INVESTIGACIÓN & DESARROLLO, No. **16**, Vol. **1**: 102 – 114 (2016)
ISSN 2518-4431

A REVIEW OF VIBRATION MACHINE DIAGNOSTICS BY USING ARTIFICIAL INTELLIGENCE METHODS

UNA REVISIÓN BIBLIOGRÁFICA DEL ANÁLISIS VIBRACIONAL PARA EL DIAGNÓSTICO DE MÁQUINAS MEDIANTE EL USO DE MÉTODOS DE INTELIGENCIA ARTIFICIAL

Grover Zurita [a)], **Vinicio Sánchez** [b)] **y Diego Cabrera** [b)]
[a)] *Centro de Investigaciones Ópticas y Energías* (CIOE), *Universidad Privada Boliviana, Bolivia*
[b)] *Department of Mechanical Engineering, Salesanien Polytechnic University. Cuenca-Ecuador*
grzurita@upb.edu
(Recibido el 12 abril 2016, aceptado para publicación el 11 de junio 2016)

ABSTRACT

In the industry, gears and rolling bearings failures are one of the foremost causes of breakdown in rotating machines, reducing availability time of the production and resulting in costly systems downtime. Therefore, there are growing demands for vibration condition based monitoring of gears and bearings, and any method in order to improve the effectiveness, reliability, and accuracy of the bearing faults diagnosis ought to be evaluated. In order to perform machine diagnosis efficiently, researchers have extensively investigated different advanced digital signal processing techniques and artificial intelligence methods to accurately extract fault characteristics from vibration signals. The main goal of this article is to present the state-of-the-art development in vibration analysis for machine diagnosis based on artificial intelligence methods.

RESUMEN

En la industria, fallas en los engranajes y rodamientos son una de las principales causas de avería en máquinas rotativas, reduciendo el tiempo de disponibilidad de los equipos en producción y generando tiempo de inactividad con costos elevados. Por tanto, existe una creciente demanda de monitoreo basado en la condición de vibración de engranajes y rodamientos, y cualquier método con el fin de mejorar la fiabilidad y exactitud del diagnóstico de fallas debe ser evaluado. Con el fin de realizar eficientemente el diagnóstico de las máquinas, investigadores desarrollan diferentes técnicas avanzadas de procesamiento digital de señales y métodos de inteligencia artificial para extraer características de las fallas en los equipos. El objetivo principal de este estudio bibliográfico es de realizar una revisión del estado del arte relacionado con el análisis de vibraciones para el diagnóstico de las máquinas usando métodos de inteligencia artificial.

Keywords: Artificial Intelligence Method, Machine Learning Method, Random Forest, Deep Learning.
Palabras Clave: Método de Inteligencia Artificial, Método de Aprendizaje de Máquina, Arboles de Decisión, Aprendizaje Profundo.

1. INTRODUCTION

To effectively and accurately diagnose rotating machines faults, researchers have extensively investigated different advanced digital signal processing techniques and artificial intelligence methods to obtain fault characteristics from vibration data signatures. Nowadays, the diagnosis techniques have been refined and data-acquisition systems have become increasingly user-friendly to make a major impact in the fault vibration diagnosis field.

This literature survey attempts to summarize and review the recent research and development of Digital Signal Processing (DSP) and Artificial Learning Methods (Machine Learning Method) in vibration fault and diagnosis of rotating machinery. It can be pointed out that the research review methodology was based on the bibliography study of Elsevier and Scopus editorials and it was also limited to papers, mostly, published in the last 10 years (from 2006 to 2016). We reviewed the state-of-the-art of all the latest research and development in this field. It aims to synthesize available information on this topic in context and provide comprehensive references to enhance some clues about this exiting vibration machine diagnosis field.

This paper essentially will focus on the application of DSP and Intelligence Learning methods to the most important components in rotating machines (Bearings and Gears). Both components play important role in the industrial machinery. Generally, when the components break down, they are subjected to the influences of different types of non-linear dynamics forces, leading to vibration signatures with a content of both deterministic and random components. Therefore, it is important to develop a suitable diagnosis method, which can handle industrial equipment with complex vibration signatures.

In order to effectively perform machine diagnosis, researchers have continuously investigated different advanced signal processing techniques such as Fast Kurtogram [1], [2], [3], [4], Fast Fourier Transform (FFT) [9], Cesptrum[5], Envelope analysis[6], [7], [8], [9] and Wavelets analysis [10], [11], [13] to accurately extract fault vibration signals characteristics from rotating machines. By applying above fault detection machine diagnosis methodologies, there are important studies to detect gear and rolling bearing faults. The first method used in machine diagnosis, maybe not the most efficient, it was the FFT however this technique is quite limited and it works efficiently only for periodic signals. In the recent years the Cepstrum analysis method has appeared, which is mainly based on its ability to detect echoes, periodicity and side bands in the spectrum [12]. The side bands could be directly related to the machine faults. Other interesting method is the Envelope Analysis or Amplitude Demodulation with Hilbert Transform; it works similar to a band pass filter. This method works by first selecting the adequate band pass filter range, so as it eliminates high amplitude signals not associated with faults, and enhance the interesting peaks related to the machine faults [13], [14]. It was used the Wavelets Packed Transform (WPT) to compress or de-noise signal data. It provides accurate data information on the energy localization content in time and frequency domain [8]. Moreover, the fast Spectrogram analysis method developed by Jerome [4], put forward the signal processing technique for fault detection. It is a powerful tool for detecting the presence of transients in a signal. An extensive tutorial of rolling bearing element diagnosis was presented by Randal *et al*. in 2011, [15]. The fault diagnosis using wavelet Envelope power spectrum of rotating machines obtained successfully results [8]. Ming *et al.*,2012, [16], applied two methods, the Spectral Kurtosis (SK) and Autoregressive (AR) models for fault diagnosis and condition monitoring of rolling bearing. The AR models are incorporated into SK as a data pre-process, performing a pre whiting of the signal to reduce some anomalies in the SK analysis.

The latest years, several analysis techniques for gears faults diagnosis have used WPT, to enhance the vibration data, which is provided by the classical statistical parameters from the vibration signal, in time and frequency domain [11], , [17], [18]. These above approaches have been very useful for implementing the Condition Based Maintenance (CBM), as presented in Jardine *et al*.[19].

It can be highlighted the latest vibration fault detection based development for DSP techniques presented by Daubechies, *et al*., [20]. It was developed the Synchrosqueezed Wavelet Transform (SWT) based on empirical mode decomposition, which was a breakthrough in this research field. Moreover, the above method is in time –frequency representation. It is a powerful method for detection of transient signals. With other words blurry signals can easily reduce the noise and enhance the signal peaks which usually are related to the faults. Li et al.,2014, [21], developed further the SWT, not only to detect machine faults, to detect also the machine speed rotation only using an accelerometer vibration signal. This means that a tachometer sensor is no longer necessary to use it.

Several reviews related to machine diagnosis based on machine learning methods have been published, the most common approaches are:

- Artificial Neural Network [22], [23], [24]
- Random Forest [25], [26], [27], [28], [29], [30]
- Support Vector Machine [31], [32], [33], [34], [35], [36], [37]
- Principal Component Analysis, and Deep Learning [38], [23], [26]

The artificial neural network and ANFIS multistage decision algorithm for detection and diagnosis of bearing faults was presented by Ertunc *et al*. 2013,[22] . An automated diagnosis of rolling bearings using neural networks was presented by Castejón *et al*., 2010, [39]. It was used the decision tree and PCA based fault diagnosis to detect a rotating machinery by Sue *et al*.,2007,[40] . There are other valuables papers regarding fault diagnosis in rotating machines. The combination of two methods ANN and Support Vector Machine (SVM) with genetic algorithm for bearing fault detection was presented by Samanta, 2006, [41]. Random Forest (RF), as a regression and classification technique, has been used for fault diagnosis in several areas of engineering. RF is a robust approach in case of having a large number of input attributes and low number of available samples for learning. In [42], a Genetic Algorithm was applied to select the best features of the samples to use the RF method and increase the classification rate value.

To the diagnosis part the PCA method was also applied [43], which is a statistical method and can reveal relationships and correlations between large numbers of variables to obtain classification models. The main advantage of this method is that it can handle correlated data and can provide valuable insights into the inter-relationships between the variables.

In order to give an overview of the content of this paper, Figure 1 illustrates the main core of this research work, such as, the artificial intelligence methods (AIM) for diagnosis of the rotating machines. Due to importance of noise reduction,

before it is processed in the AIMs the recorded vibration data with non-stationary and random signals. It was also further highlighted the part of vibration data acquisition, data pre-processing and fault detection techniques. Figure 1 denotes also, based on the bibliography, the latest approaches for fault and diagnosis in rotating machines. It denotes signal processing techniques for fault detection and machine learning methods to be used for machine diagnosis. This process is divided into four stages: 1) The part of data collection and surveys 2) Fault detection and diagnosis systems and pre-processing of data, 3) Classification of the data and diagnosis, 4) Validation and failure predictions.

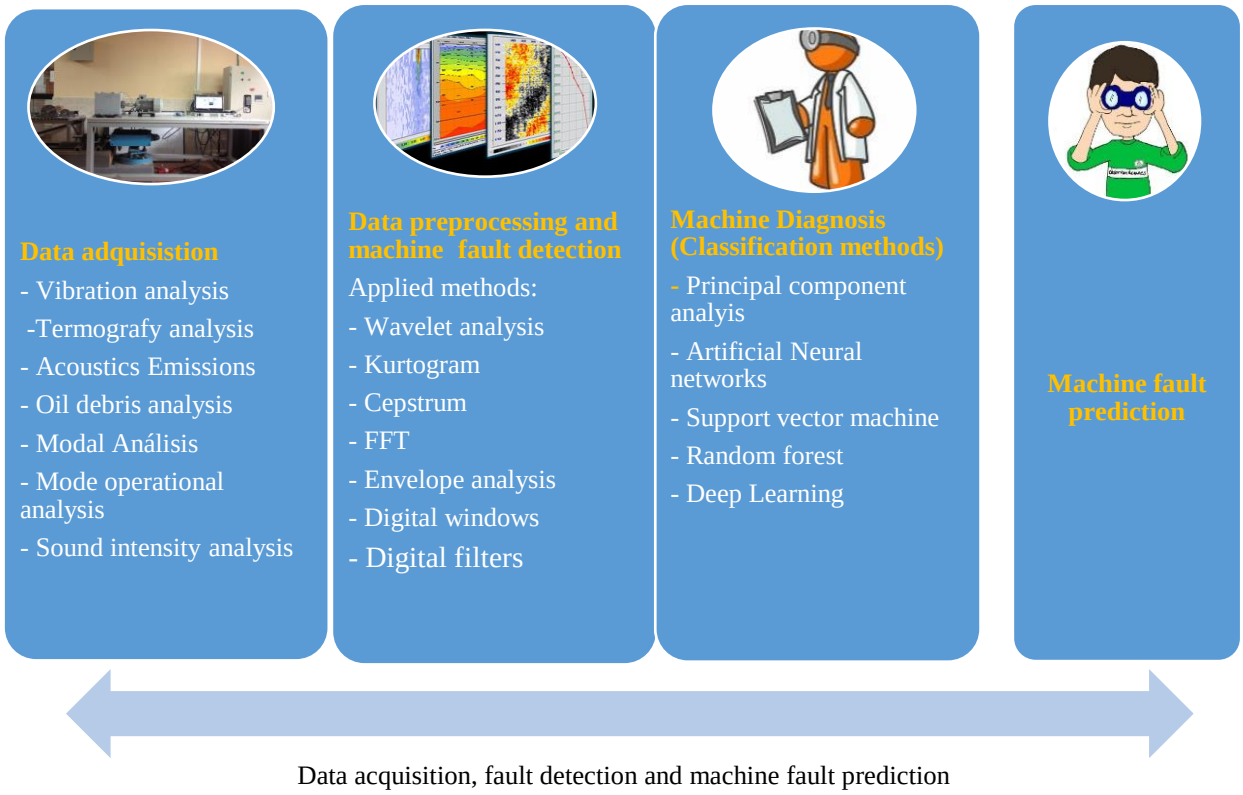

Figure 1- The flowchart of the machine diagnostics. a) Data acquisition , b) Fault detection , c) Fault classification , d) Fault prognosis.

Finally, the rest of this paper is organized as follows. Section two introduces the methods: WPT, Fast Kurtogram, ANN, SVM, RF, PCA and DL. Section three is related to the implementation of the above methods. Section four is related to discussion and summary of the paper.

2. BACKGROUND

This section carries out the literature surveys of vibration based fault detection with the most valuable methods: (a) Kurtogram, WPT Cepstrum and Enveloped analysis, b) The diagnosis techniques are carried out by machine learning methods. Due to the well proved accuracy and efficiency to handle complex vibration data, the following classification methods are recommended: RF, SVM; PCA ANN, and Deep Learning**.**

2.1 The Fast Kurtogram Method

Jerome [4] fully developed the Fast Kurtogram method based on the spectral kurtosis. The spectral kurtosis is an adequate tool for detecting the presence of transients in a signal, by specifying in which frequency bands these take place, as given in equation (1). The spectrogram analysis is based on FIR filters or STFT. As it can be seen in Figure 2, from [44], the signal are chop up along the record signal in overlapping steps with several slices by using designed FIR filters. Thereafter the spectra for each signal slices are arranged in a 3D plot with amplitude, time and frequency. Finally, it is applied to obtain a 2D diagram.

$$K(f) = \frac{\langle H^4(t,f)\rangle}{\langle H^2(t,f)\rangle^2} - 2 \tag{1}$$

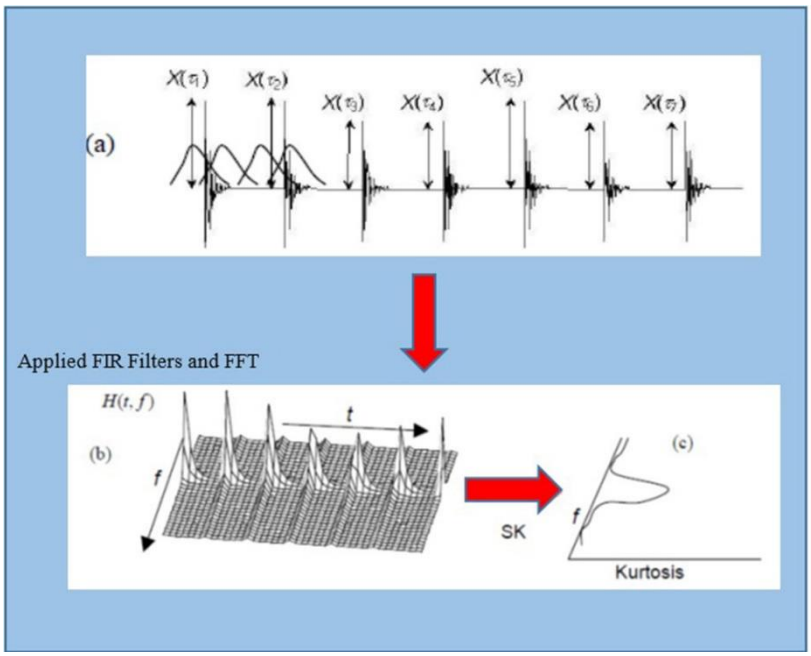

Figure 2 - Spectrum data analysis. a) Time trace signal, b) The FIR filters bank, c) The outcome of the Spectral Kurtosis analysis.

2.2 Wavelet Packed Transform (WPT)

WPT can compress or de-noise and provides accurate information on the energy localization content in time and frequency. WPT are a particular type of discrete wavelet transforms that allows one to assess the detailed information of signals in low and high frequency bands. Wavelets is a tool for signal analysis that has made great impact in various fields of engineering [10], [31],[45], [46]. The WPT, function is expressed by W(x), see Equation (2) and (3). The expression $h(k)$ and $g(k)$ are the high pass and low pass filters, respectively.

$$W_n(x) = \sqrt{2}\sum_{k=0}^{2N-1} h(k) W_n(2x-k), \quad \text{n=0,1,2} \tag{2}$$

$$W_{n+1}(x) = \sqrt{2}\sum_{k=0}^{2N-1} g(k) W_n(2x-k) \tag{3}$$

where $W_1(x) = \Psi(x)$ is the wavelet function and $W_n(x) = \varphi(x)$ is the scaling function, respectively. The scale parameter is j (j=0... J), which is the decomposition of number of levels, and the translation parameter is k (k=0,…,2^j-1).

$$W_{j,n,k}(x) = 2^{-\frac{j}{2}} W_n(2^{-j}x - k) \tag{4}$$

where $n \epsilon N$ and $(j,k)\ \epsilon Z^2$. Nowadays, the WPT has been used to explore and obtained energy features as input data to machine learning methods.

2.3 Cepstrum Method

The Complex Cepstrum is a non-linear signal processing technique with a variety of applications in areas such as reciprocating machine diagnostics, speech, and image processing. This method has proved to be effective and useful for these applications. In recent years, a great interest has focused on cepstral applications of machine diagnosis in reciprocating machines [5]. It is calculated by finding the complex logarithm of the Fourier transform of the signal in time domain x(t), then the inverse Fourier transform of the resulting sequence:

$$\hat{Y}(\tau) = FFT^{-1}\{Log(FFT(x(t))\} \tag{5}$$

The applications of the Cepstrum analysis to machine diagnostics are based mainly on its ability to detect periodicity in the spectrum, e.g., family of harmonics and sidebands.

2.4 Envelope Analysis (Amplitude Demodulation)

Envelope Detection or Amplitude Demodulation is the FFT modulating signal frequency spectrum of the modulating signal. This method is efficient to detect faults especially when the signals contained near –periodic frequency burst and generates high frequencies. The signals are generated from reciprocating machines and in some cases also from the incipient fault characteristics of the bearing faults [6], [7], [8], [9].

2.5 Principal Component Analysis (PCA)

The PCA is one of the most widely used multivariate data driven statistical techniques for vibration fault classification in rotating machine diagnosis, however it has also been used in some extend for pattern recognition (data clustering) and engine combustion fault diagnosis. PCA can work very well with, high dimensional and highly linearly correlated data. The PCA is designed to detect independent phenomena, in correlated data sets, describe all systematic variability in the data, and thereby remove noise. Data is decomposed in a number of independent components; the so-called PC. The modeling in PCA is based on the following equation (6), where X represents the variables averages.

$$X = \overline{X} + TP' + E \tag{6}$$

The matrix product TP' models the study structure. The residual matrix E contains the noise. Figure 3 illustrates, that each observation can be represented as a point in a multidimensional space, where the axes, such a, speed, load, injection timing, and fuel quality are the engine parameters. A principal component (PC) is a straight line through the observation points in the multivariable space. Moreover, PC is extracted from the residuals matrix. This iterative extraction of new PCs continues until no more systematic variation remains in the residual matrix. The first PC explains the major variation of data, all PCs are orthogonal and every PC explains a maximum in the remaining matrix variance.

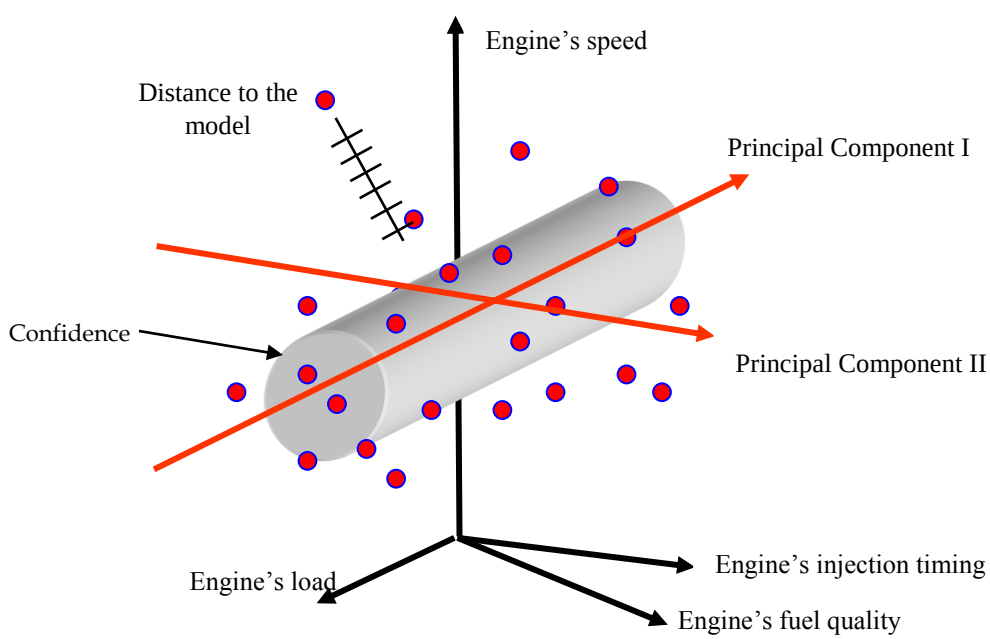

Figure 3 - Principals of the PCA.

2.6 Support Vector Machine (SVM)

SVM is a supervised learning model with associated learning algorithms that analyze data used for classification and regression analysis. In machine learning this method is also called maximum margin classifier [47]. Figure 4 illustrates the SVM procedure and it classifies data by finding the best hyperplane that separates all data points. The best hyperplane means the one with the largest margin between the two classes A and B. The support vector is the data points, which are closets to the hyperplane, these points are on the boundary are on the slab. Here only a brief review of the binary SVM classification algorithm is provided here. A thorough theoretical presentation of the method can be seen in [32], [33].

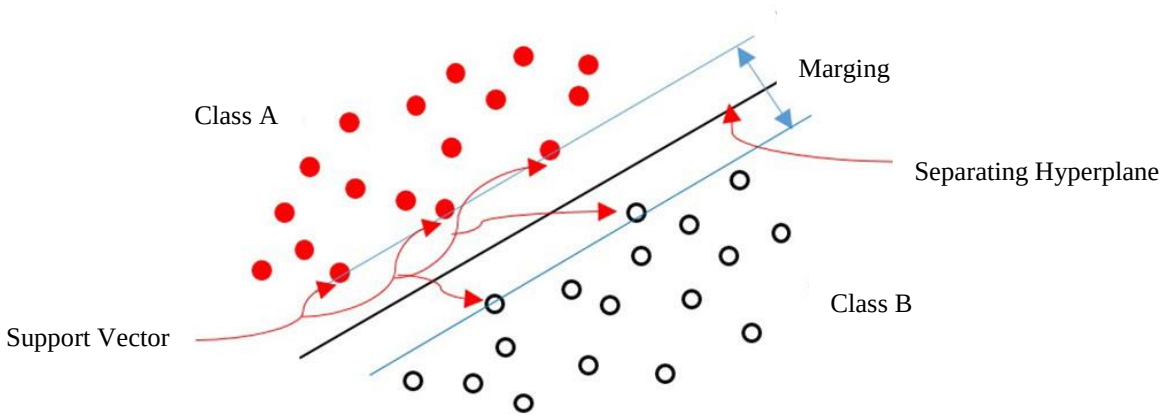

Figure 4 - Support Vector Machine (SVM) structure.

Let the data set **X** be such that:

$$\mathrm{X} = \{\{\mathrm{x1}, \mathrm{y1}\} \ldots \{\,\mathrm{xi}, \mathrm{yi}\} \epsilon R^{D} \mathrm{x}\{\pm 1\}\} \tag{7}$$

Given a kernel function $k_\theta(y_i, y_i)$ the decision function *f(x)* can be given by:

$$f(x) = \sum_{i=1}^{N} y_i \alpha_i k_\theta(y_i, y_i) + b \tag{8}$$

where b ∈R is the so-called bias term, α_i coefficients can be obtained by a quadratic optimization process. The kernel adopted is the Gaussian kernel given by:

$$K(x_i, y_j) = \exp\left(-\sum_i \frac{(x_i - x_j)^2}{2\sigma_i^2}\right) \tag{9}$$

Where σ_i^2 is the variance. In the case for multi type classification however, one needs to handle multiple classes. In such cases, the one-against-all (OAA) strategy can be used for multi-class classification. Given an input ***x***, the *i*th SVM produces the output $f_i(\boldsymbol{x})$, the final predicted class being selected by equation (10),[32],[33]:

$$\hat{i} = \arg\max_i f_i(\mathbf{x}) \tag{10}$$

2.7 Random Forests Algorithm

The Random Forest Algorithms are based on the decision trees classifier. It has been considered significantly in development of fault machine diagnostics methods through a powerful method for classification and predictions [48],[49]. One of the most widely used decision tree algorithm for classification and regression (CART) was developed by Breiman *et al*. [50]. Figure 5 gives a visual illustration of a decision tree and classification of two classes A and B. The decision trees creates a type of flowchart which consist of leafs and a set of decision to be made based off of branches, it works with a hierarchal similarity of a form of a tree. A decision tree performs by partitions of the features space in form of rectangles. Once these data subspaces have been found, decision trees can be seen as a collection set. The decision to stop or split again performs until some criteria is fulfilled.

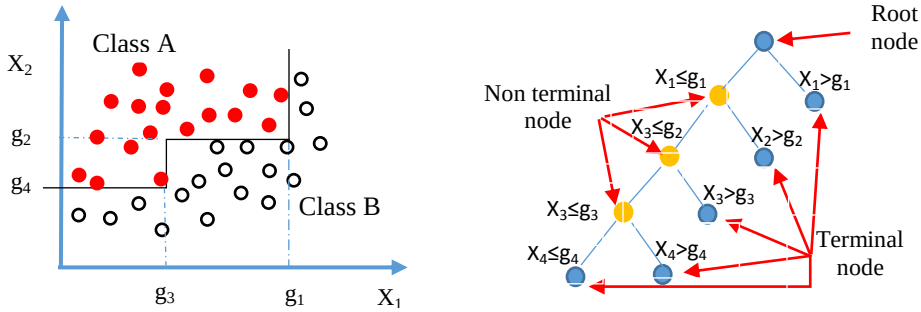

Figure 5 - Visualization of the Classification process with decision boundaries and corresponding classification tree.

Equation (11) shows the feature vector.

$$v = (x_1, x_2, x_3, \dots . x_d) \in R^d \tag{11}$$

Where x_i represents each element and d is the dimension of the feature vector. The decision trees seeks the best selection for each node j, a set of split parameter is selected resolving the optimization procedure, see equation 12:

$$\theta_j = \arg\max I\,(S_{j,}\,\theta) \qquad \theta_j \,\epsilon\, P \tag{12}$$

Where I is the fitness function S_j is the subset of training set belonging node j, P is the space of the parameter in binary tree-based model.

The RF models take the decision tree concept further by improving the classification models accuracy [25]. Figure 6 illustrates the classification procedure of RF analysis.

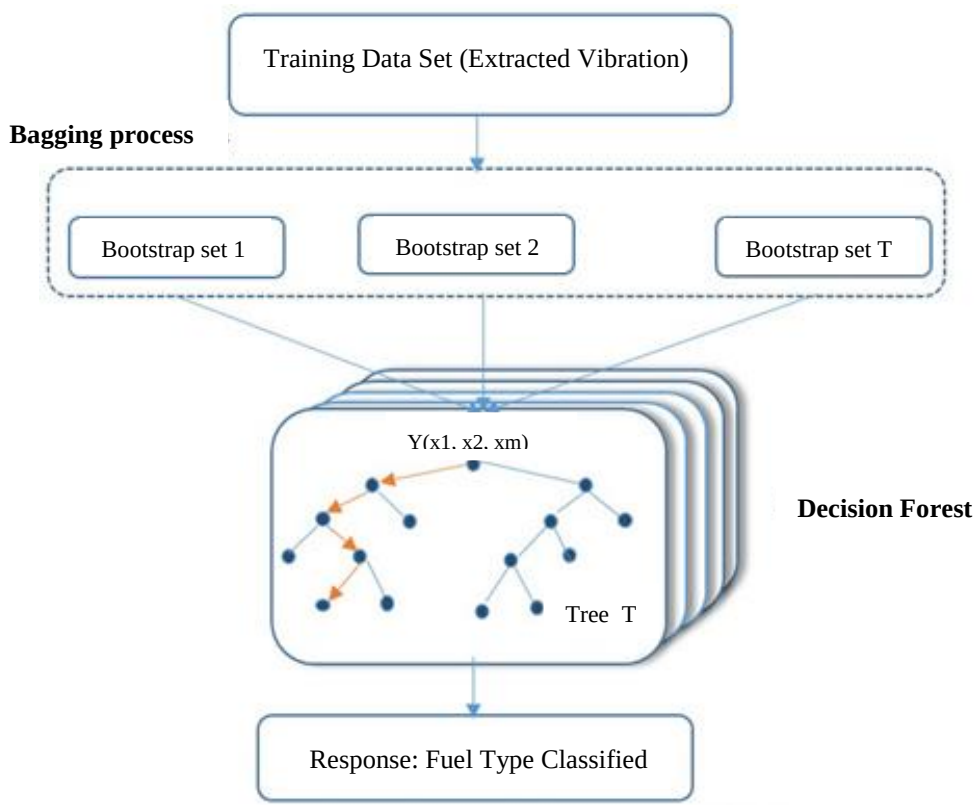

Figure 6 - The Random Forest Analysis Structure.

Breiman *et al*. [50] suggested bootstrap aggregating (bagging), to decrease the variance and reduce the risk of over fitting. It is carried out by a selection of the input variables and the random election with replacement of a data sample which is made to grow every tree in the forest. In order to obtain a variety of models that are not over fit to the available data, each component model is fit only to a bootstrap sample of the data [50]. A bootstrap sample is a sample of the same size as the original data set, but drawn with replacement. Therefore, each of those samples excludes some portion of the data, which is referred to as "out-of-bag" (OOB) data, which is a measure of the random forests prediction error [50].

2.8 Artificial Neural Networks (ANN)

The ANN is a classification method, which is well established and applied for fault detection and machine diagnosis in rotating machines[51], [52], [53], [22], [23], [24]. The design of ANNs was introduced by the structure of a real human brain, however, the processing elements and the architectures used in ANN have been developed far from their biological issue. A typical feed-forward neural network with a single hidden layer can be seen in Figure 7. The learning process may be automated by ANN, which can be configured for industrial applications, i.e., vibration data fault classification, reconstruction of the cylinder pressure using only vibration measurements, image compression, pattern recognition and miscellaneous application. Supervised back-propagation algorithms with feed- forward layer was frequently use in machine diagnosis, due to the robustness and efficiently to handle noisy data.

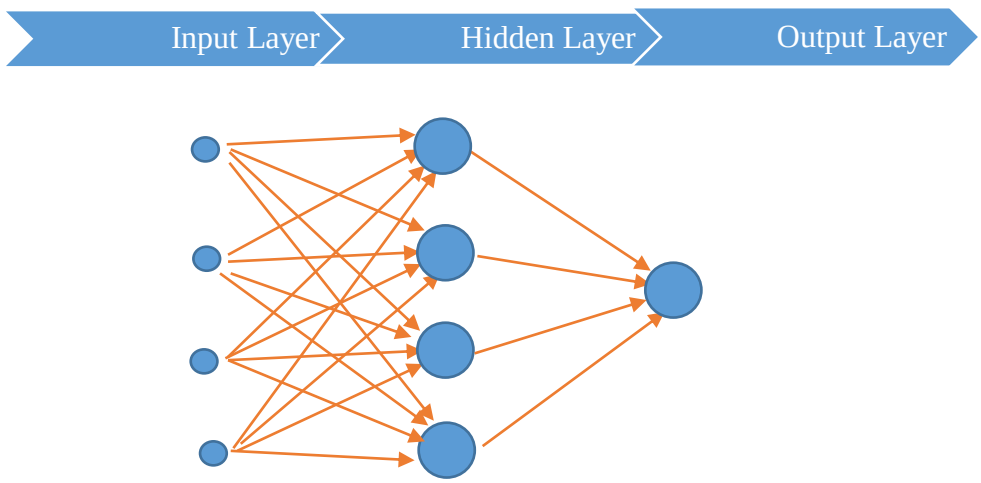

Figure 7 - Denotes a typical feed-forward neural network with a single hidden layer.

2.8 Deep Learning Method (DLM)

Deep learning Method is part of machine learning (ML) techniques and related to artificial neural networks which are composed of many layers [38],[23],[26]. It is based on a set of algorithms that attempt to model high-level abstractions in large data set, with complex structures composed of multiple non-linear transformations [38]. Cho *et al*.[54], proposed a Gaussian-Bernoulli Deep Boltzmann machine (GDBM) which was used the Gaussian neurons in the visible layer of the DBM. Equation 13 illustrates the initial equation for Deep learning analysis [55].

$$\begin{aligned} E_e(\mathbf{v},\mathbf{h}^{(1)},\ldots,\mathbf{h}^{(L)} \mid \theta) &= -\sum_{i=1}^{N_v}\sum_{j=1}^{N_1} W_{ij} v_i h_j^{(1)} / \sigma_i^2 \\ &+ \sum_{i=1}^{N_v} 2(v_i - b_i)^2 / \sigma_i^2 - \sum_{i=1}^{N_v}\sum_{j=1}^{N_l} b_j^{(l)} h_j^{(l)} \\ &- \sum_{i=1}^{L-1}\sum_{j=1}^{N_l}\sum_{k=1}^{N_l+1} w_{jk}^{(l)} h_j^{(l)} h_k^{(l+1)} \end{aligned} \tag{13}$$

where σ is the standard deviation of visible neurons, W_{ij} denotes the weight of the synaptic connection between the *i*th visible neuron and the *j*th hidden neuron, $\theta=\{\mathbf{W}, \mathbf{b}\}$ are the model parameters, , b_i represents the *i*th bias term, N_l stands for the number of neurons in the *l*th hidden layer and N_v is the number of visible neurons. The readers are refer to the bibliography to obtain an extensive description of above method [26] and [33].

3. APPLICATIONS OF THE STUDIED METHODOLOGIES

In order to effectively perform machine diagnose engine, researchers and engine´s developers have extensively investigated different advanced signal processing techniques and artificial intelligence methods to accurately extract fault vibration signals characteristics from rotating machines. In this section, is further developed the part of vibration data acquisition, data pre-processing and fault detection techniques, classification and prediction processes are highlighted. Figure 1 illustrates, based on the bibliography, the latest approaches for fault and diagnosis in rotating machines. It denotes signal processing techniques for fault detection and machine learning methods to be used for machine diagnosis. This process is divided into four stages: 1) The part of data collection and surveys 2) Fault detection and diagnosis systems and pre-processing of data, 3) Classification of the data and diagnosis, 4) Validation and failure predictions.

One main issue, which is not part of the focus of this paper, however due to the importance it could be work to mention. We are taking about the importance to perform high quality vibration measurements to obtain high quality vibration data, and consequently high classification rate values by using the machine diagnosis methods (RF, SVM, ANN and DL). Initially, the hardware (Accelerometers, Data Acquisition card) have to fulfill the required precision and accuracy, some equipment´s recommendation could be from National instruments, PCB, and Keisslers Brand.

Regarding , the digital signal processing field, there are two main phenomena that has to be take care off if you want to obtain high quality data. The first one is the alias phenomena, which is related to adequate Sampling Frequency (Fs). It has to fulfill the Nyquist criteria. The second one is the leakage phenomena, which the energy content of the signal can leakage to other frequencies. It has to be reduced to minimum. Figure 8 illustrates the flowchart of the presented review procedure to perform the vibration data acquisition, and fault detection performing by several methods.

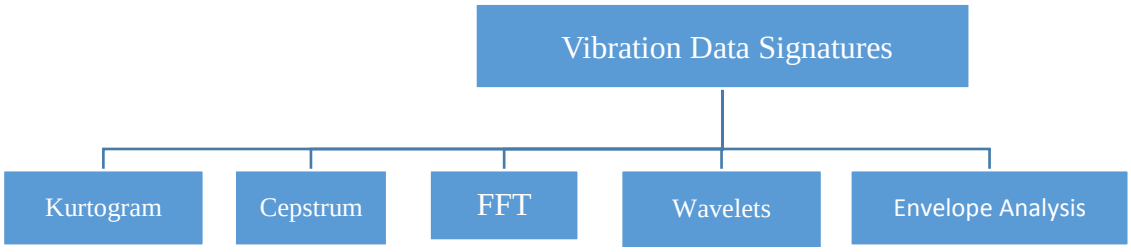

Figure 8 - Different methods for Vibration based fault detection.

3.1 Machine Fault Diagnosis

Figure 9 illustrates all the stages to perform the machine fault diagnosis procedure. It starts with the data acquisition process to collect a wide set of condition parameters. The condition parameters for fault diagnosis extracted from vibration signals are mostly related to time and frequency domains. The features extraction is carried out by applying in time domain and frequency domain. It is extracted the statistical parameters of the signals (Mean, Kurtosis, Skewness and RMS). The parameters associated to the wavelet transform domain are also used. Thereafter, the obtained features can be used as input data to Genetic Algorithms or Principal Components Analysis to extract, hopefully, only the most representative (best) features. The GA and PCA have been used as the optimization process to select the optimum features. The idea is to have an efficient process for selecting the data set of the features, which can lead to a higher diagnostic performance, and it can also reduce the input data set.

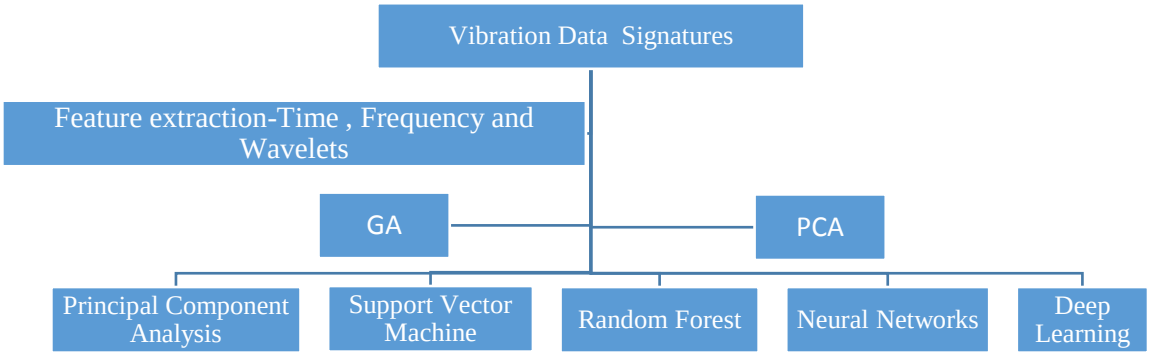

Figure 9 - Vibration based machine classification.

4. CONCLUSIONS

In this section, we compare vibration fault detection and diagnosis methods, highlighting their major characteristics. Due to the vibration signal behavior, it can be performed the selection of the signal processing analysis techniques: Periodic signals (FFT), Transient signals (Kurtogram), Periodicity, side bands and detection of echoes (Cepstrum analysis), Reduction of noise and energy localization content (Wavelets analysis), Impulse pulses at higher frequencies, Demodulation, (Envelope analysis). Table 1 illustrates a summary of the vibration fault detection and machine diagnosis methods

TABLE 1 – MAIN CHARACTERISTICS OF THE VIBRATION FAULT DETECTION METHODS

METHODS	Advantages	Disadvantages	References
KURTOGRAM	- It can detect and characterize transients signals	-It needs a precise band pass filters to be developed and incorporated into the method	[1],[2],[3]
CEPSTRUM	-Detect periodicity and side bands signals. - It can detect echo signal very well	-Working as a low pass filter could be that method averages-out the fluctuations of the curve of the spectrum.	[5]
FFT	-It works well for periodic signals	- It not suitable for non-periodic signals	
WAVELETS	-It can compress or de-noise and provides accurate information on the energy localization content in time and frequency. -Detailed information of signals in low and high frequency bands	-For fine analysis it could be computational intensive - It can take time to choose the right wavelet for a proper signal	[56], [8],[57], [12]
ENVELOPE ANALYSIS	-Efficient to detect faults, when the signals, contained near –periodic frequency burst, which generates high frequency content for repetitive impulse forces	-It is necessary to be improve the selection of band pass filters to perform the demodulation procedure	[6], [7], [8], [9]

Table 2 denotes the main characteristics of the classification methods. It is not straightforward to compare the characteristics of each classification methods. However, we will try to give some guidelines based on the studied bibliography.

TABLE 2 – MAIN CHARACTERISTICS OF THE VIBRATION FAULT VIBRATION CLASSIFICATION METHODS

METHODS	Advantages	Disadvantages	Referencies
PCA	- It can handle noisy, high dimensional and highly linearly correlated data.	- It could be difficult to be evaluated in an accurate manner the covariance matrix	[58], [43]
SVM	- Produce very accurate classifiers. - Less overfitting, robust to noise.	- Computationally expensive, thus runs slow.	[5], [32], [33], [34], [35], [36], [37]
RF	- RF is a robust approach in case of having a large number of input attributes and low number of available samples for learning	-It is high sensitivity to the selection of the input dataset with respect to the resulting tree structure. The tendency is the data could be overfitted (Solved by Bagging process)	[25], [26], [27], [28], [29], [49], [30]
ANN	-It requires less formal statistical training, -It can handle nonlinear relationships between dependent and independent variables.	- It is a "black box" - Greater computational burden, - It proneness to overfitting,	[60], [61], [62], [22], [23], [24]
DEEP LEARNING	-It is based on a set of algorithms that attempt to model high-level abstractions in large data set, with complex structures composed of multiple non-linear transformations	- It requires more input data than the other classification methods	[38], [23], [26]

Finally, based on the bibliography, we can state that the Deep Learning method outperforms the standard traditional classification methods, Li *et al*.[55]. It was carried out the classification procedure of different gears faults and compared them with other methods. However, further studies are required to perform a sensibility test to see how robust the method could be for industrial environment measurement settings.

5. BIBLIOGRAPHY

[1] D. Wang, P. W. Tse, and K. L. Tsui, "An enhanced Kurtogram method for fault diagnosis of rolling element bearings," *Mech. Syst. Signal Process.*, vol. 35, no. 1–2, pp. 176–199, Feb. 2013.

[2] Y. Lei, J. Lin, Z. He, and Y. Zi, "Application of an improved kurtogram method for fault diagnosis of rolling element bearings," *Mech. Syst. Signal Process.*, vol. 25, no. 5, pp. 1738–1749, Jul. 2011.

[3] Y. Ming, J. Chen, and G. Dong, "Spectral kurtosis based on AR model for fault diagnosis and condition monitoring of rolling bearing" Journal of Mechanicl Science and Technology. vol. 26, no. 2, 301-306. 2012

[4] A. Jerome: Fast Computing of the Kurtogram for the detection of transients faults. Mechanical system and signal processing.2007.

[5] C.-S. Park, Y.-C. Choi, and Y.-H. Kim, "Early fault detection in automotive ball bearings using the minimum variance cepstrum," *Mech. Syst. Signal Process.*, vol. 38, no. 2, pp. 534–548, Jul. 2013.

[6] A. B. Ming, W. Zhang, Z. Y. Qin, and F. L. Chu, "Envelope calculation of the multi-component signal and its application to the deterministic component cancellation in bearing fault diagnosis," *Mech. Syst. Signal Process.*, vol. 50–51, pp. 70–100, Jan. 2015.

[7] Z. Feng, X. Chen, and M. Liang, "Joint envelope and frequency order spectrum analysis based on iterative generalized demodulation for planetary gearbox fault diagnosis under nonstationary conditions," *Mech. Syst. Signal Process.*, vol. 76–77, pp. 242–264, Aug. 2016.

[8] M Lokesha and K. Ramachandran, "Fault diagnosis in gear using wavelet envelope power spectrum," *Int. J. Eng. Sci. Technol.*, vol. 3, no. 8, Aug. 2012.

[9] R. B. Randall, J. Antoni, and S. Chobsaard, "the relationship between spectral correlation and envelope analysis in the diagnostics of bearing faults and other cyclostationary machine signals," *Mech. Syst. Signal Process.*, vol. 15, no. 5, pp. 945–962, Sep. 2001.

[10] Y. Wang, G. Xu, L. Liang, and K. Jiang, "Detection of weak transient signals based on wavelet packet transform and manifold learning for rolling element bearing fault diagnosis," *Mech. Syst. Signal Process.*, vol. 54–55, pp. 259–276, Mar. 2015.
[11] W. Lu, W. Jiang, G. Yuan, and L. Yan, "A gearbox fault diagnosis scheme based on near-field acoustic holography and spatial distribution features of sound field," *J. Sound Vib.*, vol. 332, no. 10, pp. 2593–2610, May 2013.
[12] F. K. Omar and A. M. Gaouda, "Gear tooth diagnosis using wavelet multi-resolution analysis enhanced by Kaiser's windowing," *Trans. Inst. Meas. Control*, vol. 33, no. 5, pp. 573–590, Jul. 2011.
[13] F. Zuo: Gearbox fault detection using Hilbert and wavelet packet transform. Chemical Engineering Transaction2006
[14] Z. Su, Y. Zhang, M. Jia, F. Xu, and J. Hu, "Gear fault identification and classification of singular value decomposition based on Hilbert-Huang transform," *J. Mech. Sci. Technol.*, vol. 25, no. 2, pp. 267–272, Feb. 2011.
[15] R. B. Randall and J. Antoni, "Rolling element bearing diagnostics—A tutorial," *Mech. Syst. Signal Process.*, vol. 25, no. 2, pp. 485–520, Feb. 2011.
[16] A. B. Ming, Z. Y. Qin, W. Zhang, and F. L. Chu, "Spectrum auto-correlation analysis and its application to fault diagnosis of rolling element bearings," *Mech. Syst. Signal Process.*, vol. 41, no. 1–2, pp. 141–154, Dec. 2013.
[17] A. Aherwar, "An investigation on gearbox fault detection using vibration analysis techniques: A review," *Aust. J. Mech. Eng.*, vol. 10, no. 2, pp. 169–183, 2012.
[18] J. Yin, W. Wang, Z. Man, and S. Khoo, "Statistical modeling of gear vibration signals and its application to detecting and diagnosing gear faults," *Inf. Sci.*, vol. 259, pp. 295–303, Feb. 2014.
[19] K. S. Andrew, Jardine, Daming Lin, Dragan Banjevic, "A review on machinery diagnostics and prognostics implementing condition-based maintenance," *Mech. Syst. Signal Process.*, 2006.
[20] I. Debachuie: Synchrosqueezed wavelet transforms: An empirical mode decomposition-like tool," *Appl. Comput. Harmon. Anal.*, 2011.
[21] C. Li, V. Sanchez, G. Zurita, M. Cerrada Lozada, and D. Cabrera, "Rolling element bearing defect detection using the generalized synchrosqueezing transform guided by time–frequency ridge enhancement," *ISA Trans.*, vol. 60, pp. 274–284, Jan. 2016.
[22] H. Ertunc M., H. Ocak, and C. Aliustaoglu, "ANN- and ANFIS-based multi-staged decision algorithm for the detection and diagnosis of bearing faults," *Neural Comput. Appl.*, vol. 22, no. S1, pp. 435–446, May 2013.
[23] F. Jia, Y. Lei, J. Lin, X. Zhou, and N. Lu, "Deep neural networks: A promising tool for fault characteristic mining and intelligence diagnosis of rotating machinery with massive data," *Mech. Syst. Signal Process.*, vol. 72–73, pp. 303–315, May 2016.
[24] H., Zhou: Diagnosis method for gear equipment by sequential fuzzy neural network. *International Symposium on Neural Networks*. Springer Berlin Heidelberg, 2008.
[25] I. Witten, H. & Frank E: Data mining: Practical machine learning, tools and technique. Boston Morgan
[26] 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, Aug. 2016.
[27] A. Verikas, A. Gelzini , Bacauskiene: Mining data with random forest: A survey and results of new tests pattern recognition 2008. .
[28] J. Phillips, E. Cripps, J. W. Lau, and M. R. Hodkiewicz, "Classifying machinery condition using oil samples and binary logistic regression," *Mech. Syst. Signal Process.*, vol. 60–61, pp. 316–325, Aug. 2015.
[29] N. Horning, "Introduction to decision trees and random forests," *Am. Mus. Nat. Hist.*, 2013.
[30] Z. Jones and F. Linder, "Exploratory data analysis using random forests," in *Prepared for the 73rd annual MPSA conference*, 2015.
[31] W. Li, Z. Wang, Z. Zhu, G. Zhou, and G. Chen, "Design of Online Monitoring and Fault Diagnosis System for Belt Conveyors Based on Wavelet Packet Decomposition and Support Vector Machine," *Adv. Mech. Eng.*, vol. 5, no. 0, pp. 797183–797183, Jan. 2015.
[32] W. Xiao, W. Cheng, Y. Zi, C. Zhao, C. Sun, Z. Liu, J. Chen, and Z. He, "Support evidence statistics for operation reliability assessment using running state information and its application to rolling bearing," *Mech. Syst. Signal Process.*, vol. 60–61, pp. 344–357, Aug. 2015.
[33] C. Li, R.-V. Sanchez, G. Zurita, M. Cerrada, D. Cabrera, and R. E. Vásquez, "Multimodal deep support vector classification with homologous features and its application to gearbox fault diagnosis," *Neurocomputing*, vol. 168, pp. 119–127, Nov. 2015.
[34] J. Carmona: Tutorial sobre Máquinas de Soporte Vectorial 2014. Departamento Ingeniería Artificial: Universidad Nacional de Educación a Distancia Madrid (España).
[35] T. Xuan :Gear Fault Classification Using Genetic Programming and Support Vector Machines. International

Conference on Computer and Automation Engineering ICCAE 2010.
[36] J. Xuan, H. Jiang, T. Shi, and G. Liao, "Gear fault classification using genetic programming and support vector machines," *Int. J. Inf. Technol.*, vol. 11, no. 9, 2005.
[37] D. Su and W. Peng, "Online Gearbox Condition Monitoring Supported by Wireless Communication Techniques."
[38] M. Gan, C. Wang, and C. Zhu, "Construction of hierarchical diagnosis network based on deep learning and its application in the fault pattern recognition of rolling element bearings," *Mech. Syst. Signal Process.*, vol. 72–73, pp. 92–104, May 2016.
[39] C. Castejón, O. Lara, and J. C. García-Prada, "Automated diagnosis of rolling bearings using MRA and neural networks," *Mech. Syst. Signal Process.*, vol. 24, no. 1, pp. 289–299, Jan. 2010.
[40] I. Gravalos, Spyridon Loutridis,Dimitrios Moshou,Theodoros Gialamas,Dimitrios Kateris, Zisis Tsiropoulos, "Detection of fuel type on a spark ignition engine from engine vibration behaviour," *Appl. Therm. Eng.*, vol. Volume 54, no. Issue 14, 2013.
[41] K. Samanta, R. Al–Balushi, S. A. Al–Araimi,:Artificial neural networks and genetic algorithm for bearing fault detection," *Soft Comput.*, vol. 10, p., 2006.
[42] N. El Islem Karabadji , Ilyes Khelf, Hassina Seridi, Lakhdar Laouar, "Genetic Optimization of Decision Tree Choice for Fault Diagnosis in an Industrial Ventilator," *Chapter Cond. Monit. Mach. Non-Station. Oper.*
[43] W. Sue, J. Chen, and J. Li, "Decision tree and PCA-based fault diagnosis of rotating machinery," *Mech Syst Signal Process*, vol. 21.
[44] B. Sawalhi. Randall_2014_Gear parameter identification in a wind turbine gearbox using vibration signals.MSSP.2008. .
[45] D. Cabrera, F. Sancho, R.-V. Sánchez, G. Zurita, M. Cerrada, C. Li, and R. E. Vásquez, "Fault diagnosis of spur gearbox based on random forest and wavelet packet decomposition," *Front. Mech. Eng.*, vol. 10, no. 3, pp. 277–286, Sep. 2015.
[46] Y. Yang, X. Li, X. Liu, and X. Chen, "Wavelet kernel entropy component analysis with application to industrial process monitoring," *Neurocomputing*, vol. 147, pp. 395–402, Jan. 2015.
[47] C. Li, R.-V. Sanchez, G. Zurita, M. Cerrada, D. Cabrera, and R. E. Vásquez, "Multimodal deep support vector classification with homologous features and its application to gearbox fault diagnosis," *Neurocomputing*, vol. 168, pp. 119–127, Nov. 2015.
[48] Z. Jones and F. Linder, "Exploratory data analysis using random forests," in *Prepared for the 73rd annual MPSA conference*, 2015.
[49] N. Horning, "Introduction to decision trees and random forests," *Am. Mus. Nat. Hist.*, 2013.
[50] L. Breiman, "Random Forest," *Mach. Learn.*, vol. 45, 2001.
[51] C. Gershenson, "Artificial neural networks for beginners," *ArXiv Prepr. Cs0308031*, 2003.
[52] X. Zhou, H. Wang, P. Chen, and J. Song, "Diagnosis method for gear equipment by sequential fuzzy neural network," in *Advances in Neural Networks-ISNN 2008*, Springer, 2008, pp. 473–482.
[53] J. Schmidhuber, "Deep learning in neural networks: An overview," *Neural Netw.*, vol. 61, pp. 85–117, Jan. 2015.
[54] K. H. Cho, T. Raiko, and A. Ilin, "Gaussian-bernoulli deep boltzmann machine," in *Neural Networks (IJCNN), The 2013 International Joint Conference on*, 2013, pp. 1–7.
[55] C. Li, R.-V. Sanchez, G. Zurita, M. Cerrada, D. Cabrera, and R. E. Vásquez, "Multimodal deep support vector classification with homologous features and its application to gearbox fault diagnosis," *Neurocomputing*, vol. 168, pp. 119–127, Nov. 2015.
[56] B. Li, P. -l. Zhang, Q. Mao, S. -s. Mi, and P. -y. Liu, "Gear fault detection using adaptive morphological gradient lifting wavelet," *J. Vib. Control*, vol. 19, no. 11, pp. 1646–1657, Aug. 2013.
[57] Li,N.,Liu,N. Gear fault detection based on adaptive wavelet packet feature extraction. Mechanical engineering Science.2011.
[58] A. Bartkowiak and R. Zimroz, Sparse PCA for gearbox diagnostics paper 129, June 20, 2011.
[59] Y. Tian, M, Fu: Steel plates fault diagnosis on the basis of Support vector machine. Neurocomputing 2015.
[60] L. A. Wulandhari, A. Wibowo, and M. I. Desa. Hybrid artificial neural networks and genetic algorithms for bearings fault diagnosis. *Int J Appl Res Eng Sci*, vol. 1, no. 02, pp. 21–30.
[61] X. Zhou, H. Wang, P. Chen, and J. Song, "Diagnosis method for gear equipment by sequential fuzzy neural network," in *Advances in Neural Networks-ISNN 2008*, Springer, 2008, pp. 473–482.
[62] M. Barakat, M. El Badaoui, and F. Guillet, "Hard competitive growing neural network for the diagnosis of small bearing faults," *Mech. Syst. Signal Process.*, vol. 37, no. 1–2, pp. 276–292, May 2013.