, FD	2	4	2	3		Bagging: 5, AdaBoost: 5	BRB
Li et al. (2018)	Vibration/RTB	WPD	4	5	5				UB, LS, RIF
Quiroz et al. (2018)	Transient current/LS-PMS Motor	TD	93.8			96.9	99.6	CART: 98.8, Linear ridge: 93.9, LR: 94	BRB
Konar et al. (2015)	Vibration/induction motor	RST, GA, CWT, HT	99.15	97.23	98.86			Simple Logistic: 96.96	UB, BS, BRB, OF

V&C: Voltage and Current, TD: Time Domain, FD: Frequency Domain, AR: Autoregressive, HD: Hilbert Transform, WPD: Wavelet Packet Decomposition, RST: Rough Set Theory, GA: Genetic Algorithm, CWT: Continuous Wavelet Transform, ART-KNN: Adaptive Resonance Theory- Kohonen Neural Network, AE: Autoencoder, RIP: Repeated Incremental Pruning, LR: Logistic Regression, DT: Decision Tree

induction motor. They concluded the order of classifier performance as RF, SVM, k-NN, LDA, and ART-kNN (Adaptive resonance theory-Kohonen neural network), and identified flux2 as the excellent performing signal for classification. Timusk et al. (2008) experimented with transient but repeatable operating data using vibration information alone, taking into consideration time-series statistics and AR models. Eight different classifiers were tested on motor and gear datasets with and without re-sampling along with the Gabor feature statistics. The study revealed the failure of Gabor feature statistics as a feature in rotor faults. A smart sensor system architecture that acquired vibration, current and flux signals from induction motors was developed by Son et al. (2009). The SVM, LDA, k-NN, and RF were used for dealing with the fault classification section. The results showed that LDA and k-NN achieved high accuracy and performance stability, where RF performed the worst. Wang et al. (2012) worked with the envelope of the motor current and the order of classifier performance in terms of accuracy was k-NN (96.15%), NB (94.5%), and SVM (91.43%).

Cunha Palácios et al. (2015) conducted a meticulous review on machine fault analysis, including stator and rotor faults with NB, k-NN, SVM, ANN, repeated incremental pruning (RIP), and DT classifiers. Considering the rotor faults alone and in combination with other faults under the varied power supply and mechanical loading conditions, they observed that ANN and k-NN could perform with the highest accuracies 96.5% and 96.1%, respectively, where NB could only get an accuracy nearing 70%. Godoy et al. (2016) employed fuzzy ARTMAP network, SVM, k-NN, and MLP to classify broken rotor bar defects in inverter-fed induction motors analyzing the TD stator current signal, and the rate of acquisition was based on machine frequency supply. Over a wide range of frequencies and load conditions, k-NN performed the best followed by MLP, and the fuzzy-based classifier was the weakest performer.

Martin-Diaz et al. (2017b) worked on the current signals from an inverter feed induction motor with one broken rotor bar under changing operational conditions. Various TD and FD features were extracted, and feature selection task was performed with the help of the Pearson correlation coefficient. NB was able to classify the faulty data with only two instances misclassified showing the same behavior in both classes. But SVM and k-NN categorized all instances to a particular class only. The same authors (Martin-Diaz et al. 2018) continued the work by additionally including the classifiers AB, ANN, and bagging. Two load conditions and two faulty conditions like half-broken rotor bar and broken rotor bar were considered on an overall eight cases, resulting in a classifier ranking as follows- AB, bagging, MLP, NB, SVM, and k-NN. Li et al. (2018) has worked with WPD based feature extraction for classifiers BPNN, SVM, and k-NN on looseness, friction, and load unbalance faulty conditions. In the classification process, BPNN achieved 98.7% accuracy with WPD, and in degradation assessment, k-NN showed the best performance.

The start-up transient current signal of line start-permanent magnet synchronous motor (LS-PMSM) was introduced by Quiroz et al. (2018) for identifying the broken rotor bar fault using RF classifier. They compared the performance with the other five algorithms using 13 statistical TD features. The experiments were conducted with different number of trees within RF and three reduced features. RF performed well with 99.6% accuracy, which proved to be independent of the number of trees. Konar et al. (2015) applied signal processing techniques such as CWT, HT for feature extraction, and PCA, GA and RST for dimensionality reduction. They proved that powerful signal processing tools, in conjunction with the data mining technique, could reduce the burden on SVM. The classification performance turned out to be SVM, the best performer with an accuracy of 99.15%, followed by k-NN, giving 98.86%. MLP algorithm showed 97.23%, and finally, simple

logistic showed the least performance with 96.96% when worked with six types of induction motor faults.

5 Deep learning-based approaches to RFD

The abundance of data and the evolution of new algorithms supported by the GPU-like hardware expansions have accelerated the development of DL, which is essentially a nested hierarchical and more abstract representational framework. The advantage of the automatic feature extraction of DL models made them widely accepted within the field of machine fault diagnosis since this allows for avoiding the requirement of high domain expertise in feature engineering. In fact, RFD makes use of a variety of DL models, including the most popular CNNs, autoencoder based models such as SAE and SDAE, sequential models such as LSTM and GRU, the generative model like GAN, generative and probabilistic models such as DBN, and other deep neural networks (DNN). Indeed, the majority of the DL algorithms have gained traction in a short space of time. The commonly used DL algorithms, along with their features, pros and cons, relevance, and compatibility with RFD are presented in Table 4.

5.1 Convolutional neural network

The CNN is established as a convolution-based hierarchical image pattern recognizer and is the most widely used DL model in RFD. The discrepancy caused by using 1-D time-series signals in a 2-D convolution-based network is solved by different methods such as 1-D to 2-D conversion or by introducing 1-D CNNs.

In one of the early RFD attempts, Janssens et al. (2016) addressed rotor misalignment fault using various levels of lubricant degradation and other common bearing faults. They presented their work to manifest the improvements of feature-learning based approach over feature-engineering based approach by quantifying CNN's performance over linear regression based on handcrafted features. The proposed CNN contained 32 convolution filters and 200 hidden units in the fully connected layer with scaled DFT of raw data as input, and it achieved an accuracy of 93.61% over 87.25% of the conventional model.

CNN is specifically designed for handling 2-D data; therefore, several image representations of input data have been widely used with CNNs. Jeong et al. (2016) proposed orbital image-based diagnosis using CNN for identifying unbalance and misalignment. Orbital images necessitate specific preprocessing such as denoising, offset shifting, re-orienting, and rescaling before applying it to CNN. The proposed CNN was formed by a stack of two convolution layers, two max-pooling (sub-sampling) layers, followed by fully connected layers. This framework attained significantly good results in classifying orbital images. One of the successful methods of inputting data for structural fault analysis of the rotor is the IR images. CNN proved to be effective in such a scenario, by Liu et al. (2017b) through comparing CNN's performance with that of SVM on different sized input images. SVM with features extracted from the histogram of thermal images was unable to demonstrate the performance at par with CNN in terms of accuracy. A data fusion method for converting data into RGB channel false-color image and the position-based polar image was proposed by Yao et al. (2018). First, they had created the simulation of fault in the MAT-LAB/Simulink environment in which harmonics with frequencies 1x, 2x, 4x, and higher values were injected to simulate characteristic components of certain faults. The raw data

Table 4 Summary on DL classifiers

Classifier	Properties	Pros	Cons	Compactibility and relevance with SRF
CNN	# CNN is layered feedforward NN expects 2-D inputs # By convoluting filters, the patterns across an image are extracted # By weight sharing, number of parameters reduced	# Compared to NN, less complex network in terms of count of neurons # Can handle noisy and non-linear input # Many proved architectures are available openly	# Poor generalization ability, # Network complexity in terms of no. of layers is high. # The cost of computation is also increased	With an appropriate mechanism to represent time series data as a 2-D image, CNN will be a good choice. With a large amount of input data and with modified CNN architectures, it can give results in RFD
AE based Models: SAE,SDAE	# Contains an encoder(input, hidden layers) which reduces input dimensionality, followed by a decoder to recreate input # Stack of AE is called stacked AE (SAE) and if it takes corrupted input then it is called stacked denoising AE (SDAE)	# No need for prior or expert knowledge # Can deal with non-linear data even if it is complex # Can be used as a feature extractor for fault data	# Requirement of pretraining and validation # Demands more training data, processing time and hyperparameter tuning # Chances of sparse representation	No signal processing needed for input, with raw as input it guarantees to deliver features most sensitivity to faults. SDAE is recommended in noisy signal environments in RFD diagnosis.
Sequential Models: RNN and Variants: LSTM, GRU	# An NN with a cyclic structure in hidden units such that weights can be shared in the whole network # By recursive structure, it holds past information and uses back-propagation to update weights # LSTM and GRU are two variants with Gates to avoid vanishing gradient problem	# Deep and sequential model # Can exploit temporal relationship of data # Performs well with short-term information # Deal with non-linear data	# Fail in processing very long sequence with tanh activation and becomes unstable with ReLU activation # Large data needed for training. # Suffers vanishing gradient problem and cant keep long dependencies	Mostly RFDs can not provide sufficiently huge data that RNN demands a decent performance. Moreover, RFD faulty data is not depending too much on its sequential property to make decisions. So RNN is not the optimum choice

Table 4 (continued)

Classifier	Properties	Pros	Cons	Compactibility and relevance with SRF
Generative Models: DBN	# An unsupervised probabilistic DL network with multiple layers of stochastic latent variables # Multiple RBMs or AEs have stacked together this generative hybrid graphical model # DBN are pre-trained using Greedy learning algorithm	# No need for prior or expert knowledge # Can deal with non-linear data # Temporary coherence of time series is considered # Has strong representational ability and deals with local optimal parameter fixing problem	# Training requires initialization and sampling, which demands higher computation # Optimization process consumes much time # Certain signal processing methods used with DBN	DBN is used as a fault feature extractor that can be utilized in RFD. Compared to other black-box DL methods, DBN gives provision to apply fault features as input. So for RFDs, this is a good option
GAN	# Input reconstruction based semi-Supervised learning method # Contains two parts, generator(G) and the discriminator(D), where D tries to find the difference of signals generated by G, with original samples # Useful in generating additional data in fault diagnosis scarce data situations	# Works well with less data compared to other DL models # Learn the internal data representations with its distribution # Generate samples faster than fully visible belief nets	# Huge computation requirement for training # GAN training is unstable compared to VAE # Discrete data generation is harder.	Generality issues in DL due to data unbalance, scarce data, unlabelled data situations, which are common in RFDs can be solved using GANs. As a data upsampler for generating faulty data, it is highly recommended

collected in this way was then converted to color and polar images in order to connect with the practical working conditions. The method is validated with a ball screw testbed with misalignment fault.

There are several ways by which vibration data has been applied to CNNs in 2-D form. Symmetrized dot pattern (SDP) image was proved as one of the convincing ways by Zhu et al. (2019a). Normalized time waveform was mapped into an SDP on a polar graph to produce SDP images, which was demonstrated as an effective method of representing time-series data from multiple sensors. The effectiveness of using the SDP image was manifested by applying it to a CNN with two convolution and pooling layers and one fully-connected layer with faulty data on unbalance, misalignment, rub, and oil whirl. The same authors developed another method using SDP images and CNN utilizing characteristic information fusion (Zhu et al. 2019b). They operated the Hilbert vibration decomposition (HVD) technique to decompose the original signal to get modal components and then fused them based on the information fusion of SDP. A CNN with two convolution and pooling layers and one fully connected layer achieved 97.7% precision, indicating that the proposed fault diagnosis method was highly reliable. Another approach of converting raw data (that are sampled into a phonetic form) to 2-D image input for the learning of CNN, based on the spectrogram, was proposed by Yu et al. (2018). The power spectral function of the signal was converted to a 2-D image form and was applied to a batch regularized CNN with VGG16 architecture. The effectiveness of the proposed method was demonstrated by comparing its performance with BPNN, SVM, and ELM. An interesting approach for providing a 2-D representation for vibration data is the continuous wavelet transform scalogram (CWTS), which was proposed by Guo et al. (2018). CWTS was designed by discomposing vibration signals in different scales using wavelet transform with certain preprocessing, and then it was applied to parallel CNN for acquiring the scores to select the appropriate fault model. The model was applied to different testbed data and demonstrated that the proposed approach was a universal and generic solution. Dislocated time-series CNN (DTS-CNN) proposed by Liu et al. (2017a) to handle disparity of time-series industrial data (most often in 1-D form) with 2-D images, thereby helped CNNs to excel well as it did in other domains. DTS-CNN adds a dislocate layer that extracts the relationship between signals with different intervals in periodic mechanical signals. The dislocated data matrix was applied to a CNN architecture with four convolution layers, two pooling layers, and two fully connected layers resulting in superior performance over the other methods, especially in non-stationary data. Recently, Yongbo et al. (2020) developed a method using CNN that extracted fault features from the IR thermal (IRT) images and fed it into the softmax regression classifier for fault pattern identification. The proposed method could achieve an accuracy of 98.59%, outperforming DBN, DNN, and SAE. In another similar work, Liu et al. (2020), used shaft orbit images as input to the CNN. This generated the first set of features, while DWT transform of denoised swing signal and SVD created the second set of features. Finally, the CNN and DWT-SVD combined features applied to the SVM classifier to attain a better recognition effect than other methods. Presenting the vibration signal in the form of a 2-D Kurtogram to CNN was also observed in some bearing fault analysis works using CNNs (Udmale et al. 2019a, b).

Most of the CNN architectures reported in the industrial fault diagnosis system either deal with 2-D plots or convert raw data to a 2-D matrix for operational convenience. But Ince et al. (2016) introduced adaptive 1-D CNN, which used 1-D arrays instead of 2-D matrices for both kernels and feature maps. In this method, the 2-D convolution and rotation operations were replaced with corresponding 1-D operations. Moreover, instead of applying raw motor current signals to CNN, a notch filtered current spectrum was applied

as input to suppress power system fundamental frequencies. Wang et al. (2018) demonstrated another notable work utilizing 1-D CNN. They proposed an efficient way of fusing multi-sensor data for multiple motor faults with a multi-model DL containing 1-D CNN and MLP. They applied the frequency spectrum of current and vibration signals with TD values as input to the learning model. The current envelope analysis, performed by HT helped the model to identify sideband harmonics of faulty conditions. The experiments proved that the method was efficient and robust not only in rotor faults but also in bias sensor faults in varying load conditions. Vector CNN (VCNN) developed by Xiaoxun et al. (2019) was another 1-D CNN based approach. It took the original 1-D vibration signal vector into the model, avoiding the information loss observed in the signal processing operations. The model was tested with rotor rubbing, unbalance, misalignment, and whirl faults on a testbed. Liao et al. (2019) proposed a 1-D CNN-GRU model for hydroelectric generating unit (HGU) fault diagnosis in order to combine the speed and lightness of convnets with the order-sensitivity of GRU. The framework followed a three-layer structure consisting of the input layer, the hidden layer, and the output layer, where the hidden layer contained the convolution layer, pooling layer, GRU layer, and the fully-connected layer. GRU, combined with CNN, was able to achieve significant improvement in the results. A recent contribution in 1-D CNN was presented by Zhao et al. (2020a) to identify shaft misalignment and crack in the rotor system. The proposed multi-input convolutional neural network (MI-CNN) took the raw TD input signal and applied to 1-D CNNs, in which kernels of different sizes were adopted to extract the signal features of diverse dimensions. The method could achieve an accuracy of 99.42% when the fused features were applied to softmax function for classification.

Modifying CNN structure is another way by which the researchers utilize CNN in order to adapt to specific requirements of the RFD problem. Sadoughi and Hu (2019) proposed a physics-based CNN (PCNN) which modified the CNN structure by incorporating physical knowledge about the faulty component and fault characteristics. They achieved it by designing a convolutional layer that functioned based on the similarity between the input signals (data) and the fault characteristics signals (physics). The overall structure of PCNN contained five convolutional layers, three max-pooling layers, and one global max-pooling layer, as well as spectral kurtosis (SK), envelope, and FFT function layers around the first convolution layer. The CNN proposed by Zhang et al. (2017) with wide first-layer convolutional kernel and deep network structure with small convolutional layers is known as WDCNN (Wide first-layer kernels deep CNN). The most important property of WDCNN is its ability to suppress high-frequency noise while taking raw input data directly. This work utilized data augmentation to increase the number of inputs in order to improve generalization ability. It was able to achieve 100% accuracy in the public CWRU bearing dataset. Different CNN models and architectures used in RFD are shown in Table 5.

Along with structure alteration, there were several attempts to make CNN suitable for multi-sensor data through data fusion or feature fusion methods. For integrating automatic feature learning, adaptive feature fusion, and accurate fault classification of data from multiple sources, a multi-stream CNN was proposed by Yuan et al. (2019). In the feature learning layer, 1-D convolution was employed for the raw vibration signal. Then 2-D convolution was used for the infrared image, followed by a feature fusion layer and a softmax layer for classification. Misalignment, rub impact and looseness fault data from a testbed was used for demonstrating the model, and the results were compared with BPNN. Sonkul et al. (2019) developed multi-channel CNN (McCNN), which accepted multi-channel input from bearing, shaft, and gearbox with sensors arranged in horizontal and vertical directions. It took raw TD data and needed little preprocessing. Yuan et al. (2018) proposed a novel

Table 5 Summary on CNN architecture

References	Model	Architecture overview	Faults	References	Model	Architecture overview	Faults
Jeong et al. (2016)	CNN	Conv1: (7X7) #32, Conv2: (7X7) #64, Pool1: 2X2, Pool2: 4X4 FC: 1024 Actv. Fn: Sigmoid	UB MA	Sadoughi and Hu (2019)	PCNN	Conv1, FFT layers + 4 Convolution layers, 3 max pooling layers, one global max pooling layer	UB MA CF OF
Janssens et al. (2016)	CNN	Conv1: (2X64) #32 FC: 200	UB CF OF	Liao et al. (2019)	1-D CNN, GRU	I/p Layer, ConvLayer, Max Pool Layer, GRU Layer, FC Layer, Output Layer. ReLU in CNN and softmax in GRU	UB CF OF
Yuan et al. (2018)	M-CNN	1-D structure and 2-D structure. Adaptively decided	UB MA LS RIF CF	Yuan et al. (2019)	Multi-stream CNN	Conv1: (24X24)# 6, Conv2: (12X12)#6, Conv3: (8X8)#12, Conv4: (2X2)#12, Pool: 2X2, FC: 432, 1-D CNN: 3 Conv+ 3P layrs	UB MA LS RIF CF
Liu et al. (2017a)	DTS-CNN	Conv1, Conv2: (3X4)#16, P1: 2X2, Conv3, CNV4: (3X4)#32, P2: 2X2, FCs: 2, Softmax, Actv. Fn: ReLU	UB MA BS OF BRB	Lu et al. (2019b)	Enhanced CNN	Conv1: (3X3)#16 Conv2: (3X3)#32, Pool: 3x3(2Nos) , FC: 128	UB CF OF
Han et al. (2019b)	DACNN	Conv1: (64X 1)#32, Pool1: (4X1), Conv2: (34X3)#128, Pool2: (8X1), FC1: 120, FC2: 64, Softmax, Dropout: 0.5%, Actv. Fn: ReLU + 2 models of 16X1 i/p & 3, 5 conv-Pool layers	MA LS OF	Zhang et al. (2018)	SVM + CNN	Conv1-1: (3x3)#32, Conv1-2: (3x3)#32, Conv1-3 (3x3)#32, Pool1: 4x8, Conv2-1: (3x3)#32, Conv2-2: (3x3)#32, Conv2-3: (3x3)#32, Pool2: 4x8, Fc3: 64, Fc4: 7, Actv. Fn: ReLU	UB LS
Yao et al. (2018)	CNN	Conv1: (3X3)#32, Pool1: 2X2 Conv2: (3X3)#64, Pool2: 2X2 FC: 5, Softmax, Actv. Fn: ReLU	MA	Xiao et al. (2019b)	CNN + MMD reglzn.	Conv1: (64X1)#64, P1: 2X2, Conv2, Conv3, Conv4: (3X1)#32, FC1: 128, FC2: 64, FC3: 6	UB MA BS BRB OF

Table 5 (continued)

References	Model	Architecture overview	Faults	References	Model	Architecture overview	Faults
Liu et al. (2017b)	CNN	Conv1, Conv2: (5X5)#6, Pool1, Pool2: (2X2)#6, Conv3: (4X4)#12, Pool3: (2X2)#12, Conv4: (3X3)#12, Pool4: (2X2)#12 , FC: 1200, Soft-max, Activ. Fn: Sigmoid	UB MA LS RIF	Sonkul et al. (2019)	McCNN	Conv1: (40X2)#8, P1: 3X1, Conv2: (20x2)#16, P2: 4X1 Conv3: (4X2)#32, P3: 3X1 Conv4: (4X2)#32, P4: 2X1 Conv5: (2X2)#32, P5: 2X1 Conv6: (1X2)#32, P6: 2X1 FC: 12, Softmax, Activ. Fn: ReLU	UB LS BRB OF
Xiaoxun et al. (2019)	VCNN	Two conv-pool layer pairs and one FC layer	UB MA RIF	Zhu et al. (2019b)	fusion+ CNN	Conv1: (5X5)#6, Conv2: (3X3)#12 P1, P2: 2X2, one FC, softmax, Activ. Fn: ReLU	UB MA BRB
Zhang et al. (2019a)	CNN	Conv1-3 (3x3)#32, Pool1: 4x8 Conv2-1: (3x3)#32, Pool2: 4x8 Fc3: 64, Fc4: 7, Activ. Fn: ReLU	UB LS	Langarica et al. (2019a)	CNN	Conv1: 20 filters Conv2: 40 filters Pool1, Pool2: 2X2, FC1: 100, Softmax, Activ. Fn: ReLU	UB OF
Zhu et al. (2019a)	CNN	Conv1: (5X5)#6, Pool1: (2X2), Conv2: (5X5)#16, Pool2: (2X2), FC 84, softmax, Activ. Fn: ReLU	UB MA RIF	Wang et al. (2018)	1-D CNN, MLP	1-D CNN, Activ.Fn: ReLU, Final classifier: Softmax	BS BRB OF
Han et al. (2019a)	ST-CNN	Conv1, Conv2, Conv3: (3X3)#n1, Max Pool: 2X2(3), FC-3Nos, Activ.Fn: ReLU, softmax	UB MA LS OF	Liang et al. (2019a)	WT-MLCNN	Conv1, Conv2, Conv3: (3X3)#32, Conv4, Conv5, Conv6: (3X3)#64, Max Pool: 2X2, One FC, softmax	UB LS CF OF

Conv: Convolution, P: Pooling, #: No. of channels, Actv. Fn: Activation Function, OF: Other Faults, CF: Compound Fault

multi-mode convolutional neural network (M-CNN) for the fusion of both structured and unstructured, heterogeneous data. Adjustable filter banks were introduced to decide appropriate convolution mode, and t-SNE was applied to fuse the deep features. On raw vibration as well as infrared image input data, CNN achieved a classification rate of 98.97% on almost all important SRFs. The convolutional pooling architecture on CNN operated as a feature learning tool for specific other ML models. Sun et al. (2017) developed such a framework in which the local filters convolutional pooling architecture was learned by backpropagation neural network to give it discriminative power so that it derived invariant and robust features. Finally, SVM classified six faults of an induction motor with the best performance among state-of-the-art methods.

Moreover, improving the generalization ability, even with scarce labeled samples, is one of the major issues in applying DL to SRFs. Zhang et al. (2018) proposed the usage of handcrafted features generated by STFT and the labeling of unlabelled data with SVM. A CNN with six convolution layers, two pooling layers, and two fully connected layers designed to discriminate different unbalance conditions and looseness with better accuracy and generalization ability. The same work was modified (Zhang et al. 2019a) by applying the features from raw data to several candidate SVMs for selecting the best performing SVM, and it assisted in creating an augmented training set (ATS). The CNN is designed with two sets of convolution, pooling, and fully connected layers to improve the fault diagnosis results, and tested with two different datasets. Another approach towards solving the issues of insufficient and extremely imbalanced, distributed data with DL models, was presented by Xu et al. (2019) with a small data-driven convolution neural network (SDD-CNN). They adopted the label dilation method to solve the imbalance in data distribution and a semi-supervised data augmentation method for solving scarce data issues. Among the four variants of architectures demonstrated by them, SDD-inception v3, using a deep transfer learning strategy, proved to be superior with 99.56% accuracy. Transfer learning is effectively used for the same purpose in certain works. Xiao et al. (2019b) proposed transfer learning by CNN embedded with maximum mean discrepancy (MMD) to solve the DL issue of handling the unlabeled data and the unlabeled data under constantly varying working conditions (which are commonly available from real applications). The regularization term of MMD ensured CNN parameters to learn to reduce the distribution mismatch between the features in the source and target domains. The CNN was designed with four convolution layers, one pooling layer, and four fully connected layers, and was tested on the European driving cycle (NEDC) data. The results revealed that the proposed method achieved higher accuracy than the other methods.

Hardware associated approaches with CNN that directly deal with the industrial internet of things (IIoT) based hardware platforms, as well as with other embedded systems, have also been observed in RFD literature. An IIoT based application development for rotor fault diagnosis was demonstrated by Langarica et al. (2019a) with dynamic incremental PCA (DIPCA) and reconstruction-based contribution (RBC). In this approach, a statistical data-driven algorithm (for preliminary detection and identification of faults) was combined with a signal-based method that utilized a CNN for refined fault diagnosis of vibration signals. The CNN, with two convolutional layers of 20 filters and 40 filters, respectively, two pooling layers and softmax output function, achieved an identification accuracy of over 90%. The IIoT components and design principles were well defined and followed in this framework. In another hardware related approach, Lu et al. (2019b) implemented an embedded system with Raspberry Pi and signal acquisition and processing circuit for evaluating the performance of the proposed CNN in real-time data. The CNN input image was created from a vibration signal by base two logarithmic spectrum (B2LS), and the proposed CNN

structure included two convolution layers, two max-pooling layers, one fully connected layer, and a softmax layer. This work proved to be a heterogeneous computing framework for motor fault diagnosis on a small-size, flexible, and convenient handheld device.

Adding the discriminative ability to the existing classifier improves the performance to a great extent, as seen in (Sun et al. 2017) (where CNN gave discriminative power and SVM was the classifier). An almost same approach, with an adversarial network as discriminator and CNN (DACNN) as the classifier, was proposed by Han et al. (2019b). DACNN merged the idea of adversarial training into the feature learning of deep model, i.e., the parameters of trained CNN with small datasets were transferred for training with the adversarial strategy. By experimenting with a wind turbine test rig and a generic gearbox fault dataset, they proved that their method was effective and superior even with smaller datasets. The same authors (Han et al. 2019a) proposed a way to extract spatial (between measurements) and temporal (for each measurement) features using a spatio-temporal pattern network (STPN). It provided an efficient way to apply multivariate time-series to 1-D CNN in single-channel or multi-channel mode with a complexity less than that of a 1-D CNN. The authors experimentally proved that the spatiotemporal CNN (ST-CNN) outperformed SL methods like SVM and RF. Besides, dealing with compound faults is one of the research challenges still kept least addressed in RFD. Liang et al. (2019) made such an attempt with an end to end diagnosis model integrating CNN with wavelet transform and multi-label classification, namely WT-MLCNN, which directly used raw vibration signals. The main attraction of the work is that, rather than considering a compound fault as a new single fault, the model considered it as a multi-label image classification problem by converting a 1-D vibration signal to 2-D time-frequency images by wavelet transform.

The discussion on CNN can be summarised as follows. First, around 51% of the RFD-related works that use DL adopted CNN as the classifier, which indicates how CNN has been a gamechanger, even with time-series data inputs. However, the main concern with using CNN for RFD relates to the 2D representation of time-series data produced by machine health monitoring sensors. Several time-tried, efficacious methods have been involved in this conversion, including SDP, CWTS, and spectrogram. Meanwhile, rather than use time-series data as input, there are a number of methods directly acquire 2D images such as orbital images or IR images. A fair number of the related works (68%) altered the structure of CNN to make it compactable with RFD issues, while the review substantiates the fact that utilizing DFC with CNN is close to impossible. It is also clear that 1D CNNs began to gain in popularity, and adding the discriminative ability to CNN ensure the wider acceptance.

5.2 Deep belief networks

DBN is an unsupervised, probabilistic DL network with a stacked structure of RBMs or AEs, pre-trained using a greedy learning algorithm. The efficacy to provide impressive results with fault specific features rather than with raw data has offered this model much acceptance among the RFD research community.

Certain approaches became significant in the way they retreat the input so that the DBN can make more out of it. In this direction, Oh et al. (2016) developed a method for converting raw vibration signals to the image by the omnidirectional regeneration method (ODR). Histogram of oriented gradients (HOG) descriptor was presented as input to the DBN. They observed that the classification accuracy varied from 95% to 100%, depending on the number of nodes in the hidden layers of the DBN. Guan et al.

(2019) made a significant contribution to structural fault identification by combining EMD and DBN. The IMF signals generated by EMD were selected based on fault information content using sample entropy. The FFT transformation was adopted after the reconstruction of selected signals into new vibration signals and then presented it to DBN. The effectiveness of the proposed method was tested with the FD and TD structural fault features. Comparing with BPNN and CNN, this method achieved the highest diagnostic accuracy and was superior in precision. Compressed sensing has been proved as an efficient method to reduce the data capacity and improve analysis efficiency. This was effectively applied by Shao et al. (2018) along with convolutional DBN (CDBN). To enhance the feature learning ability for the compressed data, the CDBN model was constructed with Gaussian visible units. The exponential moving average (EMA) technique was employed to improve generalization performance.

The handling of heterogeneous data sources is another criterion in RFD. Yan et al. (2019a) proposed a multi-DBN model to solve multi-source heterogeneous information fusion problems. Multiple DBNs approach was adopted in such a way that one DBN was used per each sensor stream, and finally, the fused results were presented as output. The authors demonstrated their method with shaft orbit information and multi-source heterogeneous information, which were collected from a testbed on different rotor misalignment conditions. DBN was observed as a fault feature extractor by Shao et al. (2016). The features from each layer of the multi-layer representations were collected for the performance evaluation of layer-wise feature extraction. The same was utilized for comparison with the traditional features. BPNN was demonstrated as a decent classifier that is capable of classifying six different rotor faults far better than the other methods with the DBN extracted features. A DBN approach for rotor fault classification and mechanical degradation assessment was presented in the work of Li et al. (2019), who applied DBN by stacking five RBM layers with Bernoulli functions. The input vibration data of unbalance, rub, and looseness faulty conditions decomposed by WPD, was applied to the model, and its performance was compared with DNN, k-NN, SVM, and BPNN models.

We can summarise the following aspects of DBN based on the attendant discussion. First, DBN is among the preferred options for RFD researchers as it allows them to remain unrestrained in terms of applying the feature processing and can thus take advantage of the fault-specific features. Even though we were able to identify a few works related to DBN, around 44% of these exclusively focused on rotor faults. Most of them demonstrated the scope and possibilities of using DBN in terms of different aspects, including feature extraction, the fusing of DBNs for multi-sensor data, using DBN with regenerated input, and convolutional DBNs. Meanwhile, layer-wise feature extraction is another interesting aspect of utilizing DBNs. In short, the possibility of using extracted features rather than raw data is the distinguishing feature of this method.

5.3 Autoencoders and DNN models

As an unsupervised method capable of learning features, AE is widely used in many applications, including RFD. The basic AE models are enhanced by stacking multiple AEs to form a stacked autoencoder (SAE), while, through the addition of denoising capability, stacked denoising autoencoders (SDAE) can be developed.

Chen and Li (2017) proposed a simple and straightforward implementation of AE by developing a deep neural network with a stack of multiple DAEs as hidden layer. The model employed dropout regularization to avoid overfitting, and it was used for the identification of SRF with rub and oil film whirl. Even without any fault specific signal preprocessing, their method was able to incorporate the characteristics of input data effectively and extract characteristic information of different faults. Additionally, the feature learning ability of AE has been widely utilized in RFD, and several attempts were made to enhance this feature learning capability. A hybrid AE model combining one denoising autoencoder (DAE) and multiple contractive auto-encoders (CAE), was proposed by Shao et al. (2017) that enhanced the feature learning ability of standard AE. They utilized locality preserving projection (LPP) for the fusion of features to extract the most representative information and a softmax layer for the classification of SRF. They considered different inputs like raw vibration data, features extracted from certain frequency-band signals, features extracted from the most sensitive frequency band signal, etc. The performance of the model was compared with standard DAE, standard CAE, BPNN, and SVM. For good measure, learning fault specific features from TD and FD transformed data with SDAE was also studied in the RFD literature. Zhao et al. (2018) illustrated SDAE to detect the rotor faults with FFT transformed FD signal as input. The performance of SDAE was compared with the different number of hidden nodes and deep architectures, as well as with other models like BPNN and SVM, supplying TD and FD input signals. The SDAE deep architecture adaptively extracted the fault features and was able to achieve an optimum accuracy of 99.89% with softmax output function. Meng et al. (2019) identified several issues associated with deep learning architecture in SRF analysis, especially when using AE. They addressed irregularity of raw vibration, generalization issue, increased number of parameters in the DL model, etc. in their work. They proposed segmentation of raw input, data augmentation, and scaled conjugate gradient (SCG) algorithm in SAE respectively to address these issues. In induction rotating machinery fault diagnosis, the proposed method proved to be superior over the other methods, and it achieved an accuracy of 99.6%.

Furthermore, rotor misalignment through ball removal from bearing on a self-aligning bearing rotor system along with the other bearing faults, was studied by Narendiranath Babu et al. (2018). They presented two autoencoders and a softmax layer for the realization of DNN with an SCG algorithm and the cross-entropy technique. The performance was compared with ANN, and EMD implemented along with HHT achieved 100% accuracy for DNN. The same authors came up with a similar work (Babu et al. 2018) using a journal bearing dataset, and they enhanced the study with TD, FD, and TFD features on the models. The AE with an increased number of hidden layers and a final softmax layer indicated a classification accuracy of 100% again. The results highlighted that as the number of faults increases, the overlap of fault types begins, and that a deep AE performs well in such situations. An SDAE was employed by Kong et al. (2019) for identifying rotor related faults of aero-engine. The vibration input from multi-sensor data was undergone multi-modal heterologous feature extraction by SDAE, and then it was applied to support vector classifiers to fuse SDAEs in different modalities to get the acceptable classification results. Recently, Tang et al. (2020) introduced an SAE to identify contact rubbing, oil whirl, dynamic unbalance, and static unbalance faults. SAE extracted significant fault features after the statistical filtering process, which reduced the interference information and increased the signal-to-noise ratio. The results projected the dominance of the proposed method over DAE, SVM, and BPNN.

Improving the robustness of feature representation of sparse AE is very much important, and one of such works was manifested by Sun et al. (2016). A partial corruption was

added into the input with the help of the denoising coding, and the learned sparse feature representation assisted the training of a neural network with drop out regularization, for classification. During forward propagation, the average activation for sparsity was computed to calculate sparse cost function. The framework was compared with NN, LR, and SVM on rotor fault data of induction motor and was superior over the other methods. Lei et al. (2016a) proposed a two-stage learning process utilizing sparse filtering in SAE. The local discriminative features were extracted from raw vibration signals using sparse filtering. Then by averaging these local features, the learned features were obtained and applied to the softmax regression. Though the method does not deal with SRF data, we included this work since this method can handle the scarcity of labeled data issues, which is very common in RFD. Moreover, the authors give an observation that weight vectors of sparse filtering attribute similar properties to those of the Gabor filters.

Certain methods reduce the overhead of tuning the parameters of the DL models, and one such method was proposed by Yang et al. (2016). An ELM based autoencoder generated a weight matrix of the hidden layer, that transformed the input dataset into a new feature representation, and this representative learning avoided information loss compared to the traditional feature extraction methods. The last single layer of ELM performed the fault classification task. The framework was compared with different types of feature extraction methods of wind turbine data and achieved the best result. Another approach with ELM was proposed by Sharma et al. (2015), in which RMS value of 3-phase voltage and current signature was considered as input to the model. They identified six external faults of three-phase induction motor and compared ELM performance with SVM and MLP to highlight the performance of ELM. From the ELM design perspective, the work of Yang et al. (2016) was differentiated from other ELM designs by the hierarchical structure of ELM. The ELM in this work was designed by concatenating the forwarding list of ELM and were processed independently. The two crucial ELM based components in the framework were ELM based autoencoder and ELM based classifier so that the 'number of hidden nodes' was the only adjustable parameter. Another architecture modification in ELM was observed in ensemble ELM (EELM), proposed by Wang et al. (2020) for compound-fault diagnosis of rotating machinery. It consisted of an ELM based clustering sub-network that generated the Euclidean distance representations from each point to every centroid followed by another subnet that identified potential output tags through multiple-output-node multi-label learning. TD and FD features were extracted by PSO and variational mode decomposition (VMD), and the method proved its dominance on a real-world database. Recently Zhao et al. (2020b) proposed a multi-manifold deep extreme learning machine (MDELM) algorithm to classify UB, MA, LS, and rub faults. The feature extraction was performed by an ELM based modified sparse filtering, and then the mining of the intra-class and inter-class discriminant feature information was accomplished by MDELM with multi-manifold constraints. State-of-the-art algorithms were selected as the benchmark for comparative studies of the MDELM algorithms. In essence, ELM proved to be an efficient method of fault classification for any kind of rotating machinery.

Neural networks with the deep structure of multiple feed-forward layers (neither with RBMs nor with AEs) are also available in the literature of RFD. Bo et al. (2017) developed a simple DNN model that dealt with the whirl machinery rotor vibration data to identify rotor unbalance, misalignment, and the rubbing fault. Eight spectrum values, including the 1x, 2x, 3x harmonics, which have a high effect on SRFs, are presented as the input eigenvalues of the network. Li et al. (2019) applied DNN with five hidden hyperbolic tangent layers with 50, 50, 32, 32, and 32 nodes, respectively. The model was trained on coefficient and energy-based features extracted by DB4 WPD, using a parallelizing stochastic gradient

descent algorithm. The unbalance, rub, and looseness faults were implemented and were compared with DNN, k-NN, SVM, and BPNN to indicate its superiority.

The AE-based RFD research can be summarised as follows. First, autoencoder-based models are the most widely used DL models after CNNs in the RFD-related literature. In fact, since RFD uses the features learned by AE, only nominal attempts have been made to utilize DFC in the literature, and there is an ongoing trend within the research fraternity to enhance the feature-learning capability of AE. Deep AE performs well in terms of multiple-fault or overlapped-fault situations. Generally, SDAE performance has been studied with varying numbers of hidden nodes and different deep architectures. Meanwhile, ELM-based AE is becoming increasingly popular since it shifts the emphasis from the overheads involved in parameter tuning.

5.4 LSTM and GRU

An RNN is capable of learning temporal information of sequential data by holding the past information and sharing weights using its recursive structure. LSTM and GRU are two variants of RNN that are equipped with gates to avoid vanishing gradient problem. There were only a few attempts made to capture the long-term dependencies of input time-series signal using the sequential analysis techniques of AI. One such work was manifested by Lei et al. (2019), who introduced LSTM for categorizing SRFs, bearing faults, and other compound faults in a wind turbine test rig. They dealt with multivariate time-series raw data and used simple data fusing techniques, opening an opportunity for applying advanced data fusion methods. The proposed method was compared with SVM, MLP, RNN, WDCNN, and DCNN with a single sensor as well as multi-sensor data, and projected the dominance of their method. Xiao et al. (2019a) proposed a more advanced approach of hybrid feature learning that combined statistical parameters, recurrence quantification analysis (RQA), and three-layer stacked LSTM. Twenty-nine statistical features, including 16 TD features and 13 FD features and 10 RQA features were extracted and applied to the batch normalization layer. The method proved its preeminence by comparing state-of-the-art AI-based fault diagnosis methods, together with different feature extraction techniques. Similar to LSTM, GRU can also utilize long term dependency of time-series data. Liao et al. (2019) proposed a 1-D CNN-GRU architecture that adaptively learns fault factors. In this three-layer structure, the hidden layer contained a GRU layer with the convolution layer, the pooling layer, and the fully-connected layer. GRU, combined with CNN, helped to exploit the speed and lightness of convnets with the order-sensitivity of GRU. The authors proved the effectiveness of the system by implementing the model in a working hydropower station and illustrated good results in four small hydro operation monitoring systems.

We can conclude the following things about sequential learning DL approaches in RFD. Given the majority of sensing methods used in RFD produce time-series data, only a few attempts were made to pinpoint the sequential analysis for capturing the time-series behavior of RFD data, resulting in a nominal share for RNN in RFD literature. To deal with multivariate time-series, the data fusion technique has shown its effectiveness in certain works. LSTM and GRU models are being recognized in recent times, which provides a roadmap for RFD analysis to explore the long-term dependencies. These models are easily incorporated with the other models to exploit the effectiveness of both.

5.5 Generative adversarial networks

Being an input reconstruction based semi-supervised learning method, GAN turned out to be a critical solution to data-related issues in RFD. In one of such attempts, the data imbalance issue was addressed by Lee et al. (2017) by applying GANs to oversample the minority class. Their method contained two GANs for two different types of faults. They compared GAN based oversampling with standard oversampling methods and their combined effects on the performance of DNN. Based on the results, they concluded that when the imbalance ratio is minor, GAN is adequate, and if the imbalance ratio is significant, then standard methods are desirable. Their results proved that in intelligent fault detection and diagnosis using DNN with imbalanced data, GANs could be effectively utilized. Data imbalance issue was solved using an adversarial network as discriminator and CNN (DACNN) as the classifier, by Han et al. (2019b). In this model the small datasets are transferred to DACNN, and then trained with the adversarial strategy. By experimenting with a wind turbine test rig and a generic gearbox fault dataset, they proved their method effective and superior even when the dataset size is very small.

We can infer the following observations about GANs from the literature. GANs are mainly used for oversampling data in RFD research. While dealing with small-sized datasets, GANs proved to be a solution to many issues, including data imbalance. The philosophy of adversarial training is applicable in other DL models as well.

6 Summary and challenges

A systematic overview of various state-of-the-art AI-based techniques for RFD is presented in this study. Moreover, special attention has been given to introduce rotor faults with their characteristics (refer Table 1), that will be instrumental in deciding the feature selection process, preprocessing methods, and even in selecting the classifier. In this section, a summary of research progress that happened so far in RFD will be described, in line with the various phases of data-driven AI-based RFD framework. The overall statistics of the state-of-the-art research is demonstrated in the heatmap shown in Fig. 5.

Data acquisition phase: Methods in both ML and DL pursue the same process in the data acquisition phase of RFD, though they demand data at different quantity and precision. For an acceptable level of generalization, DL cannot compromise with the amount of data customarily needed by ML, but can deal with noisy data environments. Therefore the data source selection is always critical in the case of RFD. From the literature summarized in the heatmap, it is identified that around 42% of works depended on RTB method for data collection and about 20% bank on the other sources, including test rigs for bearing, gear, or wind turbine whereas a few works utilized open-source datasets as well. The 30% of works collected induction motor current and voltage as input. It is not an exciting fact from the RFD perspective as this input mainly deals with BRB fault only, though very few exceptions are notified. It is quite interesting to notice that ML had to depend hardly 13% on other open sources, while DL heavily draws around 34% for other sources. This indicates that testbed kind of data collection methods are unable to mitigate the data requirement issues of DL, which we have to analyze in conjunction with the number of works explicitly performed for SRF. A significant finding is that 55% of ML related works deal exclusively with RFD, but DL's contribution in this regard is just 36%, which indicates that the remaining works in both ML and DL

		SVM	ANN	Others	Multiple Classifiers	CNN	AEs & DNNs	Others	Overall
		Machine Learning				Deep Learning			
Faults Considered	UB	74.19	86.36	66.67	68.75	75	81.82	63.64	74.45
	MA	38.71	45.45	16.67	31.25	50	54.55	45.45	40.15
	LS	22.58	31.82	22.22	6.25	35.71	36.36	36.36	27.01
	BS	29.03	22.73	22.22	43.75	14.29	9.09	18.18	23.36
	SC	6.45	9.09	5.56	0	0	0	0	3.65
	RIF	12.9	22.73	11.11	6.25	25	18.18	18.18	16.79
	BRB	32.26	9.09	38.89	87.5	10.71	18.18	27.27	29.93
	CF	19.35	27.27	5.56	0	28.57	9.09	18.18	17.52
	OF	35.48	27.27	72.22	56.25	53.57	81.82	63.64	51.09
Signal Processing	TFD	25.81	31.82	16.67	56.25	21.43	9.09	18.18	26.28
	FD	51.61	45.45	38.89	12.5	21.43	36.36	27.27	35.04
	TD	12.9	18.18	27.78	18.75	0	0	0	11.68
	Comb.	9.68	4.55	16.67	12.5	14.29	9.09	45.45	13.87
	Raw Data	0	0	0	0	42.86	45.45	9.09	13.14
Monitoring Technique	Vibration	51.61	81.82	44.44	18.75	67.86	100	90.91	62.04
	V & C	29.03	13.64	22.22	56.25	7.14	0	9.09	20.44
	AE	3.23	4.55	11.11	0	3.57	0	0	3.65
	Temp.	6.45	0	5.56	0	3.57	0	0	2.92
	Comb. /Others	9.68	0	16.67	25	17.86	0	0	10.95
Data Source	RTB	51.61	72.73	27.78	12.5	42.86	36.36	27.27	42.34
	IM	38.71	9.09	44.44	68.75	14.29	9.09	27.27	29.93
	WT	3.23	9.09	0	6.25	10.71	9.09	18.18	7.3
	Other Source	6.45	9.09	27.78	12.5	32.14	45.45	27.27	20.44

Fig. 5 Summary heatmap of RFD (values in percentage)

consider rotor faults along with the bearing or gear faults. This limits the scope of RFD specific research improvements, which in turn causes a proportional falling off in the percentage of AI works that utilized rotor fault specific parameters (DFC, frequency harmonics, etc.) in ML with 14% and DL with 8%. This statistical analysis is shown in Fig. 6.

The facts about the data source of RFD summarized so far draws in two important conclusions. Firstly, testbed, which is recognized as the primary data source of RFD, often fails to provide sufficient data for DL so that DL methods choose some open-source datasets. Secondly, these datasets lack rotor specific data though they have a large amount of bearing or gear fault data.

The other challenges identified from the literature in acquiring faulty data are:

– In real situations, the machine runs in faulty conditions are very rare compared to its normal running conditions. So it is difficult to get faulty data keeping the balanced sampling.
– Though we can simulate faulty conditions in testbeds and acquire data, it will be a challenging and difficult task to run the testbeds for a long time in a faulty environment.
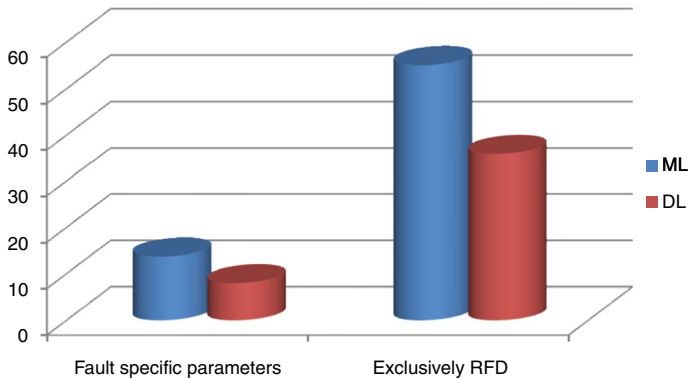
Fig. 6 RFD specific research (percentage of works)

– Commonly available testbeds find it difficult to simulate the frequently changing RPMs, load, and other environmental conditions like noise, which makes the data unrealistic most of the time.

This scenario opens a new research direction of data generation or augmentation. The augmenting dataset is familiar to the image and speech processing research community but not significantly explored in time-series data processing. The techniques proposed for augmentation frequently uses random sampling with a replacement proven to be readily implementable but has certain issues. Firstly, there is no guarantee that the generated data follows the same distribution as that of the original data. Secondly, in no way, it ensures the time-series properties (trend, seasonality, etc.) to be present in the newly synthesized data in most of the cases. In augmentation, it is also challenging to confirm the label of newly created samples, opening a new direction to the research community. Another method of generating data is using GANs as the data generator. As we mentioned earlier, the approaches using GAN is the least attended one in time-series data-based machine health monitoring. It has been proven by many researchers that the popular minority class oversampling methods, like ADASYN (He et al. 2008) and SMOTE (Chawla et al. 2002) are unable to provide good generalization in the scenario where data is sparse, and presence of outliers is more. But GANs often demonstrated much better performance in such scenarios. So proving the effectiveness of GANs in time-series data generation is also a new area of research. The methods to address unlabelled datasets or scarce labeled datasets always demand more research in small datasets situations. It invites more exploration of unsupervised learning paradigms as well. The option of 'rotating machinery libraries' that artificially creates faulty rotor data, mentioned in Sect. 2 is not in vogue, but it leaves something to count on in the present scenario.

Among the signal sensing methods in the data acquisition phase, the most commonly adopted approach is vibration analysis. It is evident that SRF is the root cause that influences the characteristics and behavior of vibration. Hence, except for broken rotor bar faults, more than 80% of works depends on vibration data. Compared to other methods, vibration sensors have to be mounted on the device, and as a result, the signals are sensitive to noise as well as the sensor mount positions. By applying deeper layers, DL overcomes these issues, but ML is not robust to noise, which justifies the comparatively higher percentage of other sensing methods with ML literature. In signal sensing, the position and

number of sensors have a high implication on SRF since different faults affect different planes (refer Table 1), and placing the sensors on that basis will help in better diagnosis results.

Meanwhile, DL prefers to have higher dimensional input data demanding more number and types of sensors, which leads to certain research problems related to the fusion of multi-sensor data. It is clear from the statistics that only around 10% of works both in ML and DL combined tries to deal with multi-sensor data, which indicates that SRF diagnosis has to take a big leap in this direction. Another important observation is that the selection of sensing methods heavily depends on the fault type. For example, broken rotor bar fault is one of the prevalent rotor faults in induction motors, and it is notified from the literature that no sensing methods apart from motor current and voltage, have been tried for this fault both in ML and DL models. Only a small fraction of research (less than 5%) accounts for acoustic emission and temperature-related RFD studies. Still, they are gaining popularity because of their low cost, ease of installation, and their non-contact and non-intrusive nature, which avoids many sensor-related hurdles.

Feature processing Phase: DL exercises the automated feature extraction and selection by learning discriminative features in an end-to-end manner, and therefore, this phase is skipped in DL. But certain signal processing techniques have been utilized in DL models for data preprocessing and proper input representation. As opposed to DL end-to-end learning, ML goes for manual feature extraction, which is summarized from RFD perspective, as follows:

– The symptom parameters of SRF are frequency harmonics. Hence, the TD methods are unable to capture and utilize these parameters, resulting in poor performance for RFD.
– The non-linearity and non-stationary nature of the rotor vibration signals complicates the accurate fault diagnosis for the FD methods. Because the Fourier transform is suitable only for stationary signal processing, and it fails to reveal the inherent information of non-stationary signals. In the case of the wavelet transform, the selection of a wavelet basis and its lack of ability to adapt to changes are the key factors adversely affecting the performance.
– While the TFD processing techniques, particularly STFT, WVD, etc. gained immense popularity in RFD by its exceptional ability to handle non-stationary signals, these are exhibiting certain deficiencies. For instance, STFT fails to produce an ideal resolution to frequency and time simultaneously. Furthermore, to implement efficient and fast STFT algorithm, SFT lacks orthogonal bases. Similarly, WVD produces inference terms in its decomposition process.

The literature review statistics indicate that 18%, 40%, 31% of works still depend on conventional TD, FD, and TFD features, respectively. Irrespective of the fact that TFD processing is best-suited for extracting SRF specific features from the non-linear and non-stationary vibration signals, a more prominent share is enjoyed by FD operations.

In such a scenario, it is evident that the fault specific DFC extraction is the key to the feature extraction of RFD. But using the aforementioned methods from a general perspective is ineffective in fault specific feature extraction as the fault categories are complex and large. The 14% of ML literature in RFD accounts for the fault specific component extraction (refer Fig. 6), which either depends on FD operations to extract the harmonics of rotation frequency or the frequency decomposition (like EMD and its variants, WPD, etc.) in the TFD. These facts signify that further research can be conducted in fault specific symptom parameter extraction by wisely utilizing conventional

Table 6 Fault specific component usage

Category	Model	Significant contributions
ML (14%)	ANN	Srinivas et al. (2010a), El-Shafei et al. (2007), Chen and Chen (2011), Bin et al. (2012) , Srinivas and Holla (2014), Cacciola et al. (2016)
	SVM	Fengqi and Meng (2006), Tang et al. (2010), Baccarini et al. (2011), Xue et al. (2011), Pang et al. (2018)
DL (8%)	CNN	Ince et al. (2016), Yao et al. (2018)
	DBN	Li et al. (2019)
	DNN	Bo et al. (2017)

FD and TFD processing methods. Table 6 lists the contribution in utilizing the rotor fault specific symptom parameters, and that proves their significance in RFD.

From the literature, it is evident that DL methods for RFD also demand certain signal processing with 26% and 18% in the FD, and TFD, respectively. Certain observations from the literature are summarized as follows:

– Raw vibration signal is 1-D time-series necessitates conversion into 2-D form for CNN to operate convenient and fast, though 1-D CNNs are also present in the literature.
– As input, deep neural networks prefer frequency decomposed components to raw vibration signals.
– The performance of DBNs with raw data or TD features are worse compared to the other domain features.
– The signal acquisition mechanism also is substantive; for example, infrared images or sound inputs need proper preprocessing before applying to the DL models.

In respect of fault specific DFC utilization in DL, DBNs, AE-based models, and 1-D CNNs have opened up some provisions in the literature, but constituted only 8% of the overall works (refer Fig. 6), and requires more attention in research. It is observed that, once the adequate quality features are extracted, the feature selection process followed in the RFD is the same as that followed in any other rotating machinery fault diagnosis process.

Classification Phase: Around 80% of ML models operated in the RFD literature is constituted by SVM (41%), ANN (29%), and k-NN. The NB and all the other classifiers provide less than 5% each in the literature. Similarly, CNN seized a predominant position among the DL methods with a 56% share in the literature followed by AE based models, while DBN also registered some significant contributions. The general features of these classifiers, their pros and cons, along with their compatibility with RFD, are already summarized in Table 2 and 4.

In terms of accuracy and classification speed, SVM and ANN are top-notch, which is evident from their wide acceptance among the ML models in RFD. But the analysis of the works of comparison study of ML classifiers had shown that even when ANN meets the expectation, the accuracy of SVM was not commendable (refer Table 3). The reason may be the architecture versatility of ANN compared to SVM, which is evident from the

fact that a higher fraction of ANNs in the literature altered their basic structure in accordance with the RFD problem. NB and k-NN are worthy of their ability to deal with overfitting and have good interpretability while SVM and ANN still have to advance a lot in this regard. Since the raw signals acquired from sensors are affected by noises, it is observed that the researchers depended on preprocessing steps to deal with noisy raw data, rather than going for NB like ML models which project staunch robustness to noise.

Regarding RFD, one of the primary focuses of our discussion is about incorporating domain-specific knowledge in decision making, which we have tried to explain through the fault specific symptom parameter extraction capability of the model. It is observed that ML is doing it a much better way (though just 14% of works tried it) by using frequency decomposition and utilizing harmonics of rotation frequency as features. But this approach possesses certain challenges:

– *Ideal single fault assumption*: Most of the works assume that only one fault occurs in the rotor at a time, which is quite impractical in real industrial scenarios. The monitored characteristic frequency will be the result of frequency interplay between different faults.
– *Unobservable faults*: The manifestation of rotor faults like corrosion and wear, and some oil-based faults with these frequency components are not much sensible.
– *Par with industrial data*: The real data with continuously changing rotational frequency and load conditions cannot be simulated in most of the testbeds, which limits the applicability of this strategy.
– *Multiple movements*: The trajectory of rotor motion need not always be maintained the same, which affects the vibrational frequency components. Environmental vibration is also an influencing factor.

The application of such domain knowledge into DL models by means of fault specific features is not always straightforward as in ML, because of the automated feature extraction and selection of end-to-end learning strategy. CNN is one example for this phenomenon, from the fact that just 7.14% of works tried to incorporate fault specific features by some data fusion and 1-D CNN model. But as we mentioned earlier, DBN, AE-based models, and DNNs give some scope in this direction since they prefer processed data to raw data. One of the remarkable features in the existing literature is that the fault specific discriminative feature extraction reduces the size of the DL model and eventually leads to better performance. Some other works indicated significant effort to change the 'black box' nature of DL by using proper visualization tools like t-SNE. Almost all DL algorithms are robust to noise, but CNNs present a bit more denoising capability compared to others, especially better than AEs. SDAE like models overcome these limitations. The performance generalization issues of DL related to the size and diversity of data, class imbalance, etc. have already been discussed at the beginning of the summary section. A few more observations we like to add here.

– Regarding the accuracy of DL model, it is around 99%, and some are achieving 100% from testbed datasets, which need not guarantee the same with industry data of varying RPM and load conditions. The unavailability of generalized datasets limits the opportunity to compare different works in terms of accuracy.
– Unbalanced datasets demand the usage of metrics such as precision, recall, F1-score, etc. along with the most commonly utilized accuracy metric.

– While the acquired signal is sequential in nature, very few works have tried to address RFD problems with sequential deep learning models like RNN.

7 Future research directions

In order to encourage research on RFD, with special emphasis on SRFs to provide a more generalized, industry conforming, and realistic solution, the current research can be modified in the following ways.

1. Synthetic data generation: Utilize synthesized data by engaging data augmentation or GAN based data generation for addressing data scarcity issues. Research can be done in time-series data augmentation with a focus on ensuring time-series properties of the data, maintaining the correlation between different columns of multivariate time-series, and ultimately confirming the labels of synthesized data. This poses a challenging research problem in time-series domain and augmentation literature. GAN based data generation has been tried in certain bearing fault diagnosis works, leaving scope for research in RFD data generation.
2. Challenging datasets: Another research lies in filling the gap between testbed data and real-world industrial data. As we have mentioned earlier, creating a complicated dataset by applying varying RPMs and load conditions in testbed simulations can create new research issues and makes the solution industry-ready.
3. Challenging Problems: Most of the RFD research deals with single faults, and only very few have tried to study compound faults, which is not on par with the real-world scenario. The research challenge in dealing with compound faults can be made more complicated by considering the compound faults as a multi labeling problem.
4. SRF specific symptom parameters: Extraction and usage of fault specific symptom parameters poses new research challenges in both 'feature engineering' as well as refining classifier architecture.
5. More learning strategies: Classifier fusion and transfer learning kind of learning strategies are least applied in RFD. Fusing new classifiers at different levels (data-level, feature-level, decision level, etc.) opens a lot of opportunities to improve accuracy. Similarly, the training with artificial faults and the transfer of the trained model for the classification of natural faults in a real-world scenario (by facilitating classifier fusion, transfer learning, and domain adaption) is a challenging problem in research. Few-shot learning is another option which is already proved to make reasonable accuracy with fewer data.
6. Sequential deep learning: The literature of RFD indicates that no significant attempts have been made yet to exploit the sequential nature of the sensed signals utilizing RNN based deep sequential models like LSTM and GRU.

8 Conclusion

This paper focused on conducting a comprehensive review of the literature related to RFD involving ML and DL strategies. We first introduced SRF along with various other rotor faults, emphasizing the attendant relevance and characteristics in order to prioritize a

fault-wise analysis rather than the currently followed practice of component-wise analysis. RFD was further discussed in relation to an AI-based framework to present the methodologies adopted in each phase, demonstrating the pros, cons, and challenges to provide an overview of the current RFD scenario. The review on the ML and DL classifier models was presented with the aim of emphasizing the rotor fault specific factors, such as the ability to incorporate fault-specific symptom parameters in feature engineering and dealing with imbalanced, scarce, or noisy data. The works that conducted a comparison of the ML models within an SRF context were summarised separately, while similar attention was also paid to the various classifier fusion methods. The summary section concluded the current trends, challenges, and recommendations in terms of all the aforementioned aspects, whereas the future direction section, mainly focused on facilitating the application of testbed solutions in real-world industrial scenarios. Ultimately, this work is presented to provide a single point of reference for RFD researchers in relation to AI with practical insights.

References

Al-Bedoor B (2001) Modeling the coupled torsional and lateral vibrations of unbalanced rotors. Comput Methods Appl Mech Eng 190(45):5999–6008

Alsalaet J (2012) Vibration analysis and diagnostic guide. http://www.jaafarali.info/files/vibration_analysis_guide.pdf. Accessed: 2019-09-30

Angelopoulos A, Michailidis ET, Nomikos N, Trakadas P, Hatziefremidis A, Voliotis S, Zahariadis T (2020) Tackling faults in the industry4.0 era—a survey of machine-learning solutions and key aspects. Sensors 20(1):109

Armaki MG, Roshanfekr R (2010) A new approach for fault detection of broken rotor bars in induction motor based on support vector machine. In: Proceedings—2010 18th Iranian Conference on Electrical Engineering, ICEE 2010 pp 732–738, https://doi.org/10.1109/IRANIANCEE.2010.5506976

Aydin I, Karaköse M, Akin E (2007) Artificial immune based support vector machine algorithm for fault diagnosis of induction motors. In: International Aegean Conference on Electrical Machines and Power Electronics and Electromotion ACEMP'07 and Electromotion'07 Joint Conference pp 217–221, https://doi.org/10.1109/ACEMP.2007.4510505

Babu NT, Aravind A, Rakesh A, Jahzan M, Prabha RD, Ramalinga Viswanathan M (2018) Automatic fault classification for journal bearings using ANN and DNN. Arch Acoust 43(4):727–738. https://doi.org/10.24425/aoa.2018.125166

Baccarini LMR, Rocha E, Silva VV, De Menezes BR, Caminhas WM (2011) SVM practical industrial application for mechanical faults diagnostic. Expert Syst Appl 38(6):6980–6984. https://doi.org/10.1016/j.eswa.2010.12.017

Bachschmid N, Pennacchi P, Vania A (2002) Identification of multiple faults in rotor systems. J Sound Vib 254(2):327–366

Bate GH (1987) Vibration diagnostics for industrial electric motor drives. In: Application Notes, Bruel & Kjaer, Nacrum

Biet M (2013) Rotor faults diagnosis using feature selection and nearest neighbors rule: application to a turbogenerator. IEEE Trans Ind Electron 60(9):4063–4073. https://doi.org/10.1109/TIE.2012.2218559

Biet M, Bijeire A (2011) Rotor faults diagnosis in synchronous generators using feature selection and nearest neighbors rule. In: SDEMPED 2011—8th IEEE Symposium on Diagnostics for Electrical Machines, Power Electronics and Drives pp 300–306, https://doi.org/10.1109/DEMPED.2011.6063640

Bin GF, Gao JJ, Li XJ, Dhillon BS (2012) Early fault diagnosis of rotating machinery based on wavelet packets—empirical mode decomposition feature extraction and neural network. Mech Syst Signal Process 27(1):696–711. https://doi.org/10.1016/j.ymssp.2011.08.002

Bo L, Jin Y, Chen Y, Fan X (2017) The study on rotating machinery fault diagnosis based on deep neural networks. In: Proceedings—2016 International Conference on Intelligent Transportation, Big Data and Smart City, ICITBS 2016 9:125–129, https://doi.org/10.1109/ICITBS.2016.67

Bognatz S (1995) Alignment of critical and non critical machines. Orbit 4:23–25

Bonnett AH (2000) Root cause ac motor failure analysis with a focus on shaft failures. IEEE Trans Ind Appl 36(5):1435–1448

Cabal-Yepez E, Fernandez-Jaramillo AA, Garcia-Perez A, Romero-Troncoso RJ, Lozano-Garcia JM (2015) Real-time condition monitoring on vsd-fed induction motors through statistical analysis and synchronous speed observation. Int Trans Electr Energy Syst 25(8):1657–1672

Cacciola S, Agud IM, Bottasso CL (2016) Detection of rotor imbalance, including root cause, severity and location. J Phys Conf Ser. https://doi.org/10.1088/1742-6596/753/7/072003

Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP (2002) Smote: synthetic minority over-sampling technique. J Artif Intell Res 16:321–357

Chen CS, Chen JS (2011) Rotor fault diagnosis system based on sGA-based individual neural networks. Expert Syst Appl 38(9):10822–10830. https://doi.org/10.1016/j.eswa.2011.02.074

Chen P (2009) Foundation and application of condition diagnosis technology for rotating machinery. San* eisha publication, Japan

Chen WY, Xu JX, Panda SK (2011) A study on automatic machine condition monitoring and fault diagnosis for bearing and unbalanced rotor faults. In: Proceedings—ISIE 2011: 2011 IEEE International Symposium on Industrial Electronics pp 2105–2110, https://doi.org/10.1109/ISIE.2011.5984486

Chen X, Wang S, Qiao B, Chen Q (2018) Basic research on machinery fault diagnostics: past, present, and future trends. Front Mech Eng 13(2):264–291

Chen Z, Li Z (2017) Research on fault diagnosis method of rotating machinery based on deep learning. 2017 Prognostics and System Health Management Conference, PHM-Harbin 2017—Proceedings. https://doi.org/10.1109/PHM.2017.8079279

Chow MY, Mangum PM, Yee SO (1991) A neural network approach to real-time condition monitoring of induction motors. IEEE Trans Ind Electron 38(6):448–453

Chu F, Lu W (2005) Experimental observation of nonlinear vibrations in a rub-impact rotor system. J Sound Vib 283(3–5):621–643

Cui L, Huang J, Zhang F (2017) Quantitative and localization diagnosis of a defective ball bearing based on vertical-horizontal synchronization signal analysis. IEEE Trans Ind Electron 64(11):8695–8706

Cunha Palácios RH, Da Silva IN, Goedtel A, Godoy WF (2015) A comprehensive evaluation of intelligent classifiers for fault identification in three-phase induction motors. Electr Power Syst Res 127:249–258. https://doi.org/10.1016/j.epsr.2015.06.008

Darpe A, Gupta K, Chawla A (2006) Dynamics of a bowed rotor with a transverse surface crack. J Sound Vib 296(4–5):888–907

Devendiran S, Manivannan K (2016) Vibration based condition monitoring and fault diagnosis technologies for bearing and gear components-a review. Int J Appl Eng Res 11(6):3966–3975

Duan L, Xie M, Bai T, Wang J (2016a) A new support vector data description method for machinery fault diagnosis with unbalanced datasets. Expert Syst Appl 64:239–246. https://doi.org/10.1016/j.eswa.2016.07.039

Duan L, Yao M, Wang J, Bai T, Zhang L (2016b) Segmented infrared image analysis for rotating machinery fault diagnosis. Infrared Phys Technol 77:267–276. https://doi.org/10.1016/j.infrared.2016.06.011

El-Shafei A, Hassan TA, Soliman AK, Zeyada Y, Rieger N (2007) Neural network and fuzzy logic diagnostics of 1x faults in rotating machinery. J Eng Gas Turbines Power 129(3):703–710. https://doi.org/10.1115/1.2227417

Fahy F, Thompson D (2016) Fundamentals of sound and vibration. CRC Press, Boca Raton. https://doi.org/10.1201/b18348

Fang R, Ma H (2006) Application of mcsa and svm to induction machine rotor fault diagnosis. In: 2006 6th World Congress on Intelligent Control and Automation, IEEE, vol 2, pp 5543–5547

Fatima S, Guduri B, Mohanty AR, Naikan VNA (2014) Transducer invariant multi-class fault classification in a rotor-bearing system using support vector machines. Meas J Int Meas Conf 58:363–374

Fei CW, Bai GC, Tang WZ, Ma S (2014) Quantitative diagnosis of rotor vibration fault using process power spectrum entropy and support vector machine method. Shock Vib 20:4. https://doi.org/10.1155/2014/957531

Fengqi W, Meng G (2006) Compound rub malfunctions feature extraction based on full-spectrum cascade analysis and SVM. Mech Syst Signal Process 20(8):2007–2021. https://doi.org/10.1016/j.ymssp.2005.10.004

Gang N, Son JD, Widodo A, Yang BS, Hwang DH, Kang DS (2007) A comparison of classifier performance for fault diagnosis of induction motor using multi-type signals. Struct Health Monitor 6(3):215–229. https://doi.org/10.1177/1475921707081110

Gangsar P, Tiwari R (2019) A support vector machine based fault diagnostics of induction motors for practical situation of multi-sensor limited data case. Measurement 135:694–711

Glowacz A (2018) Acoustic based fault diagnosis of three-phase induction motor. Appl Acoust 137(February):82–89. https://doi.org/10.1016/j.apacoust.2018.03.010

Godoy WF, da Silva IN, Goedtel A, Palácios RHC, Lopes TD (2016) Application of intelligent tools to detect and classify broken rotor bars in three-phase induction motors fed by an inverter. IET Electr Power Appl 10(5):430–439. https://doi.org/10.1049/iet-epa.2015.0469

Gohari M, Eydi AM (2020) Modelling of shaft unbalance: modelling a multi discs rotor using k-nearest neighbor and decision tree algorithms. Measurement 151:107253

Guan Z, Liao Z, Li K, Chen P (2019) A precise diagnosis method of structural faults of rotating machinery based on combination of empirical mode decomposition, sample entropy, and deep belief network. Sens Switzerland. https://doi.org/10.3390/s19030591

Guo S, Yang T, Gao W, Zhang C (2018) A novel fault diagnosis method for rotating machinery based on a convolutional neural network. Sens Switzerland 18(5):1. https://doi.org/10.3390/s18051429

Guyon I, Elisseeff A (2003) An introduction to variable and feature selection. J Mach Learn Res 3(Mar):1157–1182

Han T, Liu C, Wu L, Sarkar S, Jiang D (2019a) An adaptive spatiotemporal feature learning approach for fault diagnosis in complex systems. Mech Syst Signal Process 117:170–187. https://doi.org/10.1016/j.ymssp.2018.07.048

Han T, Liu C, Yang W, Jiang D (2019b) A novel adversarial learning framework in deep convolutional neural network for intelligent diagnosis of mechanical faults. Knowl-Based Syst. https://doi.org/10.1016/j.knosys.2018.12.019

Hariharan V, Srinivasan P (2009) Vibration analysis of misaligned shaft vibration analysis of misaligned shaft-ball bearing system ball bearing system ball bearing system. Indian J Sci Technol 2(9)

Hassan T, El-Shafei A, Zeyada Y, Rieger N (2003) Comparison of neural network architectures for machinery fault diagnosis. In: ASME Turbo Expo 2003, collocated with the 2003 International Joint Power Generation Conference, American Society of Mechanical Engineers Digital Collection, pp 415–424

He H, Bai Y, Garcia EA, Li S (2008) Adasyn: Adaptive synthetic sampling approach for imbalanced learning. In: 2008 IEEE International Joint Conference on Neural Networks (IEEE World Congress on Computational Intelligence), IEEE, pp 1322–1328

Ince T, Kiranyaz S, Eren L, Askar M, Gabbouj M (2016) Real-time motor fault detection by 1-D convolutional neural networks. IEEE Trans Ind Electron 63(11):7067–7075. https://doi.org/10.1109/TIE.2016.2582729

Ishibashi T, Han B, Kawai T (2017) Rotating machinery library for diagnosis. In: Proceedings of the 12th International Modelica Conference, Prague, Czech Republic, May 15–17, 2017, Linköping University Electronic Press, 132, pp 381–387

Janssens O, Schulz R, Slavkovikj V, Stockman K, Loccufier M, Van De Walle R, Van Hoecke S (2015) Thermal image based fault diagnosis for rotating machinery. Infrared Phys Technol 73:78–87. https://doi.org/10.1016/j.infrared.2015.09.004

Janssens O, Slavkovikj V, Vervisch B, Stockman K, Loccufier M, Verstockt S, Van de Walle R, Van Hoecke S (2016) Convolutional neural network based fault detection for rotating machinery. J Sound Vib 377:331–345. https://doi.org/10.1016/j.jsv.2016.05.027

Jeong H, Park S, Woo S, Lee S (2016) Rotating machinery diagnostics using deep learning on orbit plot images. Proc Manuf 5:1107–1118. https://doi.org/10.1016/j.promfg.2016.08.083

Jiang Q, Shen Y, Li H, Xu F (2018) New fault recognition method for rotary machinery based on information entropy and a probabilistic neural network. Sens Switzerland. https://doi.org/10.3390/s18020337

Kang M, Kim JM (2013) Singular value decomposition based feature extraction approaches for classifying faults of induction motors. Mech Syst Signal Process 41(1–2):348–356. https://doi.org/10.1016/j.ymssp.2013.08.002

Karvelis P, Georgoulas G, Tsoumas IP, Antonino-Daviu JA, Climente-Alarcón V, Stylios CD (2015) A symbolic representation approach for the diagnosis of broken rotor bars in induction motors. IEEE Trans Ind Informatics 11(5):1028–1037. https://doi.org/10.1109/TII.2015.2463680

Keskes H, Braham A, Lachiri Z (2013) Broken rotor bar diagnosis in induction machines through stationary wavelet packet transform and multiclass wavelet SVM. Electr Power Syst Res 97:151–157. https://doi.org/10.1016/j.epsr.2012.12.013

Konar P, Sil J, Chattopadhyay P (2015) Knowledge extraction using data mining for multi-class fault diagnosis of induction motor. Neurocomputing 166:14–25. https://doi.org/10.1016/j.neuco m.2015.04.040

Kong X, Peng G, Li X, Wang Z, Na X (2019) Aeroengine Fault Diagnosis Method Based on Stack Denoising Auto-Encoders Network. In: Proceedings—2018 11th International Congress on Image and Signal Processing. BioMedical Engineering and Informatics, CISP-BMEI 2018:1–9. https:// doi.org/10.1109/CISP-BMEI.2018.8633128

Kurek J, Osowski S (2010) Support vector machine for fault diagnosis of the broken rotor bars of squirrel-cage induction motor. Neural Comput Appl 19(4):557–564. https://doi.org/10.1007/s0052 1-009-0316-5

Langarica S, Ruffelmacher C, Nunez F (2019a) An Industrial Internet Application for Real-Time Fault Diagnosis in Industrial Motors. IEEE Transactions on Automation Science and Engineering pp 1–12, https://doi.org/10.1109/TASE.2019.2913628

Langarica S, Rüffelmacher C, Núñez F (2019b) An industrial internet application for real-time fault diagnosis in industrial motors. IEEE Transactions on Automation Science and Engineering

Lee GY, Kim M, Quan YJ, Kim MS, Kim TJY, Yoon HS, Min S, Kim DH, Mun JW, Oh JW et al (2018) Machine health management in smart factory: a review. J Mech Sci Technol 32(3):987–1009

Lee YO, Jo J, Hwang J (2017) Application of deep neural network and generative adversarial network to industrial maintenance: A case study of induction motor fault detection. In: Proceedings—2017 IEEE International Conference on Big Data, Big Data 2017 2018-January:3248–3253, https://doi. org/10.1109/BigData.2017.8258307

Lees A, Sinha J, Friswell M (2009) Model-based identification of rotating machines. Mech Syst Signal Process 23(6):1884–1893

Lei J, Liu C, Jiang D (2019) Fault diagnosis of wind turbine based on long short-term memory networks. Renew Energy 133:422–432. https://doi.org/10.1016/j.renene.2018.10.031

Lei Y, He Z, Zi Y, Hu Q (2007) Fault diagnosis of rotating machinery based on multiple ANFIS combination with GAs. Mech Syst Signal Process 21(5):2280–2294. https://doi.org/10.1016/j.ymssp .2006.11.003

Lei Y, He Z, Zi Y (2008) A new approach to intelligent fault diagnosis of rotating machinery. Expert Syst Appl 35(4):1593–1600. https://doi.org/10.1016/j.eswa.2007.08.072

Lei Y, Jia F, Lin J, Xing S, Ding SX (2016a) An intelligent fault diagnosis method using unsupervised feature learning towards mechanical big data. IEEE Trans Ind Electron 63(5):3137–3147. https:// doi.org/10.1109/TIE.2016.2519325

Lei Y, Jia F, Lin J, Xing S, Ding SX (2016b) An intelligent fault diagnosis method using unsupervised feature learning towards mechanical big data. IEEE Trans Ind Electron 63(5):3137–3147. https:// doi.org/10.1109/TIE.2016.2519325

Li W, Tsai Y, Chiu C (2004a) The experimental study of the expert system for diagnosing unbalances by ann and acoustic signals. J Sound Vib 272(1–2):69–83

Li W, Tsai YP, Chiu CL (2004b) The experimental study of the expert system for diagnosing unbalances by ANN and acoustic signals. J Sound Vib 272(1–2):69–83. https://doi.org/10.1016/S0022 -460X(03)00317-1

Li W, Zhu Z, Jiang F, Zhou G, Chen G (2015a) Fault diagnosis of rotating machinery with a novel statistical feature extraction and evaluation method. Mech Syst Signal Process 50:414–426

Li W, Zhu Z, Jiang F, Zhou G, Chen G (2015b) Fault diagnosis of rotating machinery with a novel statistical feature extraction and evaluation method. Mech Syst Signal Process 50–51:414–426. https:// doi.org/10.1016/j.ymssp.2014.05.034

Li X, Wang K, Jiang L (2011) The application of AE signal in early cracked rotor fault diagnosis with PWVD and SVM. J Softw 6(10):1969–1976. https://doi.org/10.4304/jsw.6.10.1969-1976

Li Z, Pedersen VGB, Wang K, He Y (2018) Fault classification and degradation assessment based on wavelet packet decomposition for rotary machinery. Lecture Notes Electr Eng 451:509–516. https ://doi.org/10.1007/978-981-10-5768-7_54

Li Z, Wang Y, Wang K (2019) A deep learning driven method for fault classification and degradation assessment in mechanical equipment. Comput Ind 104:1–10. https://doi.org/10.1016/j.compi nd.2018.07.002

Liang P, Deng C, Wu J, Yang Z, Zhu J, Zhang Z (2019) Compound fault diagnosis of gearboxes via multi-label convolutional neural network and wavelet transform. Comput Ind 113:103132. https://doi. org/10.1016/j.compind.2019.103132

Liao GP, Gao W, Yang GJ, Guo MF (2019) Hydroelectric generating unit fault diagnosis using 1-D convolutional neural network and gated recurrent unit in small hydro. IEEE Sens J 19(20):9352–9363. https:// doi.org/10.1109/JSEN.2019.2926095

Liu D, Lai X, Xiao Z, Hu X, Zhang P (2020) Fault diagnosis of rotating machinery based on convolutional neural network and singular value decomposition. Shock and Vibration 2020

Liu R, Meng G, Yang B, Sun C, Chen X (2017a) Dislocated time series convolutional neural architecture: an intelligent fault diagnosis approach for electric machine. IEEE Trans Ind Inform 13(3):1310–1320. https://doi.org/10.1109/TII.2016.2645238

Liu R, Yang B, Zio E, Chen X (2018) Artificial intelligence for fault diagnosis of rotating machinery: a review. Mech Syst Signal Process 108:33–47

Liu Z, Wang J, Duan L, Shi T, Fu Q (2017b) Infrared image combined with cnn based fault diagnosis for rotating machinery. In: Proceedings—2017 International Conference on Sensing, Diagnostics, Prognostics, and Control, SDPC 2017 2017-December:137–142, https://doi.org/10.1109/SDPC.2017.35

Lobato TH, da Silva RR, da Costa ES, Mesquita AL (2019) An integrated approach to rotating machinery fault diagnosis using, eemd, svm, and augmented data. J Vib Eng Technol pp 1–6

Lu J, Ming T, Zhang C (2019a) Simulation research of rotor misalignment fault based on adams. In: IOP Conference Series: Earth and Environmental Science, IOP Publishing, vol 252, p 022148

Lu S, Qian G, He Q, Liu F, Liu Y, Wang Q (2019b) Insitu motor fault diagnosis using enhanced convolutional neural network in an embedded system. IEEE Sens J PP(c):1–1. https://doi.org/10.1109/jsen.2019.2911299

Ma H, Zhao X, Teng Y, Wen B (2011) Analysis of dynamic characteristics for a rotor system with pedestal looseness. Shock Vib 18(1–2):13–27

Cruz MA, Marques Cardoso SAJ (2000) Rotor cage fault diagnosis in three-phase induction motors by extended park's vector approach. Electr Mach Power Syst 28(4):289–299

MacCamhaoil M (2016) Static and dynamic balancing of rigid rotors. Bruel & Kjaer application notes, BO pp 0276–12

Mais J (2002) Spectrum analysis the key features of analyzing spectra. https://www.skf.com/binary/tcm:12-113997/CM5118%20EN%20Spectrum%20Analysis.pdf

Malik H, Mishra S (2016) Proximal support vector machine (PSVM) based imbalance fault diagnosis of wind turbine using generator current signals. Energy Proc 90(December 2014):593–603. https://doi.org/10.1016/j.egypro.2016.11.228

Malik H, Mishra S (2017) Artificial neural network and empirical mode decomposition based imbalance fault diagnosis of wind turbine using TurbSim, FAST and Simulink. IET Renew Power Gener 11(6):889–902. https://doi.org/10.1049/iet-rpg.2015.0382

Mangasarian OL, Wild EW (2001) Proximal support vector machine classifiers. In: Proceedings KDD-2001: Knowledge discovery and data mining, Citeseer

Martin-Diaz I, Morinigo-Sotelo D, Duque-Perez O, De Romero-Troncoso RJ (2017a) Early fault detection in induction motors using adaboost with imbalanced small data and optimized sampling. IEEE Trans Ind Appl 53(3):3066–3075. https://doi.org/10.1109/TIA.2016.2618756

Martin-Diaz I, Romero-Troncoso RJ, Morinigo-Sotelo D, Duque-Perez O (2017b) Evaluation of intelligent approaches for motor diagnosis under changing operational conditions. 2017 IEEE 11th International Symposium on Diagnostics for Electrical Machines. Power Electronics and Drives (SDEMPED), IEEE, pp 65–71

Martin-Diaz I, Morinigo-Sotelo D, Duque-Perez O, Romero-Troncoso RJ (2018) An experimental comparative evaluation of machine learning techniques for motor fault diagnosis under various operating conditions. IEEE Trans Ind Appl 54(3):2215–2224. https://doi.org/10.1109/TIA.2018.2801863

Martínez-Morales JD, Palacios-Hernández ER, Campos-Delgado DU (2018) Multiple-fault diagnosis in induction motors through support vector machine classification at variable operating conditions. Electr Eng 100(1):59–73. https://doi.org/10.1007/s00202-016-0487-x

Meng Z, Guo X, Pan Z, Sun D, Liu S (2019) Data segmentation and augmentation methods based on raw data using deep neural networks approach for rotating machinery fault diagnosis. IEEE Access 7:79510–79522. https://doi.org/10.1109/ACCESS.2019.2923417

Moosavian A, Ahmadi H, Sakhaei B, Labbafi R (2014) Support vector machine and K-nearest neighbour for unbalanced fault detection. J Qual Maint Eng 20(1):65–75. https://doi.org/10.1108/JQME-04-2012-0016

Mosterman PJ, Zander J (2016) Industry 4.0 as a cyber-physical system study. Softw Syst Model 15(1):17–29

Nakhaeinejad M, Ganeriwala S (2009) Observations on dynamic responses of misalignments. Technologial Notes SpectraQuest Inc

Narendiranath Babu T, Aravind A, Rakesh A, Jahzan M, Rama Prabha D (2018) Application of EMD ANN and DNN for self-aligning bearing fault diagnosis. Arch Acoust 43(2):163–175. https://doi.org/10.24425/122364

Nguyen NT, Lee HH (2008) Improvement of induction motor fault diagnosis performance by using genetic algorithm-based feature selection. Proc Inst Mech Eng Part C J Mech Eng Sci 222(8):1613–1619. https://doi.org/10.1243/09544062JMES913

Nguyen NT, Lee HH (2010) Induction motor fault diagnosis based on the k-nn and optimal feature selection. Int J Electron 97(9):1071–1081

Nguyen NT, Lee HH, Kwon JM (2008) Optimal feature selection using genetic algorithm for mechanical fault detection of induction motor. J Mech Sci Technol 22(3):490–496. https://doi.org/10.1007/s12206-007-1036-3

Niu G, Han T, Yang BS, Tan ACC (2007) Multi-agent decision fusion for motor fault diagnosis. Mech Syst Signal Process 21(3):1285–1299. https://doi.org/10.1016/j.ymssp.2006.03.003

Niu G, Widodo A, Son JD, Yang BS, Hwang DH, Kang DS (2008) Decision-level fusion based on wavelet decomposition for induction motor fault diagnosis using transient current signal. Expert Syst Appl 35(3):918–928. https://doi.org/10.1016/j.eswa.2007.08.024

Oh H, Jeon BC, Jung JH, Youn BD (2016) Smart diagnosis of journal bearing rotor systems: Unsupervised feature extraction scheme by deep learning. In: Proceedings of the Annual Conference of the Prognostics and Health Management Society, PHM 2016-October: 109–116

Oztemel E, Gursev S (2020) Literature review of industry 4.0 and related technologies. J Intell Manuf 31(1):127–182

Pang B, Tang G, Zhou C, Tian T (2018) Rotor fault diagnosis based on characteristic frequency band energy entropy and support vector machine. Entropy. https://doi.org/10.3390/e20120932

Pang X, Shao J, Xue X, Jiang W (2020) Application of axis orbit image optimization in fault diagnosis for rotor system. International Journal of Rotating Machinery 2020

Patel T, Darpe A (2009a) Use of full spectrum cascade for rotor rub identification. Adv Vib Eng 8(2):139–151

Patel TH, Darpe AK (2009b) Coupled bending-torsional vibration analysis of rotor with rub and crack. J Sound Vib 326(3–5):740–752

Patel TH, Darpe AK (2009c) Vibration response of misaligned rotors. J Sound Vib 325(3):609–628

Pennacchi P, Bachschmid N, Vania A, Zanetta GA, Gregori L (2006) Use of modal representation for the supporting structure in model-based fault identification of large rotating machinery: part 1-theoretical remarks. Mech Syst Signal Process 20(3):662–681

Qin J, Liu Y, Grosvenor R (2016) A categorical framework of manufacturing for industry 4.0 and beyond. Procedia Cirp 52:173–178

Qiu Y, Rao SS (2005) A fuzzy approach for the analysis of unbalanced nonlinear rotor systems. J Sound Vib 284(1–2):299–323. https://doi.org/10.1016/j.jsv.2004.06.029

Quiroz JC, Mariun N, Mehrjou MR, Izadi M, Misron N, Mohd Radzi MA (2018) Fault detection of broken rotor bar in LS-PMSM using random forests. Meas J Int Meas Conf 116(November 2017):273–280. https://doi.org/10.1016/j.measurement.2017.11.004

Rahman AGA, Ismail Z, Noroozi S, Chao OZ (2013) Study of open crack in rotor shaft using changes in frequency response function phase. Int J Damage Mech 22(6):791–807

Rai V, Mohanty A (2007) Bearing fault diagnosis using fft of intrinsic mode functions in hilbert-huang transform. Mech Syst Signal Process 21(6):2607–2615

Reddy MCS, Sekhar AS (2013) Application of artificial neural networks for identification of unbalance and looseness in rotor bearing system. Int J Appl Sci Eng 11(1):69–84

Ricci R, Pennacchi P (2011) Diagnostics of gear faults based on emd and automatic selection of intrinsic mode functions. Mech Syst Signal Process 25(3):821–838

Roemer MJ, Hong CA, Hesler SH (1995) Machine health monitoring and life management using finite element based neural networks. In: Proceedings of the ASME Turbo Expo 5(October), https://doi.org/10.1115/95-GT-243

Ruiming F, Hongzhong M (2006) Application of MCSA and SVM to induction machine rotor fault diagnosis. Proc World Congress Intell Control Autom (WCICA) 2(2):5543–5547. https://doi.org/10.1109/WCICA.2006.1714134

Sadeghian A, Ye Z, Wu B (2009) Online detection of broken rotor bars in induction motors by wavelet packet decomposition and artificial neural networks. IEEE Trans Instrum Meas 58(7):2253–2263. https://doi.org/10.1109/TIM.2009.2013743

Sadoughi M, Hu C (2019) Physics-based convolutional neural network for fault diagnosis of rolling element bearings. IEEE Sens J 19(11):4181–4192. https://doi.org/10.1109/JSEN.2019.2898634

Saimurugan M, Ramachandran K (2014) A comparative study of sound and vibration signals in detection of rotating machine faults using support vector machine and independent component analysis. Int J Data Anal Tech Strat 6(2):188–204

Santos P, Villa LF, Reñones A, Bustillo A, Maudes J (2012) Wind turbines fault diagnosis using ensemble classifiers. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) 7377 LNAI:67–76, https://doi.org/10.1007/978-3-642-31488-9_6

Shao H, Jiang H, Wang F, Zhao H (2017) An enhancement deep feature fusion method for rotating machinery fault diagnosis. Knowl-Based Syst 119:200–220. https://doi.org/10.1016/j.knosys.2016.12.012

Shao H, Jiang H, Zhang H, Duan W, Liang T, Wu S (2018) Rolling bearing fault feature learning using improved convolutional deep belief network with compressed sensing. Mech Syst Signal Process 100:743–765. https://doi.org/10.1016/j.ymssp.2017.08.002

Shao S, Sun W, Wang P, Gao RX, Yan R (2016) Learning features from vibration signals for induction motor fault diagnosis. Int Symp Flex Autom ISFA 2016:71–76. https://doi.org/10.1109/ISFA.2016.7790138

Sharma S, Malik H, Khatri A (2015) External fault classification experienced by three-phase induction motor based on multi-class ELM. Proc Comput Sci 70:814–820. https://doi.org/10.1016/j.procs.2015.10.122

Singh S, Kumar N (2015) Rotor faults diagnosis using artificial neural networks and support vector machines. Int J Acoust Vib 20(4):153–159. https://doi.org/10.20855/ijav.2015.20.3379

Son JD, Niu G, Yang BS, Hwang DH, Kang DS (2009) Development of smart sensors system for machine fault diagnosis. Expert Syst Appl 36(9):11981–11991. https://doi.org/10.1016/j.eswa.2009.03.069

Sonkul N, Singh J, Vyas NS (2019) Deep learning protocol for condition monitoring and fault identification in a rotor-bearing system from raw time-domain data

Srinivas H, Holla SS (2014) Applications of artificial neural network and wavelet transform for condition monitoring of the combined faults of unbalance and bearing clearance

Srinivas H, Srinivasan K, Umesh K (2010a) Application of artificial neural network and wavelet transform for vibration analysis of combined faults of unbalances and shaft bow. Adv Theor Appl Mech 3(4):159–176

Srinivas HK, Srinivasan KS, Umesh KN (2010b) Role of an artificial neural network and a wavelet transform for condition monitoring of the combined faults of unbalance and cracked rotors. Int J Acoust Vib 15(3):121–127. https://doi.org/10.20855/ijav.2010.15.3264

Sun W, Shao S, Zhao R, Yan R, Zhang X, Chen X (2016) A sparse auto-encoder-based deep neural network approach for induction motor faults classification. Meas J Int Meas Conf 89:171–178. https://doi.org/10.1016/j.measurement.2016.04.007

Sun W, Zhao R, Yan R, Shao S, Chen X (2017) Convolutional discriminative feature learning for induction motor fault diagnosis. IEEE Trans Ind Inform 13(3):1350–1359

Tajik M, Movasagh S, Shoorehdeli MA, Yousefi I (2015a) Gas turbine shaft unbalance fault detection by using vibration data and neural networks. Int Conf Robot Mechatron ICROM 2015:308–313. https://doi.org/10.1109/ICRoM.2015.7367802

Tajik M, Movasagh S, Shoorehdeli MA, Yousefi I (2015b) Gas turbine shaft unbalance fault detection by using vibration data and neural networks. In: 2015 3rd RSI International Conference on Robotics and Mechatronics (ICROM), IEEE, pp 308–313

Tang H, Liao Z, Ozaki Y, Chen P (2020) Stepwise intelligent diagnosis method for rotor system with sliding bearing based on statistical filter and stacked auto-encoder. Appl Sci 10(7):2477

Tang X, Zhuang L, Cai J, Li C (2010) Multi-fault classification based on support vector machine trained by chaos particle swarm optimization. Knowl-Based Syst 23(5):486–490. https://doi.org/10.1016/j.knosys.2010.01.004

Tao H, Mo L, Shen F, Du Z, Yan R (2016) Multi-classifiers ensemble with confidence diversity for fault diagnosis in induction motors. Proc Int Conf Sens Technol ICST 2015:1–5. https://doi.org/10.1109/ICSensT.2016.7796328

Timusk M, Lipsett M, Mechefske CK (2008) Fault detection using transient machine signals. Mech Syst Signal Process 22(7):1724–1749

Uddin J, Kang M, Nguyen DV, Kim JM (2014) Reliable fault classification of induction motors using texture feature extraction and a multiclass support vector machine. Mathematical Problems in Engineering 2014

Udmale SS, Patil SS, Phalle VM, Singh SK (2019a) A bearing vibration data analysis based on spectral kurtosis and convnet. Soft Comput 23(19):9341–9359

Udmale SS, Singh SK, Singh R, Sangaiah AK (2019b) Multi-fault bearing classification using sensors and convnet-based transfer learning approach. IEEE Sens J

Vyas NS, Satishkumar D (2001) Artificial neural network design for fault identification in a rotor-bearing system. Mech Mach Theory 36(2):157–175. https://doi.org/10.1016/S0094-114X(00)00034-3

Walker RB, Vayanat R, Perinpanayagam S, Jennions IK (2014) Unbalance localization through machine nonlinearities using an artificial neural network approach. Mech Mach Theory 75:54–66. https://doi.org/10.1016/j.mechmachtheory.2014.01.006

Wang D, Tsui KL, Miao Q (2017) Prognostics and health management: a review of vibration based bearing and gear health indicators. IEEE Access 6:665–676

Wang J, Liu S, Gao RX, Yan R (2012) Current envelope analysis for defect identification and diagnosis in induction motors. J Manuf Syst 31(4):380–387. https://doi.org/10.1016/j.jmsy.2012.06.005

Wang J, Fu P, Li W, Zhang L (2018) Multimodal deep learning for multiple motor and sensor faults diagnosis. In: Proceedings of the International Symposium on Flexible Automation, The Institute of Systems, Control and Information Engineers, pp 366–372

Wang N, Jiang D, Han T (2016) Dynamic characteristics of rotor system and rub-impact fault feature research based on casing acceleration. J Vibroeng 18(3):1525–1539

Wang XB, Zhang X, Li Z, Wu J (2020) Ensemble extreme learning machines for compound-fault diagnosis of rotating machinery. Knowl-Based Syst 188:105012

Wei Y, Li Y, Xu M, Huang W (2019) A review of early fault diagnosis approaches and their applications in rotating machinery. Entropy 21(4):409

Widodo A, Yang BS (2007) Application of nonlinear feature extraction and support vector machines for fault diagnosis of induction motors. Expert Syst Appl 33(1):241–250. https://doi.org/10.1016/j.eswa.2006.04.020

Widodo A, Yang BS (2008) Wavelet support vector machine for induction machine fault diagnosis based on transient current signal. Expert Syst Appl 35(1–2):307–316. https://doi.org/10.1016/j.eswa.2007.06.018

Widodo A, Yang BS, Han T (2007) Combination of independent component analysis and support vector machines for intelligent faults diagnosis of induction motors. Expert Syst Appl 32(2):299–312. https://doi.org/10.1016/j.eswa.2005.11.031

Wu T, Chung Y, Liu C (2010) Looseness diagnosis of rotating machinery via vibration analysis through hilbert-huang transform approach. J Vib Acoust 132(3):031005

Xia Z, Xia S, Wan L, Cai S (2012) Spectral regression based fault feature extraction for bearing accelerometer sensor signals. Sensors 12(10):13694–13719

Xiao D, Huang Y, Qin C, Shi H, Li Y (2019a) Fault Diagnosis of Induction Motors Using Recurrence Quantification Analysis and LSTM with Weighted BN. Shock and Vibration 2019: https://doi.org/10.1155/2019/8325218

Xiao D, Huang Y, Zhao L, Qin C, Shi H, Liu C (2019b) Domain adaptive motor fault diagnosis using deep transfer learning. IEEE Access 7:80937–80949. https://doi.org/10.1109/ACCESS.2019.2921480

Xiaoxun Z, Jianhong Z, Dongnan H, Zhonghe H (2019) Research on mechanical rotor condition monitoring based on VCNN. Energy Procedia 158:6393–6398. https://doi.org/10.1016/j.egypro.2019.01.208

Xu BG (2012) Intelligent fault inference for rotating flexible rotors using Bayesian belief network. Expert Syst Appl 39(1):816–822. https://doi.org/10.1016/j.eswa.2011.07.079

Xu X, Zheng H, Guo Z, Wu X, Zheng Z (2019) SDD-CNN: small data-driven convolution neural networks for subtle roller defect inspection. Appl Sci (Switzerland). https://doi.org/10.3390/app9071364

Xu X, Cao D, Zhou Y, Gao J (2020) Application of neural network algorithm in fault diagnosis of mechanical intelligence. Mechanical Systems and Signal Processing p 106625

Xue H, Wang H, Song L, Chen P (2011) Structural fault diagnosis of rotating machinery based on distinctive frequency components and support vector machines. In: International Conference on Intelligent Computing, Springer, pp 341–348

Xue H, Wang H, Chen P, Li K, Song L (2013) Automatic diagnosis method for structural fault of rotating machinery based on distinctive frequency components and support vector machines under varied operating conditions. Neurocomputing 116:326–335

Yan J, Hu Y, Guo C (2019a) Rotor unbalance fault diagnosis using DBN based on multi-source heterogeneous information fusion. Procedia Manuf 35:1184–1189. https://doi.org/10.1016/j.promfg.2019.06.075

Yan X, Sun Z, Zhao J, Shi Z, Zhang CA (2019b) Fault diagnosis of active magnetic bearing-rotor system via vibration images. Sensors 19(2):244

Yang X, Yan R, Gao RX (2015) Induction motor fault diagnosis using multiple class feature selection. In: 2015 IEEE International Instrumentation and Measurement Technology Conference (I2MTC) Proceedings, IEEE, pp 256–260

Yang ZX, Wang XB, Zhong JH (2016) Representational learning for fault diagnosis of wind turbine equipment: a multi-layered extreme learning machines approach. Energies 9(6):1–17. https://doi.org/10.3390/en9060379

Yao Y, Li Y, Zhang P, Xie B, Xia L (2018) Data fusion methods for convolutional neural network based on self-sensing motor drive system. In: Proceedings: IECON 2018—44th Annual Conference of the IEEE Industrial Electronics Society 1:5371–5376. https://doi.org/10.1109/IECON.2018.8591160

Yongbo L, Xiaoqiang D, Fangyi W, Xianzhi W, Huangchao Y (2020) Rotating machinery fault diagnosis based on convolutional neural network and infrared thermal imaging. Chin J Aeronaut 33(2):427–438

Younus AM, Yang BS (2012) Intelligent fault diagnosis of rotating machinery using infrared thermal image. Expert Syst Appl 39(2):2082–2091. https://doi.org/10.1016/j.eswa.2011.08.004

Yu T, Han Q (2010) Crack fault identification in rotor shaft with artificial neural network. In: Proceedings—2010 6th International Conference on Natural Computation, ICNC 2010 3(Icnc):1629–1634, https://doi.org/10.1109/ICNC.2010.5583771

Yu W, Huang S, Xiao W (2018) Fault diagnosis based on an approach combining a spectrogram and a convolutional neural network with application to awind turbine system. Energies. https://doi.org/10.3390/en11102561

Yuan SF, Chu FL (2006) Support vector machines-based fault diagnosis for turbo-pump rotor. Mech Syst Signal Process 20(4):939–952. https://doi.org/10.1016/j.ymssp.2005.09.006

Yuan Z, Zhang L, Duan L (2018) A novel fusion diagnosis method for rotor system fault based on deep learning and multi-sourced heterogeneous monitoring data. Meas Sci Technol. https://doi.org/10.1088/1361-6501/aadfb3

Yuan Z, Zhang L, Duan L (2019) Multi-sourced monitoring fusion diagnosis for rotating machinery faults. In: Proceedings—Annual Reliability and Maintainability Symposium 2019-January(1):1–7. https://doi.org/10.1109/RAMS.2019.8769018

Yusuf S, Brown DJ, MacKinnon A, Papanicolaou R (2013) Fault classification improvement in industrial condition monitoring via hidden markov models and naïve bayesian modeling. ISIEA 2013—2013 IEEE Symposium on Industrial Electronics and Applications pp 75–80, https://doi.org/10.1109/ISIEA.2013.6738971

Zhang J, Ma W, Lin J, Ma L, Jia X (2015) Fault diagnosis approach for rotating machinery based on dynamic model and computational intelligence. Meas J Int Meas Conf 59:73–87. https://doi.org/10.1016/j.measurement.2014.09.045

Zhang J, Zhang D, Yang M, Xu X, Liu W, Wen C (2018) Fault Diagnosis for Rotating Machinery with Scarce Labeled Samples: A Deep CNN Method Based on Knowledge-Transferring from Shallow Models. ICCAIS 2018—7th International Conference on Control, Automation and Information Sciences (Iccais):482–487, https://doi.org/10.1109/ICCAIS.2018.8570515

Zhang J, Tian J, Wen T, Yang X, Rao Y, Xu X (2019a) Deep Fault Diagnosis for Rotating Machinery with Scarce Labeled Samples. arXiv preprint arXiv:190709411 1(61806064), https://doi.org/10.1049/cje.2019.0X.0XX

Zhang S, Zhang S, Wang B, Habetler TG (2019b) Machine learning and deep learning algorithms for bearing fault diagnostics-a comprehensive review. arXiv preprint arXiv:190108247

Zhang W, Peng G, Li C, Chen Y, Zhang Z (2017) A new deep learning model for fault diagnosis with good anti-noise and domain adaptation ability on raw vibration signals. Sens (Switzerland). https://doi.org/10.3390/s17020425

Zhang ZY, Wang KS (2014) Wind turbine fault detection based on scada data analysis using ann. Adv Manuf 2(1):70–78

Zhao R, Yan R, Chen Z, Mao K, Wang P, Gao RX (2019) Deep learning and its applications to machine health monitoring. Mech Syst Signal Process 115:213–237

Zhao W, Hua C, Wang D, Dong D (2020a) Fault diagnosis of shaft misalignment and crack in rotor system based on mi-cnn. In: Proceedings of the 13th International Conference on Damage Assessment of Structures, Springer, pp 529–540

Zhao X, Wu J, Zhang Y, Shi Y, Wang L (2018) Fault diagnosis of motor in frequency domain signal by stacked de-noising auto-encoder. Comput Mater Continua 57(2):223–242. https://doi.org/10.32604/cmc.2018.02490

Zhao X, Jia M, Ding P, Yang C, She D, Liu Z (2020b) Intelligent fault diagnosis of multi-channel motor-rotor system based on multi-manifold deep extreme learning machine. IEEE/ASME Transactions on Mechatronics

Zhu X, Hou D, Zhou P, Han Z, Yuan Y, Zhou W, Yin Q (2019a) Rotor fault diagnosis using a convolutional neural network with symmetrized dot pattern images. Meas J Int Meas Conf 138:526–535. https://doi.org/10.1016/j.measurement.2019.02.022

Zhu X, Zhao J, Hou D, Han Z (2019b) An SDP characteristic information fusion-based CNN vibration fault diagnosis method. Shock and Vibration 2019: https://doi.org/10.1155/2019/3926963

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