Mechanical Systems and Signal Processing 108 (2018) 33–47

Contents lists available at ScienceDirect

Mechanical Systems and Signal Processing

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

Review

Artificial intelligence for fault diagnosis of rotating machinery: A review

Ruonan Liu [a,b], Boyuan Yang [a,b], Enrico Zio [c,d], Xuefeng Chen [a,b,*]

[a] State Key Laboratory for Manufacturing Systems Engineering, Xian Jiaotong University, Xi'an 710049, China
[b] School of Mechanical Engineering, Xi'an Jiaotong University, Xi'an 710049, China
[c] Chair on System Science and the Energetic Challenge, EDF Foundation, Laboratoire Genie Industriel, CentraleSupélec, Université Paris-Saclay Grande voie des Vignes, 92290 Chatenay-Malabry, France
[d] Energy Departement, Politecnico di Milano, Milano, Italy

A R T I C L E   I N F O

Article history:
Received 26 December 2016
Received in revised form 29 October 2017
Accepted 9 February 2018
Available online 15 February 2018

Keywords:
Artificial intelligence
Fault diagnosis
k-Nearest neighbour
Naive Bayes
Support vector machine
Artificial neural network
Deep learning
Rotating machinery

A B S T R A C T

Fault diagnosis of rotating machinery plays a significant role for the reliability and safety of modern industrial systems. As an emerging field in industrial applications and an effective solution for fault recognition, artificial intelligence (AI) techniques have been receiving increasing attention from academia and industry. However, great challenges are met by the AI methods under the different real operating conditions. This paper attempts to present a comprehensive review of AI algorithms in rotating machinery fault diagnosis, from both the views of theory background and industrial applications. A brief introduction of different AI algorithms is presented first, including the following methods: *k*-nearest neighbour, naive Bayes, support vector machine, artificial neural network and deep learning. Then, a broad literature survey of these AI algorithms in industrial applications is given. Finally, the advantages, limitations, practical implications of different AI algorithms, as well as some new research trends, are discussed.

Contents

* Corresponding author at: State Key Laboratory for Manufacturing Systems Engineering, Xian Jiaotong University, Xian 710049, China.
E-mail addresses: liuruonan0914@stu.xjtu.edu.cn (R. Liu), yangboyuanxjtu@163.com (B. Yang), enrico.zio@polimi.it (E. Zio), chenxf@mail.xjtu.edu.cn (X. Chen).

https://doi.org/10.1016/j.ymssp.2018.02.016
0888-3270/

1. Introduction

With the rapid development of technology and science, mechanical equipment in modern industry becomes more and more functional and complex. Rotating machinery is among the most important equipments in modern industrial applications. Fault diagnosis of rotating machinery becomes the most critical aspect in system design and maintenance.

Fault diagnosis of rotating machinery is a technique of fault detection, isolation and identification, which can be used applied on the information about operation condition of the equipment [1]. There are three basic tasks of fault diagnosis: (1) determining whether the equipment is normal or not; (2) finding the incipient failure and its reason; (3) predicting the trend of fault development. Therefore, essentially, fault diagnosis can be regarded as a pattern recognition problem regarding the rotating machinery condition. As a powerful pattern recognition tool, artificial intelligence (AI) has attracted great attention from many researchers and shows promise in rotating machinery fault recognition applications.

Due to the variability and richness of the response signals, it is almost impossible to recognize fault patterns directly. Therefore, a common fault diagnosis system often consists of two key steps: data processing (feature extraction), fault recognition [1,2]. Most common intelligent fault diagnosis systems are built based on the preprocessing by feature extraction algorithms [2] to transform the input patterns so that they can be represented by low-dimensional feature vectors for easier match and comparison [3].

Then, the feature vectors are used as the input of AI techniques for fault recognition. The step of fault recognition amounts to mapping the information obtained in the feature space to machine faults in the fault space. Numerous AI tools or techniques have been used, including convex optimization, mathematical optimization, as well as classification-, statistical learning- and probability-based methods. Specifically, classifiers and statistical learning methods have been widely used in fault diagnosis of rotating machinery, that includes, k-nearest neighbor (k-NN) algorithms [4], Bayesian classifier [5], support vector machine (SVM) [6] and artificial neural network (ANN) [7]. Most recently, deep learning approaches have also began to be applied in the field of fault diagnosis [8].

In this paper, we aim at presenting a comprehensive survey on the recent research and development of AI methods for rotating machinery fault diagnosis, from both the views of theory and application. The rest of this paper is organized as follows. Section 2 introduces the basic theory of various AI methods. Section 3 reviews the applications of AI approaches in rotating machinery fault diagnosis. Prospects of AI methods in fault diagnosis are discussed in Section 4. Concluding remarks are drawn in Section 5.

2. Theoretical background of AI approaches

AI algorithms for fault diagnosis of rotating machinery have become popular due to their robustness and adaptation capabilities. Also, they do not require full prior physical knowledge, which may be difficult to obtain in practice. Among the various AI algorithms, k-NN, Naive Bayes classifier, SVM and ANN algorithms have been applied most commonly in fault diagnosis.

2.1. k-Nearest neighbour

k-NN is an instance-based learning algorithm based on the principle that the instances within a dataset will generally exist in close proximity to other instances with similar properties [9]. For a given training set of classified instances $T = \{(x_1, y_1), (x_2, y_2), \ldots, (x_N, y_N)\}$, where x_i is the feature vector of the unlabeled instance, y_i is the label and $y_i = c_1, c_2, \ldots, c_K,\ i = 1, 2, \ldots N$. For a training sample (x, y), the k-NN algorithm searching for the k nearest instances to x based on a given distance metric. The neighbourhood containing these k instances is represented by $N_k(x)$. Then, the label of test sample x can be calculated based on decision rules:

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

where I is the indicator function.

If the instances are tagged with a classification label, then the label of an unclassified instance can be determined by observing the class of its nearest neighbours as shown in Fig. 1 [10].

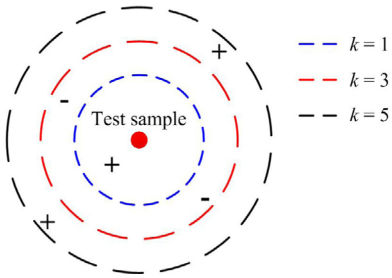

Fig. 1. The diagram of *k*-NN. The equidistant lines are represented by dotted lines. It can be seen that when *k* = 1 or *k* = 5, the test sample is classified to be positive sample; when *k* = 3 is classified to be negative sample.

There are three basic elements in the *k*-NN algorithm: the number of measured instances *k*, the distance metric and the decision rule for classification. Compared with other AI algorithms, *k*-NN shows an advantage of simple implementation.

2.2. Naive Bayes classifier

The Naive Bayes method is a classification method based on Bayes' Theorem and the conditional independence assumption [11]. For a given training set $T = \{(x_1, y_1), (x_2, y_2), \ldots, (x_N, y_N)\}$ with label $y, y_i = c_1, c_2, \ldots, c_K,\ i = 1, 2, \ldots N$, assume there are S_l possible values for $x^l,\ l = 1, 2, \ldots, n$; and there are K possible values for Y. Naive Bayes first learns the joint probability distribution $P(X, Y)$ of the input and output by the conditional probability distribution based on the conditional independence assumption:

$$\begin{aligned} P(X = x|Y = c_j) &= P(X^{(1)} = x^{(1)}, \ldots, X^{(n)} = x^{(n)}|Y = c_j) \\ &= \Pi_{l=1}^{n} P(X^{(l)} = x^{(l)}) \qquad j = 1, 2, \ldots, K \end{aligned} \tag{2}$$

Then, based on the learnt model, the output label y with the biggest posterior probability for the given input x can be calculated via Bayes' Theorem:

$$P(Y = c_j|X = x) = \frac{P(X = x|Y = c_j)P(Y = c_j)}{\sum_j P(X = x|Y = c_j)P(Y = c_j)} \tag{3}$$

and

$$y = \arg\max_{c_j} P(Y = c_j)\Pi_l P(X^{(l)} = x^{(l)}|Y = c_j) \tag{4}$$

Naive Bayes classifier is a common method for classification because it is easy to implement with high efficiency.

2.3. Support vector machine

SVM is a computational learning method for small samples classification [12]. Algorithmically, SVM builds optimal separating hyperplane $f(x) = 0$ between data sets by solving a constrained quadratic optimization problem based on the structural risk minimization (SRM) [13,14].

$$y = f(x) = W^T x + b = \sum_{i=1}^{N} W_i x_i + b \tag{5}$$

where W is a N-dimensional vector and b is a scalar. The optimal separating hyperplane is the separating hyperplane that creates the maximum distance between the plane and the nearest data, that is, the maximum margin as shown in Fig. 2. By converting the optimization problem with Kuhn-Tucker condition into the equivalent Lagrangian dual quadratic optimization problem, the classifier based on the support vector can be obtained.

2.4. Artificial neural network

ANN is believed to be the most commonly used algorithm. For its most popular form, there are three components in an ANN: input layer, hidden layer and output layer. Units in hidden layer are called hidden units, because their values are not observed. ANN is an intelligence technique based on a number of simple processors or neurons, as shown in Fig. 3. The circles labeled "+1" are intercept terms and are called bias units. Fig. 3(a) is a human neuron and Fig. 3(b) is a simple model of ANN. a_{ij} represents the jth neuron unit in the ith layer.

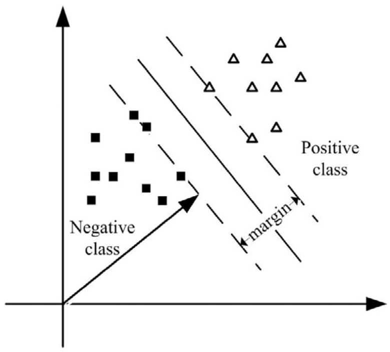

Fig. 2. The optimal hyperplane for a binary classification by SVM.

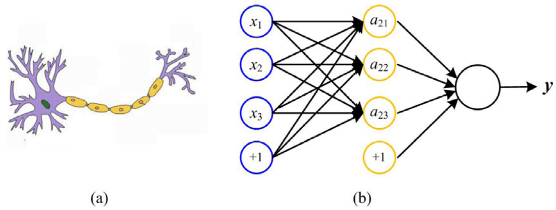

Fig. 3. Human neuron and a MLP with two hidden layers.

The "neuron" in ANN is a computational unit that takes as input x_1, x_2, x_3 and an intercept term. The output y can be obtained by:

$$y = f(W^T x) = f\left(\sum_{i=1}^{3} W_i x_i + b\right) \tag{6}$$

where f is called the activation function, often chosen to be the sigmoid function; W are the ANN model parameters (or weights); b is a scalar. An ANN interconnects many "neurons" and the output of a neuron can be the input of another. The weights W are obtained by an iterative training procedure, based on known input–output patterns.

ANNs implement algorithms that attempt to achieve a neurological related performance, such as learning from experience, making generalizations from similar situations and judging states where poor results were achieved in the past.

2.5. Deep learning

For many tasks, it is difficult to know what features should be extracted to feed to the AI algorithms [3]. Aiming at learning feature hierarchies with features from higher levels of the hierarchy formed by the composition of lower level features [15], deep learning methods have the potential to overcome the aforementioned deficiencies in current intelligent fault diagnosis methods [16]. The Venn diagram of the relationship between different AI disciplines is shown in Fig. 4. In deep learning methods, automatically learning features at multiple levels of abstraction allows a system to learn complex functions mapping the input to the output directly from data. In order to learn these complicated functions, deep architectures are needed, which are composed of multiple levels of non-linear operations. Through the deep architectures, deep learning-based methods are able to adaptively capture the representation information from natural input signals through non-linear transformations and approximate complex non-linear functions with a small error.

Recent methods based on deep learning have demonstrated state-of-the-art performance in a wide variety of tasks, such as computer visual, audio recognition, natural language processing, as well as fault diagnosis [17,8]. Thus far, deep learning methods such as autoencoder, restricted boltzmann machine (RBM) and deep belief network (DBN) have also been used for rotating machinery fault diagnosis.

An autoencoder is an unsupervised learning algorithm that applies backpropagation, setting the target values to be equal to the inputs. As depicted in Fig. 5, an autoencoder NN comprises two parts: encoder network and decoder network. The encoder part transforms the input data to a low-dimensional space and the decoder part reconstructs the inputs from the corresponding codes.

The autoencoder tries to learn a function $h_{W,b} \approx x$. Therefore, if there is structure or relationship in the data, the autoencoder will be able to discover those correlations and learn a low-dimensional representation of the original data. Multiple

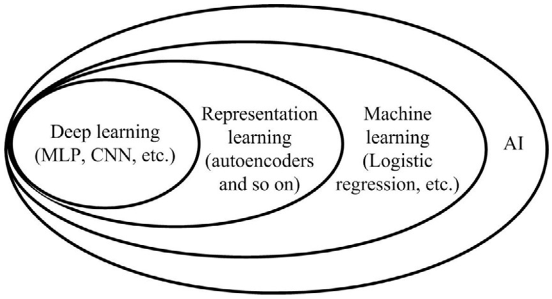

Fig. 4. The Venn diagram of the relationship between different AI disciplines.

layers of sparse autoencoders, in which the outputs of each layer is wired to the inputs of the successive layer, can constitute a stacked autoencoder.

DBN is also a useful tool for fault diagnosis, which is a deep network constructed by multilayer Restricted Boltzmann machine (RBM). Boltzmann machine (BM) is an energy-based model, which is a particular form of log-linear Markov Random Field (MRF). That is, the energy function is linear in its free parameters. To make it powerful enough to represent complicated relations, some hidden variables are considered. By increasing the number of hidden variables (or hidden units), the modeling capacity of the BM can be improved. DBM is an undirected bipartite graphical model which further restrict BM to those without visible-visible and hidden-hidden connections, as shown in Fig. 6.

The energy function $E(v,h)$ of an RBM is defined as:

$$E(v,h) = -b'v - c'h - h'Wv \tag{7}$$

where W are the weights connecting hidden and visible units. b and c are the offsets of the visible and hidden layers, respectively. Eq. (7) can be translated to the free energy formula:

$$F(v) = -b'v - \sum_i log \sum_{h_i} e^{h_i(c_i + W_i v)} \tag{8}$$

Due to the specific structure of RBM, visible and hidden units are conditionally independent. Therefore, we can get:

$$p(h|v) = \prod_i p(h_i|v)$$
$$p(v|h) = \prod_j p(v_j|h)(9)$$

By increasing the number of hidden layers, deep Boltzmann machine (DBM) can be obtained. To obtain DBN, Bayes belief network is emploied at the part closer to the visible layer, while RBM is used at the part away from the visible layer, as shown in Fig. 7. In Deep Belief Networks (DBNs) [18–20], the key principles are: (1) unsupervised learn representations to pre-train each layer; (2) unsupervised train one layer at a time, on top of the previously trained ones; (3) fine-tune all the layers by supervised training.

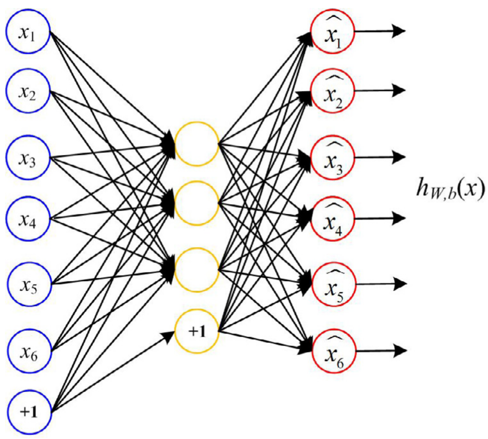

Fig. 5. A graphical representation of an autoencoder.

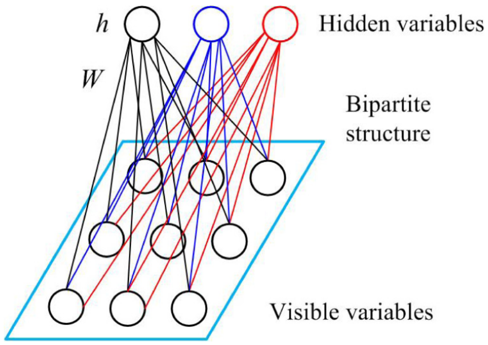

Fig. 6. A graphical depiction of an RBM.

The performance of most deep learning algorithms is enhanced on large datasets and strong computation ability [21]. As shown in Fig. 8, for small datasets in area I, the performance of deep learning is not different from the traditional machine learning; as the dataset dimension increases in area II, the performance of deep learning also increases, while traditional algorithm changes little.

3. Applications of AI in fault diagnosis of rotating machinery

This section presents the applications of AI approaches in fault diagnosis of rotating machinery.

3.1. Preprocessing for fault diagnosis of rotating machinery

Most AI techniques are applied in practical applications of industrial systems, in combination with a feature extractor [22]. According to the literature, many signal-processing methods have been applied for preprocessing for fault recognition, including time-domain analysis, frequency-domain analysis, time–frequency analysis. Time-domain analysis, such as mean, variance, kurtosis estimation, is directly based on the time signal itself. Empirical mode decomposition (EMD) is a most commonly used time-domain method for feature extraction [23]. Frequency-domain analysis, such as fast Fourier transform (FFT) and bispectrum analysis, is based on the transformed signal in frequency domain. Time–frequency analysis such as short time Fourier transform (STFT), Hilbert-Huang transform (HHT) [24], wavelet transform [25], wavelet packet transform (WPT) and sparse decomposition [26] are often used for feature extraction.

On the other hand, with the recent development of the Internet of Things, wireless communications, e-commerce, and smart manufacturing, the amount of data collection has grown in an exponential manner [27]. Therefore, data fusion methods and feature dimensionality reduction methods are also in urgent need for simplifying fault detection and big data handling. Refs. [28,29] have summarized the state of the art of data fusion methods, as well as their benefits and challenging aspects. Recently, data fusion, as well as feature dimensionality reduction techniques has been applied widely in industrial systems. For example, Ref. [30] has proposed a way to reduce the number of vibration sensors by advanced signal processing and data fusion techniques, which can in turn reduce the dependency on the experience and subjective judgments in practical applications. A new method has been proposed, which can be used to construct a single composite spectrum using all the measured vibration data set via the data fusion technique [31]. In [32], a composite spectrum data fusion method is proposed for rotating machines fault diagnosis, which can simplify the diagnosis and obtain increased performance. A data fusion technique has been applied to compute composite higher-order spectra, and PCA has been used for fault separation

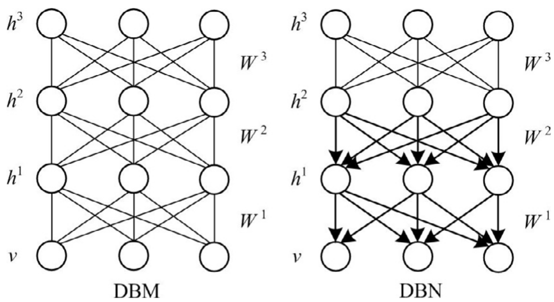

Fig. 7. DBM and DBN.

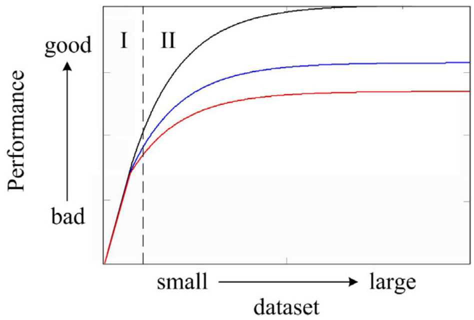

Fig. 8. Performance comparison of traditional machine learning methods and deep learning methods.

and diagnosis [33,34]. The sensitivity of the method has been examined through experimental and industrial cases to verify its robustness [35].

3.2. k-NN in fault diagnosis of rotating machinery

In literature, *k*-NN has been proved to work well and applied widely in fault diagnosis. As an instance-based algorithm, *k*-NN can be used both for classification and regression.

Wang [4] identifies 5 different gear crack levels via *k*-NN and the redundant statistical features constructed by Daubechies 44 binary WPT. Jung and Koh [36] exploit the multiscale energy analysis of discrete wavelet transformation to obtain a low-dimensional feature subset of data as the input of *k*-NN for bearing fault classification. In [37], Pandya et al. present a bearing fault diagnosis technique, which uses the features extracted by HHT as the input of *k*-NN classifiers. This paper also compares the performance of various AI methods, such as Naive Bayes, ANN, and weighted ANN. He [38] proposes a two-step approach for plastic bearing fault diagnosis, which firstly extracts frequency and time domain features by envelope analysis and EMD; then, the frequency domain features are used to identify bearing outer race faults, and time domain features are used to build a *k*-NN classifier to identify other types of bearing faults. Yaqub et al. [39] propose a inchoate fault detection framework based on *k*-NN classifier and an adaptive feature extractor, which is built based on higher order cumulants and wavelet transform. Lei and Zuo [40] propose a gear crack level identification method based on a two-stage feature selection via Euclidean distance evaluation technique and a weighted *k*-NN classifier. Based on the discovered characteristic vibration source by kernel independent component analysis, Li et al. [41] propose a gearbox multi-fault diagnosis method, which uses locally linear embedding (LLE) to reduce the dimension of the original fault feature vector extracted by WPT and, then, *k*-NN is applied to reduce the feature vector to identify the fault pattern of the gearbox.

k-NN is always applied in combination to different reduction methods, such as principal component analysis (PCA), kernel PCA (KPCA) and contribution analysis (CA). Li et al. [42] firstly compress the high-dimensional time-frequency domain feature sets into low-dimensional eigenvectors by supervised uncorrelated orthogonal locality preserving projection and, then, the low-dimensional eigenvectors are utilized as the input of *k*-NN for life grade recognition of rotating machinery. With the feature extracted via spectral kurtosis and cross correlation, Jing et al. [43] form a health index using PCA and *k*-NN to detect bearing faults and monitor the degradation of bearings. Safizadeh [44] utilizes PCA to reduce the dimension of features as the input of *k*-NN to identify the condition of the ball bearing. Zhou et al. [45] provide a contribution analysis-based fault isolation method by decomposing the *k*-NN distance used as the detection index.

There are also some literature works in which *k*-NN and other classifiers are applied on the same applications and compared together. In [46], Moosavian compares two classifiers, *k*-NN and ANN, based on features extracted by power spectral density. Dou and Zhou [47] compare the classification results of *k*-NN, ANN and SVM for intelligent fault diagnosis of rotating machinery.

As most instance-based classifiers, the key issue of *k*-NN is the selection of *k*, which may greatly influence the algorithm performance and should be selected carefully in real applications.

3.3. Naive Bayes classifier in fault diagnosis of rotating machinery

Naive Bayes is a generative model with high learning and predicting efficiency, which is easy to implement [10]. Due to the strong assumption of conditional independence and the need of prior knowledge, naive Bayes method is only appropriate for independent feature vector. However, different from text features, industrial signal features are often interdependent with each other, e.g. the most commonly used statistical features. Therefore, the applications of naive Bayes are often implemented after some dimensionality reduction operation or whitening preprocessing.

For transformer fault diagnosis, Zhao et al. [48] propose a combinatorial method based on Bayesian network and AdaBoostMl algorithm. The analysis results are compared with naive Bayes and other methods. Muralidharan and Sugumaran propose a fault diagnosis system for monoblock centrifugal pumps [49]. In this paper, wavelet analysis is employed for feature extraction, naive Bayes and Bayes network algorithms are used for fault diagnosis. In [50], Seshadrinath et al. compare the induction machines diagnosis results from feedforward ANN and naive Bayes, based on features extracted by the dual tree complex wavelet transform.

The effectiveness of naive Bayes has also been compared with other AI methods. Wang et al. [51] propose a motors defect diagnosis approach, which extracts features from the envelope of the motor current, instead of the motor current itself, as the inputs of three pattern classifiers: naive Bayes, *k*-NN and SVM; the classification effectiveness has been investigated and compared by experiment. Palacios et al. [52] propose a comprehensive evaluation of pattern classification methods for motors fault identification, including naive Bayes, *k*-NN, SVM and ANN, based on the amplitudes of current signals in the time domain. Phuong and Kim [53] present a comprehensive multifault diagnosis method for incipient bearing failures, which firstly extracts features by a WPT-based kurtogram; then, LDA is used to select discriminant features as the input of a naive Bayes classifier to classify the bearing fault conditions. Wan et al. [54] compare the prediction accuracies of some common statistical models such as naive Bayes and linear discriminant analysis (LDA), based on the features processed by dimensional reduction methods such as PCA, KPCA, Laplacian Eigenmaps and local linear embedding. Flett and Bone [55] propose a fault detection and diagnosis system for a diesel internal combustion engine valve train and compare the results of five classification methods, including naive Bayes, ANN, decision trees, *k*-NN and LDA. Duan et al. [56] present a segmented infrared image analysis for rotating machinery fault diagnosis, which applies an image segmentation approach to enhance the feature extraction to infrared image analysis. Then, a feature fusion method is applied to obtain features from selected regions for fault diagnosis by both naive Bayes classifier and SVM.

3.4. SVM in fault diagnosis of rotating machinery

In SVM, the selection of the kernel is very important and directly influences the separability of the samples and the classification performance. Moreover, due to the lack of feature extraction ability, SVM is generally used in combination with signal processing techniques for feature extraction. For different problems, several SVM architectures and algorithms have been developed in literature.

Wu and Meng applied SVM to analyze the full-spectrum experimental data for the diagnosis of rotor malfunctions [57], which indicates that the full-spectrum cascade can also give useful features for SVM. Tang et al. [58] propose a multi-fault classification model of rotating machines by training SVM using chaos particle swarm optimization (CPSO-SVM). Salahshoor et al. [59] integrate a common framework into a SVM classifier with an adaptive neuro-fuzzy inference system classifier, to enhance the fault detection and diagnostic tasks. With a statistical feature selection by decision tree, SVM is used by Saimurugan et al. [60] to determine the condition of rotating machines. Along with continuous wavelet transform (CWT), SVM is used for bearing fault detection of induction motors [61]. Instead of using vibration signals in traditional methods, Li et al. used SVM to analyze the acoustic emission signals based on pseudo Wigner-Ville Distribution (PWVD) for the diagnose and prediction of different rotor cracks depth [62]. Li et al. [63] present a method for mechanical fault diagnosis, based on the redundant second generation WPT, neighborhood rough set (NRS) and SVM for fault detection, attribute reduction and pattern classification. Liu et al. [64] propose an intelligent method based on a short-time matching atom decomposition method and SVM for bearing fault diagnosis. Banerjee and Das [65] investigate a hybrid method for fault signal classification based on sensor data fusion by using SVM and STFT techniques. Soualhi et al. [66] use the HHT to extract health indicators from vibration signals to tack the degradation of critical components in bearings; then, the degradation states are detected by SVM.

In order to meet the demands of real engineering applications, many improved SVMs for fault diagnosis of rotating machinery are also proposed. Combining EMD and multi-class transductive support vector machine (TSVM), Shen et al. [67] present a novel model for fault diagnosis, which is applied to diagnose the faults of a gear reducer. Gryllias and Antoniadis [68] propose a hybrid two-stage, one-against-all Support Vector Machine (SVM) approach for automated diagnosis of defective rolling element bearings. Liu et al. [69] present a wavelet support vector machine (WSVM) multi-fault classification model to analyze vibration signals from rolling element bearings. Zhang and Zhou [70] present a procedure based on ensemble empirical mode decomposition (EEMD) and optimized SVM for multi-fault diagnosis of rolling element bearings. Shen et al. [71] propose a new intelligent fault diagnosis scheme, based on the extraction of statistical parameters from WPT, a distance evaluation technique and a support vector regression (SVR)-based generic multi-class solver; the effectiveness of the intelligent fault diagnosis scheme is validated separately using datasets from bearing and gearbox test rigs. Zhao et al. [72] propose the combination of an analytic selection method of prior selection followed by a genetic algorithm (ASGA) for SVR parameters optimization. The algorithm, named ASGA-SVR, is used for forecasting the reliability of a system. Xu and Chen [73] present an intelligent fault identification method of rolling bearings based on least squares support vector machine optimized by improved particle swarm optimization (IPSO-LSSVM). A modified PSO algorithm is used to optimize the parameters of LSSVM and, then, the optimized model is used to identify the fault patterns of rolling bearings. Tang et al. [74] propose a novel fault diagnosis method based on manifold learning and Shannon wavelet SVM for wind turbine transmission systems. Based on the techniques of Hilbert Huang transform (HHT) and SVM, Wang et al. [75] develop a noise-based intelligent HHT-SVM method for engine fault diagnosis.

The effectiveness of SVM has also been compared with other AI methods applied to the same practical problems. By utilizing time-domain features as inputs, B. Samanta [76] compares the performance of gear fault detection using ANN and SVM: the results give the classification accuracy of SVM compared to that of ANN. With wavelet-based features, the relative efficiency of ANN and proximal support vector machine (PSVM) in classifying faults in the bevel gear box is compared in [77]. Kankar et al. [78] compare the effectiveness of ANN and SVM for fault diagnosis of ball bearings, based on statistical features. With the kernel neighborhood rough sets feature selection method, Zhu [79] uses three classification algorithms, namely, classification and regression trees and radial basis function support vector machine (RBFSVM), to test 10 fault datasets of rolling bearings.

3.5. ANN in fault diagnosis of rotating machinery

Numerous research activities have shown that ANN has powerful pattern classification and pattern recognition capabilities. As a result, ANN is one of the classifiers most commonly used in intelligent fault diagnosis. Like SVM, the application of ANN is also along with proper feature extractors.

Multi-layer perceptron (MLP), is an ANN made of units arranged in layers with only forward connections to units in subsequent layers [80], which has been used in several applications. Rafiee et al. [81] adopt a MLP network with a 16:20:5 structure of input-hidden-output layers for fault detection and identification of gearboxes. The standard deviation of wavelet packet coefficients of preprocessed signals are used as feature vectors. The MLP is small in size but with 100% accuracy. As the model uncertainty and disturbances of ANNs are usually difficult to eliminate, Mrugalski [82] applies the outer bounding ellipsoid algorithm to estimate the parameters of the MLP and the corresponding model uncertainty, for application of robust fault detection. Sadeghian et al. [83] present an algorithm for induction motors online detection of rotor bar breakage, based on the combination of wavelet packet decomposition (WPD) and ANN. Sanza et al. [84] use a MLP network to identify the damage severity and mesh stiffness reduction. This network is fed with statistical parameters obtained from the wavelet coefficients derived for the most sensitive levels of decomposition to damage; the output is the drop in the averaged torsional meshing stiffness when a failure appears, which is highly related with local failure. Bin et al. [85] utilize the feature extracted by WPT and EMD, as the inputs of a classical three-layers Back-Propagation (BP) neural network to identify rotor lateral early crack in rotating machinery. Rotor-related faults also account for a huge proportion of industrial machinery downtime [86,87]. According to the literature, ANN shows advantages in rotor-related faults diagnosis. Refs. [88,89] have summarized several techniques for common rotor faults diagnosis, including ANN, PCA, SVM and so on. And Kankar et al. [90] compared the performances of ANN and SVM for fault diagnosis of rotor bearing systems: the experimental results indicate that ANN has a higher classification accuracy than SVM in the cases considered in the study.

The radial basis function (RBF) network has a feedforward structure, consisting of only one hidden layer with no weighted connections and fully interconnected to the output layer. The architecture of RBF network is illustrated in Fig. 9. Compared with MLP, RBF network is faster to train [91]. Wu and Tommy [92] develop a RBF network-based fault detection system for machine fault detection and propose a cell-splitting grid algorithm so that the optimal network architecture can be determined automatically. This is tested with unbalanced electrical faults and mechanical faults operating at different rotating speeds. Lei et al. [93] apply an intelligent fault diagnosis method based on WPT, EMD and RBF network to slight rub fault diagnosis of a heavy oil catalytic cracking unit. Zhanga et al. [94] propose a high-dimensional machinery fault diagnosis method based on the combination of a hybrid model which combines multiple feature selection models and a weighted voting scheme based on the accuracy measurement of RBF network.

Probabilistic neural network (PNN) is similar with MLP in structure. The basic differences are the use of exponential activation functions and the connection patterns between neurons. The neurons of PNN at the hidden layers are not fully connected, as the structure in Fig. 10 shows. Due to the smaller number of connections, PNN is normally easier to train than MLP. Li et al. [95] utilize WPT, EMD, Wigner-Ville distributions and an autoregressive algorithm for feature extraction. Then, the extracted features are used as the inputs of PNN for gears fault diagnosis. Wang and Chen [96] propose a sequential diagnosis technique with a fuzzy PNN to sequentially identify fault types. Yaghobi [97] constructs a Meyer wavelet probabilistic neural network (MWPNN) by integrating PNN with discrete wavelet transform, and applies it for internal faults diagnosis of a generator. Wang et al. [98] combine the tensor manifold time-frequency characteristic parameters and PNN to classify the rolling-bearing failure samples.

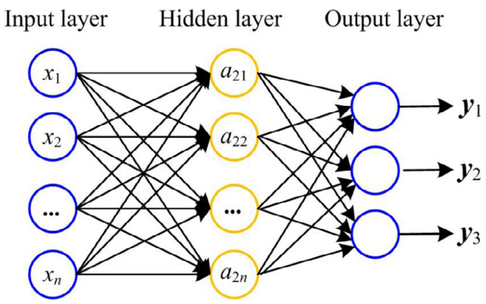

Fig. 9. RBF network structure.

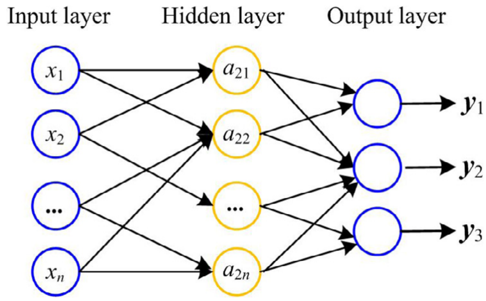

Fig. 10. PNN network structure.

XSamanta et al. [99] use time-domain features as inputs of three types of ANN: MLP, RBF and PNN. The performance of these three methods are presented and compared. With the help of GA as feature selection algorithm, the classification accuracy are almost always 100%, with just 3–6 training samples.

There are also many other improved ANNs for fault diagnosis of rotating machinery. Recurrent neural network (RNN) is a network in which the outputs of some neurons are fed back to the same neurons or neurons in preceding layers; this provides the ANN with a dynamic memory [100]. After feature extraction and selection by a nonlinear manifold learning technique, Prieto et al. [101] use a hierarchical neural network to perform classification for bearing fault detection. Hong et al. [102] apply the EMD-self-organizing map (EMD-SOM) method to analyze vibration signals and calculate a confidence value on the bearing health state.

In practice, a regularization term and prior knowledge are often added to the ANN model to avoid over-fitting and achieve better accuracy.

3.6. Deep learning in fault diagnosis of rotating machinery

Deep learning methods are usually based on deep architectures of computational elements. A classic and common example of such an element is ANN [15], which can be used to build a deep neural network (DNN) with deep architecture.

Since the idea of deep learning appeared [16], it has attracted a lot of attention from researchers in the field of computer vision, speech recognition and natural language processing. Various deep learning algorithms, such as autoencoders, stacked autoencoders [103], DBM and DBN [16], have been applied successfully also in fault diagnosis.

Recently, Lei et al. [8] have used stacked autoencoder to learn features from mechanical vibration signals directly; then, the softmax regression is employed to classify the health conditions. The combination of stacked autoencoder and softmax regression is able to obtain high accuracy for bearing fault diagnosis. Li and Wang [104] use stack autoencoders to initialize the initial weights and offsets of the MLP and provide expert knowledge for spacecraft conditions. Jia et al. [105] utilize frequency spectra to train a stacked autoencoder for fault diagnosis of rotating machinery. The dimension of the output layer is determined according to the number of conditions. The effectiveness of the stacked autoencoder is validated by four roller bearing datasets and a planetary gearbox dataset.

By collecting DBNs by layer and extracting the wavelet packet energy as feature, Gan et al. [17] propose a novel hierarchical diagnosis network with a two-layer HDN for the hierarchical identification of mechanical systems. The first layer aims at identifying fault types and the second one is developed to further recognize fault severity ranking from the result of the first layer. Shao et al. [106] propose an optimization DBN for rolling bearing fault diagnosis. In the paper, stochastic gradient descent is used to fine-tune the W of RBM. Then, particle swarm is used to decide the optimal structure of the trained DBN. In [107], Fink and Zio et al. propose a fuzzy classification approach applying a combination of Echo-State Networks and a RBM for predicting potential railway rolling stock system failure. Chuan Li et al. [108] address a multimodal deep support vector classification approach, which employs separation fusion-based deep learning in order to perform fault diagnosis tasks for gearboxes.

Because the use of deep learning-based methods for fault diagnosis has developed recently, it is not as widely used as in other fields. Nevertheless, it holds great promises due to the excellent performance it owns thus far. As a word of caution, in practice, due to the deep architecture, the number of parameters increases, leading to the risk of over-fitting. To avoid this problem, many tricks are developed, including early stopping, regularization, drop out, and so on.

4. Discussion, limitation and future trends

Various AI techniques of feature extraction and pattern recognition, as well as their applications to fault diagnosis, have been discussed. It can be seen that they have been applied in a wide range of problems of rotating machinery fault diagnosis. To sum up, they have their own advantages and weaknesses.

Table 1
List of AI algorithms and their advantages and limitations.

Algorithm	Advantages	Limitations
k-NN	1. Mature theory and easy to implement 2. Can be used for both classification and regression	1. Large computation 2. Needs lots of storage space 3. The selection of k influence rightarrow too much
Naive Bayes	1. Robust to missing values 2. Requires little storage space 3. Good physical explanation ability	1. Strong prior assumptions 2. Combinatorial explosion and computation problem 3. Need of prior probability
SVM	1. High classification accuracy 2. Can deal with high-dimensional features	1. Low efficiency for big data 2. No physical meaning
ANN	1. High classification accuracy 2. Good approximation of complex nonlinear function	1. Many parameters and easy to over-fitting 2. No physical meaning 3. Training procedure cannot be seen
Deep learning	1. Learning features and recognizing fault automatically 2. Learning more complex structure from data due to the deep architecture 3. Do not need the feature extractor	1. Needs of large samples 2. No physical meaning 3. Training for long time

Table 2
Performance comparison of AI algorithms.

	ANN	SVM	k-NN	Naive Bayes	Deep learning
Accuracy in general	***	****	**	*	****
Classification speed	****	****	*	****	**
Robustness to noise	**	**	*	***	****
Dealing with overfitting	*	**	***	***	***
Physical explanation	*	*	***	****	*
Robustness to parameters	*	*	***	****	**

k-NN is an instance-based learning algorithm that delays the induction or generalization process until classification is performed. Compared with k-means, it only computes the instance of the nearest points instead of global distance. Therefore, it requires less computation time during the training phase than eager-learning algorithms such as Bayes nets and ANN, but more computation time during the classification process.

Naive Bayes algorithm is characterized by the explicit underlying probability model, differently from other AI techniques; it provides a probability that an instance belongs in each class, rather than simply a classification.

SVM has excellent performance in generalization, also with few training data and thanks to its appropriate nonlinear mapping using kernel functions, data from two or more categories can always be separated by a hyperplane [109]. Thus it can produce high accuracy in classification tasks for rotating machinery fault diagnosis and condition monitoring.

ANN is a computational model that mimics the human brain structure, which consists of simple processing elements connected in a complex layer structure which enables the model to approximate a complex non-linear function with multi-input and multi-output. The structure of ANN can be versatile and by changing its architecture, it can achieve good fault diagnosis performance in many rotating machinery applications.

Deep learning provides an effective way to learn features automatically at multiple levels of abstraction, allowing to learn complex input-to-output functions directly from data, without depending on feature extractors, which can be of great benefit for industrial rotating machinery fault diagnosis.

Generally, SVM, ANN and deep learning methods tend to perform better when dealing with multi-dimensions and continuous features; while k-NN and naive Bayes tend to perform better when dealing with discrete features [110]. On the other hand, k-NN and naive Bayes algorithms are all explainable with clear physical meaning, whereas SVM, ANN and deep learning methods have poor interpretability.

Depending on the application, the performance are different from different algorithms. Table 1 lists a brief summary for different AI techniques, including their strengths and weaknesses. As no single learning algorithm can uniformly outperform other algorithms over all datasets. Features of learning techniques are also compared in Table 2.

In the future, AI techniques still need more attention because they offer a promising way to deal with industrial data. Besides those applications summarized in the previous section, there are some new trends in AI methods for rotating machinery fault diagnosis. In the following are list some directions:

(1) It seems important to design and develop hybrid systems for industrial applications, to benefit from the different characteristics of different algorithms.

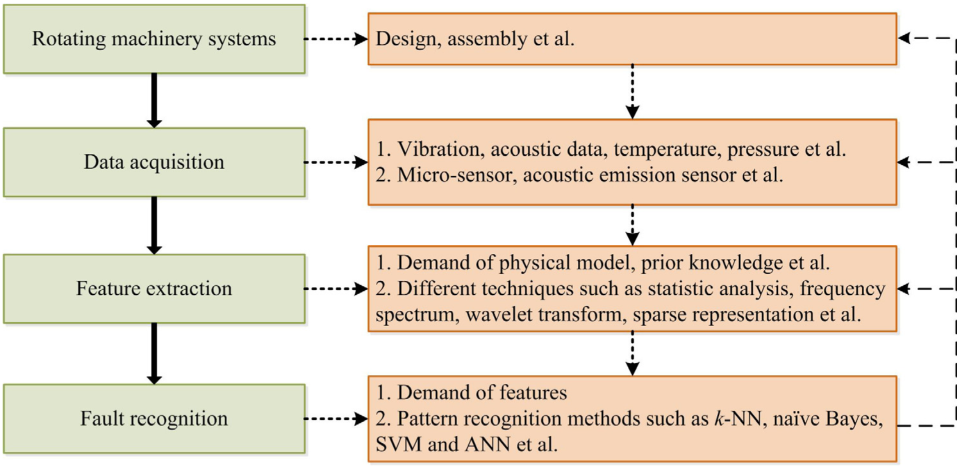

Fig. 11. Fault diagnosis framework to operation and maintenance of engineered systems.

(2) With the development of AI techniques and the rise of deep learning, intelligent diagnosis is going to be the future direction of fault diagnosis development. On the other hand, in the future diagnostic systems, not only data-driven AI methods, but also the consideration of failure mechanism and prior knowledge should be utilized and integrated closely to improve diagnostic performance.

(3) At present, fault diagnostic systems are mostly built as the combination of individual parts, such as data collection, feature extraction and dimensionality reduction, fault recognition, as shown in Fig. 11, with little consideration of the whole diagnostic system. Deep learning techniques provide a way to integrate the feature extraction part and pattern recognition part into a system. More than this, a complete integrated and automated diagnostic system should be paid more attention.

5. Conclusion

Fault diagnosis for rotating machinery is vital to reducing maintenance costs, operation downtime and safety hazards. In this paper, a number of AI techniques have been surveyed for rotating machinery diagnosis. *k*-NN-based, Naive Bayes-based, SVM-based, ANN-based and deep learning-based fault diagnosis for rotating machinery have been summarized both from the theoretical background and industrial application points of view. With the AI techniques becoming more and more mature, it is believed that AI techniques will continue to be attractive and powerful for rotating machinery fault diagnosis.

Acknowledgment

This work is supported by the National Natural Science Foundation of China (No. 51335006) and National Key Basic Research Program of China (No. 2015CB057400).

References

[1] A.K. Jardine, D. Lin, D. Banjevic, A review on machinery diagnostics and prognostics implementing condition-based maintenance, Mech. Syst. Sign. Process. 20 (7) (2006) 1483–1510.
[2] H. Yang, J. Mathew, L. Ma, Fault diagnosis of rolling element bearings using basis pursuit, Mech. Syst. Sign. Process. 19 (2) (2005) 341–356.
[3] Y. LeCun, L. Bottou, Y. Bengio, P. Haffner, Gradient-based learning applied to document recognition, Proc. IEEE 86 (11) (1998) 2278–2324.
[4] D. Wang, K-nearest neighbors based methods for identification of different gear crack levels under different motor speeds and loads: revisited, Mech. Syst. Sign. Process. 70 (2016) 201–208.
[5] P. Baraldi, L. Podofillini, L. Mkrtchyan, E. Zio, V.N. Dang, Comparing the treatment of uncertainty in bayesian networks and fuzzy expert systems used for a human reliability analysis application, Reliab. Eng. Syst. Saf. 138 (2015) 176–193.
[6] V. Vapnik, The Nature of Statistical Learning Theory, Springer Science & Business Media, 2013.
[7] S. Haykin, N. Network, A comprehensive foundation, Neural Netw. 2 (2004).
[8] Y. Lei, F. Jia, J. Lin, S. Xing, S. Ding, An intelligent fault diagnosis method using unsupervised feature learning towards mechanical big data.
[9] T. Cover, P. Hart, Nearest neighbor pattern classification, IEEE Trans. Inform. Theory 13 (1) (1967) 21–27.
[10] S.B. Kotsiantis, I. Zaharakis, P. Pintelas, Supervised Machine Learning: A Review of Classification Techniques; 2007.
[11] T. Mitchell, Generative and Discriminative Classifiers: Naive Bayes and Logistic Regression; 2005. Manuscript available at <http://www.cs.cm.edu/~tom/NewChapters.html>.
[12] A. Widodo, B.-S. Yang, Support vector machine in machine condition monitoring and fault diagnosis, Mech. Syst. Sign. Process. 21 (6) (2007) 2560–2574.
[13] N. Cristianini, J. Shawe-Taylor, An Introduction to Support Vector Machines and Other Kernel-based Learning Methods, Cambridge University Press, 2000.
[14] B. Scholkopf, A.J. Smola, Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond, MIT Press, 2001.

[15] Y. Bengio, Learning deep architectures for ai, Found. Trends®Mach. Learn. 2 (1) (2009) 1–127.
[16] G.E. Hinton, R.R. Salakhutdinov, Reducing the dimensionality of data with neural networks, Science 313 (5786) (2006) 504–507.
[17] M. Gan, C. Wang, et al, Construction of hierarchical diagnosis network based on deep learning and its application in the fault pattern recognition of rolling element bearings, Mech. Syst. Sign. Process. 72 (2016) 92–104.
[18] G.E. Hinton, S. Osindero, Y.-W. Teh, A fast learning algorithm for deep belief nets, Neural Comput. 18 (7) (2006) 1527–1554.
[19] Y. Bengio, P. Lamblin, D. Popovici, H. Larochelle, et al, Greedy layer-wise training of deep networks, Adv. Neural Inform. Process. Syst. 19 (2007) 153.
[20] C. Poultney, S. Chopra, Y.L. Cun, et al., Efficient learning of sparse representations with an energy-based model, in: Advances in Neural Information Processing Systems, 2006, pp. 1137–1144.
[21] Y. LeCun, Y. Bengio, G. Hinton, Deep learning, Nature 521 (7553) (2015) 436–444.
[22] R. Liu, G. Meng, B. Yang, C. Sun, X. Chen, Dislocated time series convolutional neural architecture: an intelligent fault diagnosis approach for electric machine, IEEE Trans. Ind. Infor. 13 (3) (2017) 1310–1320, https://doi.org/10.1109/TII.2016.2645238.
[23] Y. Lei, J. Lin, Z. He, M.J. Zuo, A review on empirical mode decomposition in fault diagnosis of rotating machinery, Mech. Syst. Sign. Process. 35 (1) (2013) 108–126.
[24] Z. Peng, W.T. Peter, F. Chu, A comparison study of improved Hilbert–Huang transform and wavelet transform: application to fault diagnosis for rolling bearing, Mech. Syst. Sign. Process. 19 (5) (2005) 974–988.
[25] R. Yan, R.X. Gao, X. Chen, Wavelets for fault diagnosis of rotary machines: a review with applications, Sign. Process. 96 (2014) 1–15.
[26] B. Yang, R. Liu, X. Chen, Fault diagnosis for a wind turbine generator bearing via sparse representation and shift-invariant K-SVD, IEEE Trans. Ind. Infor. 13 (3) (2017) 1321–1331, https://doi.org/10.1109/TII.2017.2662215.
[27] Y. Lei, F. Jia, J. Lin, S. Xing, S.X. Ding, An intelligent fault diagnosis method using unsupervised feature learning towards mechanical big data, IEEE Trans. Ind. Electron. 63 (5) (2016) 3137–3147.
[28] B. Khaleghi, A. Khamis, F.O. Karray, S.N. Razavi, Multisensor data fusion: a review of the state-of-the-art, Inform. Fusion 14 (1) (2013) 28–44.
[29] A.D. Nembhard, J.K. Sinha, A. Yunusa-Kaltungo, Development of a generic rotating machinery fault diagnosis approach insensitive to machine speed and support type, J. Sound Vib. 337 (2015) 321–341.
[30] J.K. Sinha, K. Elbhbah, A future possibility of vibration based condition monitoring of rotating machines, Mech. Syst. Sign. Process. 34 (1) (2013) 231–240.
[31] K. Elbhbah, J.K. Sinha, Vibration-based condition monitoring of rotating machines using a machine composite spectrum, J. Sound Vib. 332 (11) (2013) 2831–2845.
[32] A. Yunusa-Kaltungo, J.K. Sinha, K. Elbhbah, An improved data fusion technique for faults diagnosis in rotating machines, Measurement 58 (2014) 27–32.
[33] A. Yunusa-Kaltungo, J.K. Sinha, A.D. Nembhard, A novel fault diagnosis technique for enhancing maintenance and reliability of rotating machines, Struct. Health Monit. 14 (6) (2015) 604–621.
[34] A. Yunusa-Kaltungo, J.K. Sinha, A.D. Nembhard, Use of composite higher order spectra for faults diagnosis of rotating machines with different foundation flexibilities, Measurement 70 (2015) 47–61.
[35] A. Yunusa-Kaltungo, J.K. Sinha, Sensitivity analysis of higher order coherent spectra in machine faults diagnosis, Struct. Health Monit. 15 (5) (2016) 555–567.
[36] U. Jung, B.-H. Koh, Wavelet energy-based visualization and classification of high-dimensional signal for bearing fault detection, Knowl. Inform. Syst. 44 (1) (2015) 197–215.
[37] D. Pandya, S. Upadhyay, S. Harsha, Fault diagnosis of rolling element bearing with intrinsic mode function of acoustic emission data using APF-KNN, Expert Syst. Appl. 40 (10) (2013) 4137–4145.
[38] D. He, R. Li, J. Zhu, Plastic bearing fault diagnosis based on a two-step data mining approach, IEEE Trans. Ind. Electron. 60 (8) (2013) 3429–3440.
[39] M.F. Yaqub, I. Gondal, J. Kamruzzaman, Inchoate fault detection framework: adaptive selection of wavelet nodes and cumulant orders, IEEE Trans. Instrum. Meas. 61 (3) (2012) 685–695.
[40] Y. Lei, M.J. Zuo, Gear crack level identification based on weighted k nearest neighbor classification algorithm, Mech. Syst. Sign. Process. 23 (5) (2009) 1535–1547.
[41] Z. Li, X. Yan, Z. Tian, C. Yuan, Z. Peng, L. Li, Blind vibration component separation and nonlinear feature extraction applied to the nonstationary vibration signals for the gearbox multi-fault diagnosis, Measurement 46 (1) (2013) 259–271.
[42] F. Li, J. Wang, B. Tang, D. Tian, Life grade recognition method based on supervised uncorrelated orthogonal locality preserving projection and k-nearest neighbor classifier, Neurocomputing 138 (2014) 271–282.
[43] J. Tian, C. Morillo, M.H. Azarian, M. Pecht, Motor bearing fault detection using spectral kurtosis-based feature extraction coupled with k-nearest neighbor distance analysis, IEEE Trans. Ind. Electron. 63 (3) (2016) 1793–1803.
[44] M. Safizadeh, S. Latifi, Using multi-sensor data fusion for vibration fault diagnosis of rolling element bearings by accelerometer and load cell, Inform. Fusion 18 (2014) 1–8.
[45] Z. Zhou, C. Wen, C. Yang, Fault isolation based on k-nearest neighbor rule for industrial processes, IEEE Trans. Ind. Electron. 63 (4) (2016) 2578–2586.
[46] A. Moosavian, H. Ahmadi, A. Tabatabaeefar, M. Khazaee, Comparison of two classifiers; k-nearest neighbor and artificial neural network, for fault diagnosis on a main engine journal-bearing, Shock Vib. 20 (2) (2013) 263–272.
[47] D. Dou, S. Zhou, Comparison of four direct classification methods for intelligent fault diagnosis of rotating machinery, Appl. Soft Comput. 46 (2016) 459–468.
[48] W. Zhao, Y. Zhang, Y. Zhu, Diagnosis for transformer faults based on combinatorial Bayes network, in: 2nd International Congress on Image and Signal Processing, 2009, CISP'09, IEEE, 2009, pp. 1–3.
[49] V. Muralidharan, V. Sugumaran, A comparative study of Naïve Bayes classifier and Bayes net classifier for fault diagnosis of monoblock centrifugal pump using wavelet analysis, Appl. Soft Comput. 12 (8) (2012) 2023–2029.
[50] J. Seshadrinath, B. Singh, B. Panigrahi, Vibration analysis based interturn fault diagnosis in induction machines, IEEE Trans. Ind. Infor. 10 (1) (2014) 340–350.
[51] J. Wang, S. Liu, R.X. Gao, R. Yan, Current envelope analysis for defect identification and diagnosis in induction motors, J. Manuf. Syst. 31 (4) (2012) 380–387.
[52] R.H.C. Palácios, I.N. da Silva, A. Goedtel, W.F. Godoy, A comprehensive evaluation of intelligent classifiers for fault identification in three-phase induction motors, Electr. Power Syst. Res. 127 (2015) 249–258.
[53] P.H. Nguyen, J.-M. Kim, Multifault diagnosis of rolling element bearings using a wavelet kurtogram and vector median-based feature analysis, Shock Vib. 2015 (2015) 14, https://doi.org/10.1155/2015/320508, Article ID 320508.
[54] X. Wan, D. Wang, W.T. Peter, G. Xu, Q. Zhang, A critical study of different dimensionality reduction methods for gear crack degradation assessment under different operating conditions, Measurement 78 (2016) 138–150.
[55] J. Flett, G.M. Bone, Fault detection and diagnosis of diesel engine valve trains, Mech. Syst. Sign. Process. 72 (2016) 316–327.
[56] L. Duan, M. Yao, J. Wang, T. Bai, L. Zhang, Segmented infrared image analysis for rotating machinery fault diagnosis, Infrared Phys. Technol. 77 (2016) 267–276.
[57] W. Fengqi, G. Meng, Compound rub malfunctions feature extraction based on full-spectrum cascade analysis and SVM, Mech. Syst. Sign. Process. 20 (8) (2006) 2007–2021.
[58] X. Tang, L. Zhuang, J. Cai, C. Li, Multi-fault classification based on support vector machine trained by chaos particle swarm optimization, Knowl.-Based Syst. 23 (5) (2010) 486–490.

[59] K. Salahshoor, M. Kordestani, M.S. Khoshro, Fault detection and diagnosis of an industrial steam turbine using fusion of SVM (support vector machine) and ANFIS (adaptive neuro-fuzzy inference system) classifiers, Energy 35 (12) (2010) 5472–5482.
[60] M. Saimurugan, K. Ramachandran, V. Sugumaran, N. Sakthivel, Multi component fault diagnosis of rotational mechanical system based on decision tree and support vector machine, Expert Syst. Appl. 38 (4) (2011) 3819–3826.
[61] P. Konar, P. Chattopadhyay, Bearing fault detection of induction motor using wavelet and support vector machines (SVMs), Appl. Soft Comput. 11 (6) (2011) 4203–4211.
[62] X. Li, K. Wang, L. Jiang, The application of AE signal in early cracked rotor fault diagnosis with PWVD and SVM, JSW 6 (10) (2011) 1969–1976.
[63] N. Li, R. Zhou, Q. Hu, X. Liu, Mechanical fault diagnosis based on redundant second generation wavelet packet transform, neighborhood rough set and support vector machine, Mech. Syst. Sign. Process. 28 (2012) 608–621.
[64] R. Liu, B. Yang, X. Zhang, S. Wang, X. Chen, Time-frequency atoms-driven support vector machine method for bearings incipient fault diagnosis, Mech. Syst. Sign. Process. 75 (2016) 345–370.
[65] T.P. Banerjee, S. Das, Multi-sensor data fusion using support vector machine for motor fault detection, Inform. Sci. 217 (2012) 96–107.
[66] A. Soualhi, K. Medjaher, N. Zerhouni, Bearing health monitoring based on Hilbert–Huang transform, support vector machine, and regression, IEEE Trans. Instrum. Meas. 64 (1) (2015) 52–62.
[67] Z. Shen, X. Chen, X. Zhang, Z. He, A novel intelligent gear fault diagnosis model based on EMD and multi-class TSVM, Measurement 45 (1) (2012) 30–40.
[68] K.C. Gryllias, I.A. Antoniadis, A support vector machine approach based on physical model training for rolling element bearing fault detection in industrial environments, Eng. Appl. Artif. Intell. 25 (2) (2012) 326–344.
[69] Z. Liu, H. Cao, X. Chen, Z. He, Z. Shen, Multi-fault classification based on wavelet SVM with PSO algorithm to analyze vibration signals from rolling element bearings, Neurocomputing 99 (2013) 399–410.
[70] X. Zhang, J. Zhou, Multi-fault diagnosis for rolling element bearings based on ensemble empirical mode decomposition and optimized support vector machines, Mech. Syst. Sign. Process. 41 (1) (2013) 127–140.
[71] C. Shen, D. Wang, F. Kong, W.T. Peter, Fault diagnosis of rotating machinery based on the statistical parameters of wavelet packet paving and a generic support vector regressive classifier, Measurement 46 (4) (2013) 1551–1564.
[72] W. Zhao, T. Tao, E. Zio, System reliability prediction by support vector regression with analytic selection and genetic algorithm parameters selection, Appl. Soft Comput. 30 (2015) 792–802.
[73] H. Xu, G. Chen, An intelligent fault identification method of rolling bearings based on LSSVM optimized by improved PSO, Mech. Syst. Sign. Process. 35 (1) (2013) 167–175.
[74] B. Tang, T. Song, F. Li, L. Deng, Fault diagnosis for a wind turbine transmission system based on manifold learning and Shannon wavelet support vector machine, Renew. Energy 62 (2014) 1–9.
[75] Y. Wang, Q. Ma, Q. Zhu, X. Liu, L. Zhao, An intelligent approach for engine fault diagnosis based on Hilbert–Huang transform and support vector machine, Appl. Acoust. 75 (2014) 1–9.
[76] B. Samanta, Gear fault detection using artificial neural networks and support vector machines with genetic algorithms, Mech. Syst. Sign. Process. 18 (3) (2004) 625–644.
[77] N. Saravanan, V.K. Siddabattuni, K. Ramachandran, Fault diagnosis of spur bevel gear box using artificial neural network (ANN), and proximal support vector machine (PSVM), Appl. Soft Comput. 10 (1) (2010) 344–360.
[78] P. Kankar, S.C. Sharma, S. Harsha, Fault diagnosis of ball bearings using machine learning methods, Expert Syst. Appl. 38 (3) (2011) 1876–1886.
[79] X. Zhu, Y. Zhang, Y. Zhu, Intelligent fault diagnosis of rolling bearing based on kernel neighborhood rough sets and statistical features, J. Mech. Sci. Technol. 26 (9) (2012) 2649–2657.
[80] Y.H. Hu, J.-N. Hwang, Handbook of Neural Network Signal Processing, CRC Press, 2001.
[81] J. Rafiee, F. Arvani, A. Harifi, M. Sadeghi, Intelligent condition monitoring of a gearbox using artificial neural network, Mech. Syst. Sign. Process. 21 (4) (2007) 1746–1754.
[82] M. Mrugalski, M. Witczak, J. Korbicz, Confidence estimation of the multi-layer perceptron and its application in fault detection systems, Eng. Appl. Artif. Intell. 21 (6) (2008) 895–906.
[83] A. Sadeghian, Z. Ye, B. Wu, Online detection of broken rotor bars in induction motors by wavelet packet decomposition and artificial neural networks, IEEE Trans. Instrum. Meas. 58 (7) (2009) 2253–2263.
[84] J. Sanz, R. Perera, C. Huerta, Gear dynamics monitoring using discrete wavelet transformation and multi-layer perceptron neural networks, Appl. Soft Comput. 12 (9) (2012) 2867–2878.
[85] G. Bin, J. Gao, X. Li, B. Dhillon, Early fault diagnosis of rotating machinery based on wavelet packetsłempirical mode decomposition feature extraction and neural network, Mech. Syst. Sign. Process. 27 (2012) 696–711.
[86] T.H. Patel, A.K. Darpe, Vibration response of a cracked rotor in presence of rotor–stator rub, J. Sound Vib. 317 (3) (2008) 841–865.
[87] N.H. Chandra, A. Sekhar, Fault detection in rotor bearing systems using time frequency techniques, Mech. Syst. Sign. Process. 72 (2016) 105–133.
[88] P. Jayaswal, S. Verma, A. Wadhwani, Application of ann, fuzzy logic and wavelet transform in machine fault diagnosis using vibration signal analysis, J. Qual. Maint. Eng. 16 (2) (2010) 190–213.
[89] M.R. Mehrjou, N. Mariun, M.H. Marhaban, N. Misron, Rotor fault condition monitoring techniques for squirrel-cage induction machine? A review, Mech. Syst. Sign. Process. 25 (8) (2011) 2827–2848.
[90] P.K. Kankar, S.C. Sharma, S.P. Harsha, Vibration-based fault diagnosis of a rotor bearing system using artificial neural network and support vector machine, Int. J. Modell. Ident. Contr. 15 (3) (2012) 185–198.
[91] M.R. Meireles, P.E. Almeida, M.G. Simões, A comprehensive review for industrial applicability of artificial neural networks, IEEE Trans. Ind. Electron. 50 (3) (2003) 585–601.
[92] S. Wu, T.W. Chow, Induction machine fault detection using SOM-based RBF neural networks, IEEE Trans. Ind. Electron. 51 (1) (2004) 183–194.
[93] Y. Lei, Z. He, Y. Zi, Application of an intelligent classification method to mechanical fault diagnosis, Expert Syst. Appl. 36 (6) (2009) 9941–9948.
[94] K. Zhang, Y. Li, P. Scarf, A. Ball, Feature selection for high-dimensional machinery fault diagnosis data using multiple models and radial basis function networks, Neurocomputing 74 (17) (2011) 2941–2952.
[95] Z. Li, X. Yan, C. Yuan, J. Zhao, Z. Peng, The fault diagnosis approach for gears using multidimensional features and intelligent classifier, Noise Vib. Worldwide 41 (10) (2010) 76–86.
[96] H. Wang, P. Chen, Intelligent diagnosis method for rolling element bearing faults using possibility theory and neural network, Comput. Ind. Eng. 60 (4) (2011) 511–518.
[97] H. Yaghobi, H.R. Mashhadi, K. Ansari, Artificial neural network approach for locating internal faults in salient-pole synchronous generator, Expert Syst. Appl. 38 (10) (2011) 13328–13341.
[98] F. Wang, S. Chen, J. Sun, D. Yan, L. Wang, L. Zhang, Time-frequency fault feature extraction for rolling bearing based on the tensor manifold method, Math. Probl. Eng. 2014 (2014) 15, https://doi.org/10.1155/2014/198362, Article ID 198362.
[99] B. Samanta, K.R. Al-Balushi, S.A. Al-Araimi, Artificial neural networks and genetic algorithm for bearing fault detection, Soft Comput. 10 (3) (2006) 264–271.
[100] D. Pham, P. Pham, Artificial intelligence in engineering, Int. J. Mach. Tools Manuf. 39 (6) (1999) 937–949.
[101] M.D. Prieto, G. Cirrincione, A.G. Espinosa, J.A. Ortega, H. Henao, Bearing fault detection by a novel condition-monitoring scheme based on statistical-time features and neural networks, IEEE Trans. Ind. Electron. 60 (8) (2013) 3398–3407.
[102] S. Hong, Z. Zhou, E. Zio, W. Wang, An adaptive method for health trend prediction of rotating bearings, Digit. Sign. Process. 35 (2014) 117–123.

[103] Y. Bengio, A. Courville, P. Vincent, Representation learning: a review and new perspectives, IEEE Trans. Pattern Anal. Mach. Intell. 35 (8) (2013) 1798–1828.
[104] K. Li, Q. Wang, Study on signal recognition and diagnosis for spacecraft based on deep learning method, in: 2015 Prognostics and System Health Management Conference (PHM), IEEE, 2015, pp. 1–5.
[105] F. Jia, Y. Lei, J. Lin, X. Zhou, N. Lu, Deep neural networks: a promising tool for fault characteristic mining and intelligent diagnosis of rotating machinery with massive data, Mech. Syst. Sign. Process. 72 (2016) 303–315.
[106] H. Shao, H. Jiang, X. Zhang, M. Niu, Rolling bearing fault diagnosis using an optimization deep belief network, Meas. Sci. Technol. 26 (11) (2015) 115002.
[107] O. Fink, E. Zio, U. Weidmann, Fuzzy classification with restricted Boltzman machines and echo-state networks for predicting potential railway door system failures, IEEE Trans. Reliab. 64 (3) (2015) 861–868.
[108] C. Li, R.-V. Sanchez, G. Zurita, M. Cerrada, D. Cabrera, R.E. Vásquez, Multimodal deep support vector classification with homologous features and its application to gearbox fault diagnosis, Neurocomputing 168 (2015) 119–127.
[109] J. Liu, E. Zio, Feature vector regression with efficient hyperparameters tuning and geometric interpretation, Neurocomputing 218 (2016) 411–422.
[110] S.B. Kotsiantis, I.D. Zaharakis, P.E. Pintelas, Machine learning: a review of classification and combining techniques, Artif. Intell. Rev. 26 (3) (2006) 159–190.