

An intelligent bearing fault diagnosis system: A review

S.R.Sauffi^{1,*}, Z.A.B Ahmad¹, M.S Leong¹ and M.H Lim¹

¹Institute of Noise and Vibration, Universiti Teknologi Malaysia, 54100 Kuala Lumpur, Malaysia

Abstract. Rolling element bearing (REB) is a well-known component that most extensively used in the industry. They operate in extreme condition (high temperature, dirty environment) which may lead to unexpected failure after the certain operation. Faulty on bearing cause severe equipment damage, financial loss and threaten people's life. Development of proper fault diagnosis system of REB capable of preventing unexpected failure from occurs and maintain the machine work in the healthy state. Over a few decades, machine learning is introduced to provide a consistent fault diagnosis result. Hence, this paper reviewed the development of bearing diagnosis method using machine learning models.

1 Introduction

Machine learning model is a sub-category of artificial intelligence. It has the ability to automatically learn and generalize the data. The google trend of machine learning model in science and technology application are shown in Figure 1. In past 8 years, the use of artificial neural network is quite extensive compared to other machine learning model. However, in 2016, the deep learning model start to increase drastically and outperformed the ANN model. Thus, this study tends to summarise the previous of the usage of machine learning model in rolling element bearing fault diagnosis.

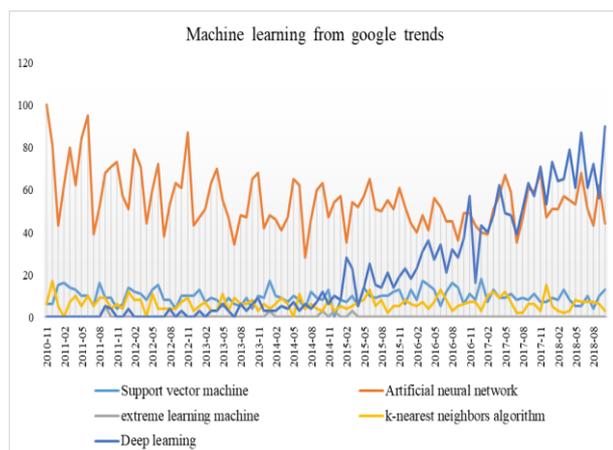


Fig 1. Google trend analysis on the use of machine learning in technology and science application

Rolling element bearing (REB) is broadly used in domestic and industrial applications. The main purpose of the bearing is to enable linear and rotational movement in order to reduce friction and handling load. Working condition of these appliances depends on the smooth and quiet running of the bearings. Moreover, bearing extensively utilised in practical equipment, such as induction machines [1] wind turbines [2], helicopters [3], and trains [4] which consider as a critical component in a majority of machines where bearing component is exposure to extreme conditions (dirty conditions, high

temperature, overstress) continuously. Dirty environment contaminates bearing's lubrication lead to the formation of the defect. Lubrication plays an important role to ensure bearing operate properly. In industry, lubrication problems contribute to 80 percent of bearing failure [5]. Further, the bearing fault is the common cause of rotating machinery failure [6] where bearing problems is the main contributions to wind turbine gearbox failure as illustrated in figure 2. Statistics show that 90 percent of the rotating machinery failure due to the bearing fault and 30 percent of the bearing failure occur on inner and outer race [7]. According to ISO 15242:2004(E), there is six common failure mode on rolling element bearing such as fatigue, corrosion, wear, plastic deformation, fracture and electrical erosion. The failure of bearing in the industry may lead to the long-term breakdown of the machine and high-cost of maintenance. Over the last two decades, the development of machine learning is quite extensive in machinery diagnosis. Machine learning is introduced in this field to assist human to make a robust and consistent decision diagnostic result.

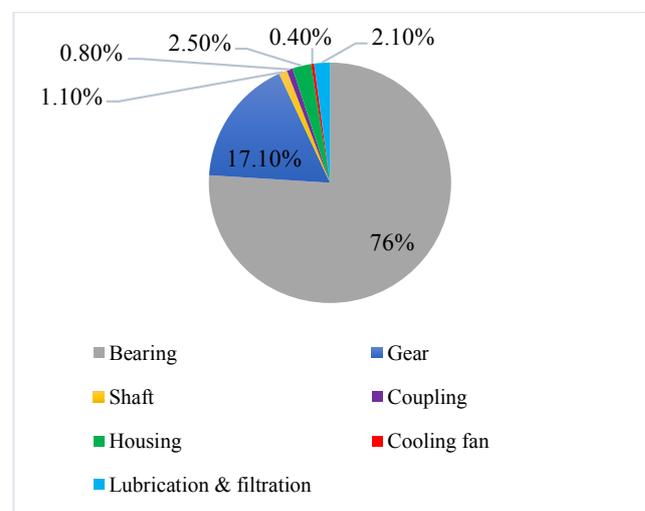


Fig 2. Major causes of wind turbine gearbox failure. (Source: Office of Energy efficiency and Renewable energy; September 2015) [8]

* Corresponding author: msramadhan93@gmail.com

A lot of research has been published related to bearing fault diagnosis by various type of methods. Bearing diagnosis works have been extensively reviewed by some researchers. One of the earliest reviews on rolling element bearing fault diagnosis using vibration analysis and acoustic emission have been done by Tandon and Nakra [9]. Nandi et al. reviewed on electrical motor components condition monitoring via motor current signature analysis (MCSA) [10]. On the other hand, Jardine et al. did a comprehensive review on rotating machinery diagnostic and prognostics by employing condition-based maintenance [11]. A general review of bearing condition monitoring in an electric motor is summarised by Zhou et al. [12]. Sidar et al. reviewed on the vibration analysis methods for condition monitoring of rolling element bearings [13]. Salah et al. discussed on acoustic emission techniques on rotating machinery monitoring with the implementation of artificial intelligence [14]. A brief review of condition monitoring on wind turbine bearing is presented by Azevedo et al. [15]. Furthermore, Cubillo et al. did a detailed review of condition monitoring with the application of bearing and gear. The objective of this paper is to update the past review paper by including recent works and the advanced techniques implemented in bearing defect detection [16]. This paper is structured in four sections.

2 Bearing fault diagnosis based on machine learning models

Over the last few decades, the massive development of machine learning on bearing fault diagnosis has been done. Figure 3 shows the machine learning employed in bearing fault diagnosis applications.

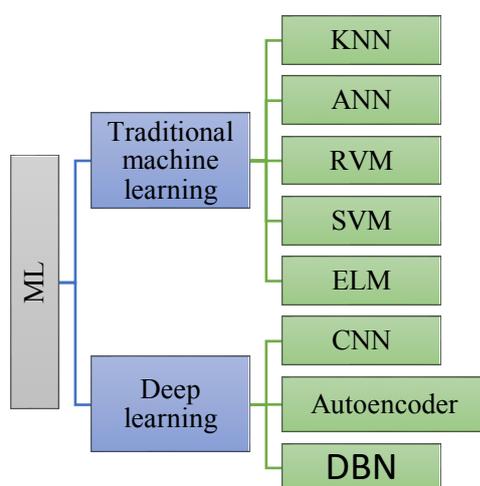


Fig 3. Machine learning models on bearing fault diagnosis

2.1 K-Nearest Neighbour

K-nearest-neighbor (K-NN) classifier is one of the most basic and simple classification techniques where the algorithm classifying objects based on the shortest distance of training examples in the feature space [17]. KNN applicable to those whose do not have any idea

about the data distribution. The advantages of this method are no assumptions is needed about the characteristics of the concepts and complex concepts can be learned by local approximation using simple procedures. The optimal value of K can be determined by resampling method such as example bootstrap or cross validation. Another important factor that affects the accuracy of the algorithm is a distance of the metric which depending on the data set. The most broadly used distance metric are Euclidean, Chebyshev, Manhattan, Minkowski, Cityblock and hamming [17], [18]. But, the Euclidean distance between a sample of test data and the samples of training data is commonly used in KNN [19]. However, KNN is incapable for the large dataset due to computationally expensive since the performance depends on the number of dimensions. Yigit proposed Artificial Bee Colony is used to find the optimal parameter of KNN [20]. KNN have been utilised in bearing applications. For example, Pandya et al. diagnosed rolling element bearing with Hilbert-Huang transform of acoustic emission data using APF-KNN [21]. Safizadeh and Latifi fused data from multi-sensor for vibration fault diagnosis of bearings by the accelerometer and load cell where the data is classified using KNN [22]. Baraldi et al. utilised K-Nearest Neighbour for bearing diagnostics under varying conditions by integrating KNN with Binary differential evolution (BDE) algorithm [23]. The proposed scheme has the satisfactory performance to diagnose bearing under variable conditions.

2.2 Artificial Neural Network

ANN has shown its capability in few decades in classify the health condition of the rotating machinery. By mimicking the human brain principle, the ANN forecast the target value with abundant of data by learning the pattern of historical that comprises an input layer, a hidden layer and output layers. Information moves through different layers from the input to the output where each layer is connected to neuron to the next layer. The neural network works when a particular input is inserted into the network then a unique output will produce [24]. The estimated output is compared to the desired output using loss function. The network is tuned so that the estimated output match the desired output. This method is trained via Back-propagation (BP) algorithm where training data set and the output is compared with the desired output until the system produce minimum mean squared error (MSE) as a stopping criterion. There are several types of a back-propagation algorithm that can be used during ANN training phases such as Levenberg-Marquardt, gradient descent, gradient descent with adaptive learning rate, scaled conjugate gradient and BFGS quasi-Newton. The back-propagation process contains two stages which are feed-forward and back-propagation. The initial stage is a feed- forward stage where the network is assigned randomly that will result in an error value [25]. Then the error value will be taken to adjust the weights in the back- propagation stage. The process is repeated until the error reduce. Due to the capability of ANN in past few decades, this method has been used in bearing fault diagnosis. For example, Yu et.

al diagnosed bearing conditions (healthy, outer race fault & inner race fault) using energy feature based on EMD method as input for ANN [26]. The result produced by ANN is good enough to classify the type of bearing failure. Further, Ben Ali et al. utilised the ANN method to diagnose bearing with the use of energy entropy from the IMFs signal which two additional parameters is proposed [27]. Based on the comparison with the previous research, this author able to examine bearing degradation without any human intervention. In the same year, the same author proposed automated nearly online bearing fault diagnosis to classify vibration signals' parameter using a combination of PNN and Simplified Fuzzy Adaptive Resonance Theory Map (SFAM). The application of ANN method able to recognise the different type of bearing defect [28]. Rao et al. utilised feed forward four layers ANN to analysed bearing defect size by using load, revolution per minutes (RPM) and RMS velocity as input to ANN [29]. Authors reported that ANN gives better defect size prediction with 3.75 percent of mean percentage error. Some authors argue with the successful implementation of ANN since the system depends so much on the proper selection of the network structure and need a large amount of training data. In reality, to obtain the abundant of real case data is not always available. Therefore, it is not suitable to apply ANN in real application otherwise applicable for experiment study.

2.3 Support Vector Machine

Last few years, the limited application of ANN on bearing diagnosis is found where a suitable classifier to mitigate the limitation of ANN is developed by Vapnik called Support vector machine (SVM) [30]. The main purpose is to separate the data by a maximum wide gap for different possible class by determining the separators (hyperplane) on search space [31] and the data are first mapped into a higher dimensional feature space through feature mapping function $\phi(x)$. SVM reduces the occurrence of risk during the training phase since it follows the Structural Risk Minimization (SRM) principle which has good performance in term of dealing with small data size. Kernel functions used in SVM mainly consists of four types of such as linear, polynomial, RBF and sigmoid. RBF by far is the most kernel used for classification in SVM. Jiang et al. used a combination of time domain feature and support vector machine (SVM) to classify the features from the multisensory where the selected feature able to identify defect from gear, bearing and rotor [32]. SVM method depends on the number vote to classify the data into each class. However, if the data contains an equivalent number of vote, SVM unable to made any decision but it will choose the class that come first in the arrangement of the classes. Hui et al. proposed an automated fault diagnosis on bearing components method by using the combination method of SVM-Dempster Shafer (DS) to classify the healthy and faulty data more accurate compared to conventional SVM [33]. The authors succeed to increase the accuracy from 76% to 94% and demonstrated the effectiveness of DS evidence theory by eliminating the conflict decision. Zeng et al. utilised one-class classification based on SVM for bearing

problems and compare with one-class classification based on the convex hull [34].

2.4 Relevance Vector Machine

Relevance Vector Machine (RVM) is a supervise learning that contains sparse kernel model that represents a Bayesian inference for probabilistic classification and regression with prior that result in the sparse in a sparse representation [31]. The RVM contains identical functional form as support vector machine but it provides probabilistic interpretation for its output [35]. RVM is based on a hierarchical prior, where in the first level, the weighted parameter is an independent Gaussian prior while in the second level an independent Gamma hyperprior is used for the variance parameters. Relevance vector machine does not suffer from the limitation of SVM where kernel function of SVM must satisfy the Mercer's conditions [36]. In addition, the RVM model provides a probabilistic output that contains less relevance vector for a given dataset and does not need to define regularisation parameter (C) [37]. Maio et al. predicted the residual life of bearing components using the combination of RVM and exponential regression which then the proposed scheme outperforms another model-based method [38]. In addition, Tran et al. utilised thermal imaging techniques to monitor the bearing state condition using RVM and generalised discriminant analysis [39]. Fei forecasted the kurtosis of bearing vibration signal using a combination of RVM and artificial bee colony (ABC) algorithm. ABC is employed to select the kernel parameter of RVM [40].

2.5 Extreme Learning Machine

Extreme learning machine (ELM) also has been introduced in rotating machinery diagnosis. ELM is the improvement of SVM method which proposed to increase the adaptivity of SVM which aims to focus on regression applications. ELM algorithm provides an easy of implementation, quick learning rate and less human intervention where the user does not need to select proper kernel function where random kernel ELM is introduced that tend to provide smallest training error and also the smallest norm of output weights [41]. In addition, its maximises the distance of the separating margins of the two different classes in the ELM feature space by minimising the norm of output weights while the hidden neurons (contains almost nonlinear activation functions) can be randomly generated independently of training samples. The hidden layer of ELM does not require tuning step [42]. There are several improvements of ELM as illustrated in table 1. Lendasse et al. did an extensive review on ELM where more information about ELM can refer to this works [41]. In addition, ELM also utilised for bearing state diagnosis. Tian et al. diagnosed various types of bearing fault under variable conditions by adopting ELM [43]. Also, Liu et al. effectively estimate the bearing useful life by using ELM [44]. Luo et al. did a research on 56 working conditions of bearing by utilising the combination of ELM and Gravitational search algorithms [45].

2.6 Deep learning

Maintenance decision making requires a machine learning algorithm to acquire the behaviour of the machine based on the feature extraction during data processing. ANN and SVM are broadly used in this area due to its performance. Recently, machine learning has moved toward deep machine learning. The difference of machine learning model and deep learning model is shown in Figure 4. The deep learning capable of reducing three processes of traditional fault diagnosis. One of the promises of deep learning is the capability to replace classical feature engineering with efficient algorithms for unsupervised or semi-supervised feature learning and hierarchical feature extraction where it can adaptively learn useful parameters from the raw data through multiple nonlinear transformations [46]. Deep learning actually utilise functions of higher complexity in order to deal with more complex problems [47]. Yu and Deng described two additional key of deep learning [48]. Firstly, the model is described as generative nature which naturally requires an extra top layer to perform the discriminative task. Secondly, deep learning is useful for big amounts of unlabeled training data for mining the structures and consistencies in the input features since its provide an unsupervised pretraining step.

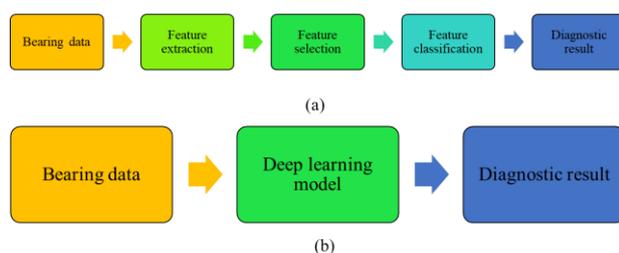


Fig 4. a) Traditional fault diagnosis and b) advance fault diagnosis

In general, deep learning is invented from neural network which recently is utilized in many applications image recognition, speech recognition and etc. Several deep learning model can be used in bearing fault diagnosis such as restricted Boltzmann machine (RBM) [50], Convolutional neural network (CNN) [51], [52], deep autoencoder [53] and deep belief network (DBN) [54]. DNN able to adaptively analyse raw data information through multiple layers of non-linear transformations and approximate complex nonlinear function with small error. autoencoder are trained by two main procedures. Firstly, each of autoencoder layers is pre-train with unsupervised techniques, like autoencoders. Then, the autoencoder structure is fine-tuned with back-propagation (BP) algorithm to achieve high accuracy of prediction [55]. A trained autoencoder can produce the input's representation by reconstructing the original data. In contrast, ANN requires the feature input that is extracted and selected manually by the user from the signal that acquires from a sensor which largely depending on prior knowledge about signal processing techniques and user expertise. Jia et al. applied autoencoder for bearing diagnosis to reveal fault characteristics of bearings under

variable operation conditions [55]. DBN is a generative graphical model which composed of several stacked RBMs. RBM comprises of two layers which are visible nodes layer and hidden nodes layer where each layer is denoted by a random binary vector and it is a bipartite undirected graphical model. He et al. did a research on bearing fault diagnosis using RBM [50]. The authors compare Gaussian RBM classifier with other classifiers, such as extreme learning machine, support vector machine, and deep belief network. The result indicates that the proposed scheme capable of diagnosing bearing state conditions accurately. Li et al. combined deep Boltzmann Machine (DBM) with random forest by fusing acoustic and vibration signal to diagnose gearbox condition [57].

In addition, CNN also received the amount of attention which has been proven successful in any types of domains. In details, the CNN structure contains multiple layers including convolutional layer, sub-sampling or pooling layer, LRN layer, fully-connected layer and the output layer [60]. Haedong et al. diagnosed rotating machinery fault based on orbit plot and the orbit plot images are used as input in CNN. The proposed method enable to classify fault mode via orbit plot [61]. Janssens et al. applied CNN for bearing fault diagnosis [62]. The results indicate the proposed scheme expressively out-performs the handcraft feature based approach which requires manual selection of feature. Recently, Guo et al. treated bearing problems using a deep convolutional neural network where the proposed scheme outperformed the traditional method like SVM [63].

3 Conclusion

In this paper, a machine learning model that was used on bearing fault diagnosis has been reviewed. Among the models, SVM achieved the highest usage in bearing fault diagnosis. However, the future trend shows that deep learning will receive a lot of attention by researcher due to its capability of providing automated feature extraction and feature selection.

References

- [1] A. Bouzida, O. Touhami, R. Ibtouen, A. Belouchrani, M. Fadel, and a. Rezzoug, "Fault diagnosis in industrial induction machines through discrete wavelet transform," *IEEE Trans. Ind. Electron.*, vol. 58, no. 9, pp. 4385–4395, 2011.
- [2] W. Y. Liu, B. P. Tang, J. G. Han, X. N. Lu, N. N. Hu, and Z. Z. He, "The structure healthy condition monitoring and fault diagnosis methods in wind turbines: A review," *Renew. Sustain. Energy Rev.*, vol. 44, pp. 466–472, 2015.
- [3] X. Qi, D. Theilliol, J. Qi, Y. Zhang, and J. Han, "A literature review on Fault Diagnosis methods for manned and unmanned helicopters," 2013

- Int. Conf. Unmanned Aircr. Syst. ICUAS 2013 - Conf. Proc.*, pp. 1114–1118, 2013.
- [4] M. Zhao, J. Lin, Y. Miao, and X. Xu, “Detection and recovery of fault impulses via improved harmonic product spectrum and its application in defect size estimation of train bearings,” *Meas. J. Int. Meas. Confed.*, vol. 91, pp. 421–439, 2016.
- [5] C. Radu, “the Most Common Causes of Bearing Failure and the Importance of,” *Rev. Lit. Arts Am.*, vol. 2010, no. February, p. 2010, 2010.
- [6] M. Elforjani and D. Mba, “Accelerated natural fault diagnosis in slow speed bearings with Acoustic Emission,” *Eng. Fract. Mech.*, vol. 77, no. 1, pp. 112–127, 2010.
- [7] L. Shuang, “Bearing Fault Diagnosis Based on PCA and SVM,” *2007 Int. Conf. Mechatronics Autom.*, no. 1, pp. 3503–3507, 2007.
- [8] “Statistics Show Bearing Problems Cause the Majority of Wind Turbine Gearbox Failures,” 2015. [Online]. Available: <https://energy.gov/eere/wind/articles/statistics-show-bearing-problems-cause-majority-wind-turbine-gearbox-failures>.
- [9] N. Tandon and A. Choudhury, “Review of vibration and acoustic measurement methods for the detection of defects in rolling element bearings,” *Tribol. Int.*, vol. 32, no. 8, pp. 469–480, 1999.
- [10] S. Nandi, H. A. Toliyat, S. Member, X. Li, and S. Member, “Condition Monitoring and Fault Diagnosis of Electrical Motors — A Review,” vol. 20, no. 4, pp. 719–729, 2005.
- [11] A. K. S. Jardine, D. Lin, and D. Banjevic, “A review on machinery diagnostics and prognostics implementing condition-based maintenance,” *Mech. Syst. Signal Process.*, vol. 20, no. 7, pp. 1483–1510, 2006.
- [12] W. Zhou, T. G. Habetler, and R. G. Harley, “Bearing Condition Monitoring Methods for Electric Machines: A General Review,” *2007 IEEE Int. Symp. Diagnostics Electr. Mach. Power Electron. Drives*, pp. 3–6, 2007.
- [13] R. Sidar, P. K. Sen, and G. Sahu, “Review of Vibration Based Fault Diagnosis in Rolling Element Bearing and Vibration Analysis Techniques,” vol. 4, no. 10, pp. 998–1003, 2015.
- [14] Y. H. Ali and S. M. Ali, “Acoustic Emission and Artificial Intelligent Methods in Condition Monitoring of Rotating Machine – A Review,” pp. 212–219, 2016.
- [15] H. D. M. De Azevedo, A. M. Araújo, and N. Bouchonneau, “A review of wind turbine bearing condition monitoring: State of the art and challenges,” *Renew. Sustain. Energy Rev.*, vol. 56, pp. 368–379, 2016.
- [16] A. Cubillo, S. Perinpanayagam, and M. Esperon-Miguez, “A review of physics-based models in prognostics: Application to gears and bearings of rotating machinery,” *Adv. Mech. Eng.*, vol. 8, no. 8, pp. 1–21, 2016.
- [17] S. B. Imandoust and M. Bolandraftar, “Application of K-Nearest Neighbor (KNN) Approach for Predicting Economic Events : Theoretical Background,” *Int. J. Eng. Res. Appl.*, vol. 3, no. 5, pp. 605–610, 2013.
- [18] H. S. Khamis, K. W. Cheruiyot, and S. Kimani, “Application of k- Nearest Neighbour Classification in Medical Data Mining,” vol. 4, no. 4, pp. 121–128, 2014.
- [19] J. Tian, C. Morillo, M. H. Azarian, and M. Pecht, “Kurtosis-Based Feature Extraction Coupled With K -Nearest Neighbor Distance Analysis,” vol. 63, no. 3, pp. 1793–1803, 2016.
- [20] H. Yigit, “ABC-based distance-weighted k NN algorithm,” *J. Exp. Theor. Artif. Intell.*, vol. 3079, no. June, pp. 1–10, 2014.
- [21] D. H. Pandya, S. H. Upadhyay, and S. P. Harsha, “Fault diagnosis of rolling element bearing with intrinsic mode function of acoustic emission data using APF-KNN,” *Expert Syst. Appl.*, vol. 40, no. 10, pp. 4137–4145, 2013.
- [22] M. S. Safizadeh and S. K. Latifi, “Using multi-sensor data fusion for vibration fault diagnosis of rolling element bearings by accelerometer and load cell,” *Inf. Fusion*, vol. 18, no. 1, pp. 1–8, 2014.
- [23] P. Baraldi, F. Cannarile, F. Di Maio, and E. Zio, “Hierarchical k-nearest neighbours classification and binary differential evolution for fault diagnostics of automotive bearings operating under variable conditions,” *Eng. Appl. Artif. Intell.*, vol. 56, pp. 1–13, 2016.
- [24] H. Sadegh, A. N. Mehdi, and A. Mehdi, “Classification of acoustic emission signals generated from journal bearing at different lubrication conditions based on wavelet analysis in combination with artificial neural network and genetic algorithm,” *Tribol. Int.*, vol. 95, pp. 426–434, 2016.
- [25] Y. H. Ali, R. Abd Rahman, and R. I. R. Hamzah, “Artificial neural network model for monitoring oil film regime in spur gear based on acoustic emission data,” *Shock Vib.*, vol. 2015, 2015.
- [26] Y. Yu, YuDejie, and C. Junsheng, “A roller bearing fault diagnosis method based on EMD energy entropy and ANN,” *J. Sound Vib.*, vol. 294, no. 1–2, pp. 269–277, 2006.
- [27] J. Ben Ali, N. Fnaiech, L. Saidi, B. Chebel-Morello, and F. Fnaiech, “Application of empirical mode decomposition and artificial neural network for automatic bearing fault diagnosis based on vibration signals,” *Appl. Acoust.*, vol. 89, pp. 16–27, 2015.
- [28] J. Ben Ali, L. Saidi, A. Mouelhi, B. Chebel-Morello, and F. Fnaiech, “Linear feature selection and classification using PNN and SFAM neural networks for a nearly online diagnosis of bearing naturally progressing degradations,” *Eng. Appl. Artif. Intell.*, vol. 42, pp. 67–81, 2015.
- [29] C. Vana, Vital Rao; Ratnamb, “Estimation of Defect Severity in Rolling Element Bearings using Vibration Signals with Artificial Neural

- Network,” *Jordan J. Mech. Ind. Eng.*, vol. 9, no. 2, pp. 113–120, 2015.
- [30] C. Cortes and V. Vapnik, “Support-Vector Networks,” vol. 297, pp. 273–297, 1995.
- [31] M. Rafi and M. S. Shaikh, “A comparison of SVM and RVM for Document Classification,” *Procedia Comput. Sci.*, vol. 00, pp. 3–8, 2013.
- [32] L. L. Jiang, H. K. Yin, X. J. Li, and S. W. Tang, “Fault diagnosis of rotating machinery based on multisensor information fusion using SVM and time-domain features,” *Shock Vib.*, vol. 2014, 2014.
- [33] K. H. Hui, M. H. Lim, M. S. Leong, and S. M. Al-Obaidi, “Dempster-Shafer evidence theory for multi-bearing faults diagnosis,” *Eng. Appl. Artif. Intell.*, vol. 57, no. October 2016, pp. 160–170, 2017.
- [34] M. Zeng, Y. Yang, S. Luo, and J. Cheng, “One-class classification based on the convex hull for bearing fault detection,” *Mech. Syst. Signal ...*, vol. 81, pp. 274–293, 2016.
- [35] M. Tipping, “Sparse Bayesian Learning and the Relevance Vector Machine,” *J. Mach. Learn. Res.*, vol. 1, pp. 211–244, 2001.
- [36] N. Ceryan, “Application of support vector machines and relevance vector machines in predicting uniaxial compressive strength of volcanic rocks,” *J. African Earth Sci.*, vol. 100, pp. 634–644, 2014.
- [37] M. Pal, “Kernel Methods in Remote Sensing: a Review,” *ISH J. Hydraul. Eng.*, vol. 15, no. sup1, pp. 194–215, 2009.
- [38] F. Di Maio, K. L. Tsui, and E. Zio, “Combining Relevance Vector Machines and exponential regression for bearing residual life estimation,” *Mech. Syst. Signal Process.*, vol. 31, pp. 405–427, 2012.
- [39] V. T. Tran, B. S. Yang, F. Gu, and A. Ball, “Thermal image enhancement using bi-dimensional empirical mode decomposition in combination with relevance vector machine for rotating machinery fault diagnosis,” *Mech. Syst. Signal Process.*, vol. 38, no. 2, pp. 601–614, 2013.
- [40] S. Fei, “Kurtosis forecasting of bearing vibration signal based on the hybrid model of empirical mode decomposition and RVM with artificial bee colony algorithm,” *Expert Syst. Appl.*, vol. 42, no. 11, pp. 5011–5018, 2015.
- [41] A. Lendasse, V. C. Man, Y. Miche, and G. Bin Huang, “Advances in extreme learning machines (ELM2014),” *Neurocomputing*, vol. 174, pp. 1–3, 2016.
- [42] G.-B. Huang, H. Zhou, X. Ding, and R. Zhang, “Extreme learning machine for regression and multiclass classification,” *IEEE Trans. Syst. Man, Cybern. Part B, Cybern.*, vol. 42, no. 2, pp. 513–29, 2012.
- [43] Y. Tian, J. Ma, C. Lu, and Z. Wang, “Rolling bearing fault diagnosis under variable conditions using LMD-SVD and extreme learning machine,” *Mech. Mach. Theory*, vol. 90, pp. 175–186, 2015.
- [44] Y. Liu, B. He, F. Liu, S. Lu, Y. Zhao, and J. Zhao, “Remaining useful life prediction of rolling bearings using PSR, JADE, and extreme learning machine,” *Math. Probl. Eng.*, vol. 2016, 2016.
- [45] M. Luo, C. Li, X. Zhang, R. Li, and X. An, “Compound feature selection and parameter optimization of ELM for fault diagnosis of rolling element bearings,” *ISA Trans.*, vol. 65, pp. 556–566, 2016.
- [46] Y. LeCun, Y. Bengio, and G. Hinton, “Deep learning,” *Nature*, vol. 521, no. 7553, pp. 436–444, May 2015.
- [47] M. Bianchini and F. Scarselli, “On the complexity of neural network classifiers: A comparison between shallow and deep architectures,” *IEEE Trans. Neural Networks Learn. Syst.*, vol. 25, no. 8, pp. 1553–1565, 2014.
- [48] D. Yu and L. Deng, “Deep Learning and Its Applications to Signal and Information Processing [Exploratory DSP],” *Signal Process. Mag. IEEE*, vol. 28, no. 1, pp. 145–154, 2011.
- [49] L. Deng, G. Hinton, and B. Kingsbury, “New types of deep neural network learning for speech recognition and related applications: an overview,” *2013 IEEE Int. Conf. Acoust. Speech Signal Process.*, pp. 8599–8603, 2013.
- [50] X. He, D. Wang, Y. Li, and C. Zhou, “A Novel Bearing Fault Diagnosis Method Based on Gaussian Restricted Boltzmann Machine,” vol. 2016, 2016.
- [51] L. Jing, T. Wang, M. Zhao, and P. Wang, “An Adaptive Multi-Sensor Data Fusion Method Based on Deep Convolutional Neural Networks for Fault Diagnosis of Planetary Gearbox,” *Sensors*, vol. 17, no. 3, p. 414, 2017.
- [52] H. Shao, H. Jiang, F. Wang, and H. Zhao, “An enhancement deep feature fusion method for rotating machinery fault diagnosis,” *Knowledge-Based Syst.*, vol. 119, pp. 200–220, 2016.
- [53] C. Lu, Z. Y. Wang, W. L. Qin, and J. Ma, “Fault diagnosis of rotary machinery components using a stacked denoising autoencoder-based health state identification,” *Signal Processing*, vol. 130, pp. 377–388, 2017.
- [54] J. Yin and W. Zhao, “Fault diagnosis network design for vehicle on-board equipments of high-speed railway: A deep learning approach,” *Eng. Appl. Artif. Intell.*, vol. 56, no. August, pp. 250–259, 2016.
- [55] F. Jia, Y. Lei, J. Lin, X. Zhou, and N. Lu, “Deep neural networks: A promising tool for fault characteristic mining and intelligent diagnosis of rotating machinery with massive data,” *Mech. Syst. Signal Process.*, vol. 72–73, no. November, pp. 303–315, 2016.
- [56] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, “Learning representations by back-propagating errors.”
- [57] C. Li, R. V. Sanchez, G. Zurita, M. Cerrada, D.

- Cabrera, and R. E. Vásquez, “Gearbox fault diagnosis based on deep random forest fusion of acoustic and vibratory signals,” *Mech. Syst. Signal Process.*, vol. 76–77, pp. 283–293, 2016.
- [58] A. This, “Synchronization for Uncertain Neural Networks With Mixed Time Delays,” vol. 40, no. 1, pp. 173–185, 2010.
- [59] L. Shao, L. Liu, and X. Li, “Feature Learning for Image Classification Via Multiobjective Genetic Programming,” *Neural Networks Learn. Syst. IEEE Trans.*, vol. 25, no. 7, pp. 1359–1371, 2014.
- [60] J. Sun, Z. Xiao, and Y. Xie, “Automatic multi-fault recognition in TFDS based on convolutional neural network,” *Neurocomputing*, vol. 222, no. August 2016, pp. 127–136, 2017.
- [61] H. Jeong, S. Park, S. Woo, and S. Lee, “Rotating Machinery Diagnostics Using Deep Learning on Orbit Plot Images,” *Procedia Manuf.*, vol. 5, pp. 1107–1118, 2016.
- [62] O. Janssens *et al.*, “Convolutional Neural Network Based Fault Detection for Rotating Machinery,” *J. Sound Vib.*, vol. 377, pp. 331–345, 2016.
- [63] X. Guo, L. Chen, and C. Shen, “Hierarchical adaptive deep convolution neural network and its application to bearing fault diagnosis,” *Meas. J. Int. Meas. Confed.*, vol. 93, pp. 490–502, 2016.