Copyright © 2017 Tech Science Press SDHM, vol.12, no.1, pp.43-67, 2017

Brake Fault Diagnosis Through Machine Learning Approaches – A Review

Alamelu Manghai T.M.[1], Jegadeeshwaran R.[2] and Sugumaran V.[3]

Abstract Diagnosis is the recognition of the nature and cause of a certain phenomenon. It is generally used to determine cause and effect of a problem. Machine fault diagnosis is a field of finding faults arising in machines. To identify the most probable faults leading to failure, many methods are used for data collection, including vibration monitoring, thermal imaging, oil particle analysis, etc. Then these data are processed using methods like spectral analysis, wavelet analysis, wavelet transform, short-term Fourier transform, high-resolution spectral analysis, waveform analysis, etc., The results of this analysis are used in a root cause failure analysis in order to determine the original cause of the fault. This paper presents a brief review about one such application known as machine learning for the brake fault diagnosis problems.

Keywords: Vibration analysis, machine learning, feature extraction, feature selection, feature classification, Brake fault diagnosis.

1 Introduction

Brakes are one of the most important control components which bring the vehicle to rest within a reasonable distance even under the most adverse conditions. It is also desirable that the retardation should be smooth and the rate of retardation should be proportional to the pedal effort. This means that whilst the effort required by the driver to operate the

[1] School of Mechanical and Building Sciences, VIT University Chennai Campus, Chennai, Tamil Nadu, India – 600 127.
E-mail: alamelumangai.m@vit.ac.in; Phone: +91 44 39931335; Fax: +91 44 39932555
[2] School of Mechanical and Building Sciences, VIT University Chennai Campus, Chennai, Tamil Nadu, India – 600 127.
E-mail: krjegadeeshwaran@gmail.com; Phone: +91 44 39931335; Fax: +91 44 39932555
[3] School of Mechanical and Building Sciences, VIT University Chennai Campus, Chennai, Tamil Nadu, India – 600 127. Phone: +91 44 39931335; Fax: +91 44 39932555

brakes shall not be excessive. The brake system should be very reliable to promote the highest degree of safety on the road. Hence, an efficient brake system is responsible for the safety and stability of the vehicle. However, there are moving components involved; they are bound to get faulty due to various reasons, viz. wearing, air leak, fade, etc. When such things occur, the effectiveness of the brake reduces resulting in accidents. Hence, it is necessary that they are monitored all the time and diagnosed when faults occur. Monitoring of brakes is a separate area of concern in the contemporary automotive world. There are many analysis techniques, such as shock pulse method, wear debris analysis, acoustic emission, vibration analysis are available for the fault diagnosis problem. The review about one such analysis techniques is discussed in the following sections.

The shock pulse method is basically a vibration-monitoring technique. The impacts caused by damage in rolling element, such as bearings, brakes, gears, etc., generate shock pulses in the ultrasonic frequency band (Butler, 1973). The shock pulse method gives an indication of the condition of rotating components by measuring the magnitude of the mechanical impacts. These shock pulses can be measured using special accelerometers. However, it does not give much of diagnostic information. Wear debris analysis is the study of component wear particles in the lubricant to determine the condition of the machine parts (Hunt and Trevor, 1993). Excessive concentration of the wear particles in the lubricant signifies abnormal wear. It is relatively less costly; however, it is an off-line process and does not give much diagnostic information. The phenomenon of sound generation in materials under stress is termed as acoustic emission (AE). When a structure is subjected to an external stimulus, localized sources trigger the release of energy in the form of stress waves which propagate to the surface. Plastic deformation of cracks is the main sources of AE in metals. AE can detect the growth of subsurface cracks (Roberts and Talebzadeh, 2003). AE requires sophisticated signal processing systems. The whole system for fault diagnosis is more costly compared to systems based on vibration. The condition of moving components of a machine is assessed from the amount and nature of vibration, they generate. Deterioration in brake condition produces an increase in vibration level. Thus, an increase in the overall level of the vibration indicates a mechanical deterioration of one or more elements of the brake. Because the vibration waveform will contain a spectrum of frequencies associated with the various brake elements, knowledge of the frequencies within the spectrum at which a significant increase in vibration level has occurred can help to diagnose the causes of the deterioration (Luo, Osypiw, Irle, 2000). A comparative study suggests vibration signal as a suitable tool for the fault diagnosis problem compared to AE Signals (Al-Ghamd and David Mba, 2006).

2 Vibration analysis

Vibration analysis is one of the main methods in fault diagnosis. A review of the works in the area of vibration analysis for condition monitoring of brake components forms the theme of this section. The review of the literature made under the subsections of conventional methods, time domain analysis and frequency domain analysis and pattern recognition are described below.

2.1 Time domain analysis

Time domain analysis is the analysis of mathematical functions or physical signals with respect to time. A time-domain graph shows how a signal changes over time, whereas a frequency-domain graph shows how much of the signal lies within each given frequency band over a range of frequencies. There are many techniques available for the analysis. Some of the frequently used techniques have been discussed in this section.

2.1.1 Overall vibration level

Measuring overall vibration level over a broad band of frequencies is one of the most basic vibration techniques. The trend of overall vibration level is plotted against time and it acts as an indicator of deterioration. The overall level is often referred to as the signal RMS value. As peak is significantly affected by noise, RMS level is generally preferred in machine condition monitoring applications. The overall RMS level is a good indicator of machine condition in the case of simple machines, although it does not give much of diagnostic information. However, the same is not suited for complex machinery. In fact, localized faults here may go undetected until a significant secondary damage or catastrophic failure occurs (Colin Mercer, 2001).

2.1.2 Wave shape metrics

Faults which produce short-term impulses such as bearing faults may not significantly alter the overall vibration level; however, may cause a statistically significant change in the shape of the signal. With a number of fault types, the shape of the signal is a better indicator of damage than the overall vibration level.

Crest factor or kurtosis is often used as a non-dimensional measure of the shape of the signal. Both of these signal metrics increase in value as the ‘spikiness’ of the signal increases (i.e., as the signal changes from a regular continuous pattern to one containing isolated peaks). Kurtosis, being a purely statistical parameter, is usually preferable to crest factor in machine condition monitoring applications; for the same reasons, RMS is preferable to peak. However, the crest factor is in more widespread use because meters which record crest factor are more common and affordable than kurtosis meters. The wave shape metrics will not detect faults unless the amplitude of the vibration from the faulty component is large enough to cause a significant change in the total vibration level of the signal. This limits their use.

2.1.3 Time synchronous averaging

Time synchronous averaging is used to eliminate signal components that are not synchronous with the shaft rate of rotation. Eliminated components include electrical noise, bearing vibrations and vibrations related to other shafts or nearby machinery. The idea of TSA is to take ensemble average of the raw signal over a number of evolutions in an attempt to remove or reduce noise from other sources, so as to enhance the signal components of interest. This constitutes a pre-processing of the vibration signal and hence, can add to the reliability or authenticity of other techniques (Miller, 1999; Dalpiaz et al., 2000; Badaoui, 2001).

2.1.4 Descriptive Statistics

Descriptive statistics is the discipline of quantitatively describing the main features of a collection of information or the quantitative description itself. Descriptive statistics summarize a sample, rather than use the data to learn about the population that the sample of data is thought to represent. Some measures that are commonly used to describe a data set are measures of central tendency and measures of variability or dispersion. Measures of central tendency include the mean, median and mode, while measures of variability include the standard deviation (or variance), the minimum and maximum values of the variables, kurtosis, and skewness (Prem S. Mann, 1995).

2.1.5 Time series modeling

A time series is a sequence of data points, typically consisting of successive measurements made over a time interval. Time series analysis comprises methods for analyzing time series data in order to extract meaningful statistics and other characteristics of the data. The main idea of time series modeling is to fit the waveform data to a parametric time series model and extract features based on this parametric model. There are two popular mathematical models, namely, Auto-Regressive (AR) model and the Auto-Regressive Moving Average (ARMA) model. Upadhyaya (1980) compared the prediction error of an ARMA model of the observed multivariate signal in both faults free and fault situation as a decision base to detect sensor faults. Auto-regressive (AR) model is established by time difference and vibration amplitude. As the AR model utilized the mathematical method for fitting the variable, the AR coefficients represent the signal features and can be used to determine the fault types (Bailllie and Mathew, 1996). Dron, RasolofondraibeL, Couet, Pavan., (1998) developed a method for fault detection in ball bearings based on the estimation of an autoregressive model for the vibration signals and compared the first and second order statistical properties of the estimation error. Poyhonen, Jover, Hyotyniemi (2004) applied AR model to vibration signals collected from an induction motor and used the AR model coefficients as extracted features. In practice, however, application of the AR model or ARMA model is difficult due to the complexity in modeling, especially the need to determine the order of the model (Andrew et al, 2006). Among these techniques, descriptive statistics are successfully used to extract information as features from the vibration signal in many fault diagnosis study (Sakthivel, Sugumaran, Babudeva senapathy, 2010, 2012; Sugumaran and Ramachandran, 2011; Jegadeeshwaran and Ramachandran, 2015).

2.2 Frequency domain analysis

The vibration signal from a machine can be looked upon as its signature. Every component with its own spring-mass properties has its own characteristic frequency / frequencies; excitation of these contributes to the overall signature. As such the vibration level is only a total indicator of machine condition. A better approach will be to segregate vibration into distinct frequency regions and identify the vibration changes in each such region separately. In a way this amounts to the characteristic frequencies being given importance. In turn, this can lead to more specific information regarding machine condition. The frequency domain analysis techniques are centred on this idea. The principal advantage of the frequency domain analysis is that the repetitive nature of the vibration signal is clearly displayed as peaks in the frequency spectrum at the frequencies where the repetition takes place. This allows for faults, which usually generate specific characteristic frequency responses, to be detected early, diagnosed accurately and trended over time as the condition deteriorates. However, the disadvantage of the frequency domain analysis is that a significant amount of information may be lost during the transformation process. This information is non-retrievable unless a permanent record of the raw vibration signal has been made. Frequency domain analysis is otherwise known as 'spectral analysis'. It is commonly used for fault diagnosis of rotary machines (Randall, 1982, 1987), (Fansen K. and Ruheng, 2004). The characteristic defect frequency of vibration of different components can be computed for the machines operating at constant speed. The change in level of frequency of a particular band (or frequencies) can be associated to a component of a machine. In the case of component level condition monitoring study, the change in the level of frequency of a particular band can be associated with a particular condition of the component. The change in the level of frequency of a particular band gives an indication of the type of fault, thus providing required diagnostic information.

2.2.1 Fast fourier transform

In practice, the vibration signal is acquired and converted to digital form by a data acquisition system. The Discrete Fourier Transform is used to transform this signal into a digital form to the frequency domain. An efficient algorithm was introduced by Cooley and Tukey (1965) to perform DFT called Fast Fourier Transform (FFT). It is used in many of the modern spectrum analyzers, which converts the time domain signal into the frequency domain signal. There are many frequency domain analysis, namely band-pass analysis, shock pulse (spike energy), enveloped spectrum, signature spectrum, and cascades (waterfall plots). These are described in the subsequent sections.

2.2.1.1 Band-pass analysis

The frequency spectrum gives earlier warnings than monitoring overall vibration. The level of overall vibration only increases after an increasing component has become the highest peak in the spectrum. Whenever an increase of the baseline (reference) level is detected, a further analysis is carried out for fault diagnosis. The frequency range in which the levels have exceeded gives an indication of what type of faults to expect (Taylor, 1980; Mathew and Alfredson, 1984; Dadbin and Wong, 1991). The band-pass analysis involves filtering the vibration signal above or below specific frequencies in order to reduce the amount of information presented in the spectrum to a set band of frequencies. These frequencies are typical where fault characteristic responses are anticipated. Changes in the vibration signal outside the frequency band of interest are no displayed (Rades, 2008).

2.2.1.2 Spike energy

Spike Energy is a measure of the intensity of energy generated by repetitive mechanical impacts of pulses that occur as a result of surface flaws or insufficient lubrication to machine components. These impacts tend to excite the resonance response of machine components. A signal measured near rotating machine elements appears as periodic spikes of high-frequency energy and can be measured by accelerometers (Sheam and Taylor, 1990; Julien and Ming, 1995, Xu, 1995).

2.2.1.3 Envelop analysis

It is a signal processing technique, which uses a filter and rectification pre-processing of a standard accelerometer signal to reveal the bearing defects at its fundamental frequency. Hence it is referred to as the high-frequency resonance technique (Mc Fadden and Smith, 1984; Mignano, 1997). In this technique, an analog filter is used to extract the resonance excited by the bearing fault from the frequency spectrum and the detector detects the envelope. In the modern signal analysis, the useful part of the frequency spectrum is extracted through zooming around the resonance excited by the bearing defects. Hilbert transform is used to generate the envelope of the signal. The spectrum of the envelope is calculated to show the repetition frequency of the fault generated pulses.

2.2.1.4 Shock pulse method (SPM)

A shock pulse involves the analysis of the high frequency (ultrasonic) shock waves generated by metal to metal impacts in a rotating bearing, where most of the information about bearing damage can be found (Butler, 1973). Shock waves that result from metal to metal contact are short duration bursts of energy that travel at the speed of sound through the material (Lundy, 2006). As the wave travels, it dissipates energy through the structure, thereby reducing the wave pulse. The SPM is designed to detect the weak shock pulse signals using an accelerometer with a natural frequency of 36 kHz, ideally placed very close to the subject bearing. In fact, a patented design called Tandem-Piezo is used which enables the accelerometer to measure accurately both shock pulse and vibrations. A bandpass filter with 36 kHz shock pulse signal is used to isolate the shock pulse from vibration. Finally, the waveform is converted to analog pulses (Lee, 2015). This process provides a signal that can be processed to determine bearing condition.

2.2.2 Higher-order spectra

High-order spectrum, i.e., bi-spectrum or tri-spectrum, can provide more diagnostic information than power spectrum for non-Gaussian signals. The use of higher-order spectra in vibration signal is of three folds.

(a) To extract information deviating from Gaussian (normality): The higher order spectra have an important property of being identically zero for a zero – mean Gaussian process. Thus a non-zero bi-spectrum can be used to measure deviation from normality. In a normal status, the vibration signal of the mechanical system is Gaussian distribution. When faults emerge, the vibration signal will deviate from normality. Thus, the bi-spectrum is very sensitive to the working conditions of machinery.

(b) To analyze the vibration signals of machinery: It often necessary to know whether the signals contain a second order harmonic. Some faults such as misalignment, bowl shaft, cracked shaft, etc., can be monitored using second order harmonic. This second order harmonic is then tested using power spectrum averaging. However, the power spectrum averaging gives the biased estimates of the power spectrum. So it fails to test weak harmonics in low signal to noise cases, while the bi-spectrum can provide much better results.

(c) To detect and characterize the non-linear properties: When faults emerge in mechanical systems, the signal will be displayed as a non-linear state. This quadratic nonlinearity will create a quadratic phase coupling. The degree of phase coupling describes machinery fault through bi-coherence spectrum study.

In the literature, higher-order spectrum has been identified as higher-order statistics (Wang and Too, 2002). This term has its origin in the fact that bi-spectrum and tri-spectrum are actually the Fourier transforms of the third- and fourth-order statistics of the time waveform, respectively. Bi-spectrum analysis has been shown to have wide application in machinery diagnostics for various mechanical components (Yang et al., 2002) and systems (Parker et al., 2000; Choe and Fei, 1995; Arthur and Penmaan, 1995).

2.2.3 Cepstral analysis

Cepstrum has the capability to detect harmonics and sideband patterns in power spectrum. There are several versions of the definition of cepstrum (Harris and Piersol, 2002). Amongst them, the most commonly used version is power cepstrum. The power cepstrum is defined as the inverse Fourier transform of the logarithmic power spectrum (Boogert, 1963). In the power cepstrum, cepstral analysis is done in the quefrency domain and gives a measure of periodic structures in the spectrum. This series of harmonically related structures are reduced to predominantly one 'quefrency' at the reciprocal of the harmonic spacing. A high value of quefrency points to a spectrum envelope with fast changes of power spectrum density resulting from a large number of harmonics, whereas a low value of quefrency testifies to the envelope changing slowly, which results from the small number of harmonics. Cepstral analysis has proved to be a useful tool in the detection of bearing faults (Boogert, 1963), determination of voice pitch in speech analysis (Luke, 1969; Mathew, 1989) etc., The periodicity of the excitation is commonly evident in the 'quefrency' domain; however, in the frequency domain, it appears as a number of low-level sidebands (separated by the frequency of the impulses and centered about each of the resonant frequencies) which are often difficult to detect.

2.2.4 Fault detection

A fault is an unexpected change or malfunction in a system. The term 'fault detection' is defined as the process of finding the condition of the bearing - good or faulty. Fault diagnosis is a closely associated term. It is defined as the process of identifying the condition of the element under study and cause of the problem. Key fault detection techniques are presented in this sub-section. Fault diagnosis techniques are presented in subsequent sections.

2.2.4.1 Spectral comparison

A component like bearing can be considered to be in a good working environment when by itself it is free of faults and is playing its assigned role in the machinery under normally accepted working conditions. The baseline power (magnitude square) spectrum is taken for the vibration signal under these conditions. This 'baseline' spectrum is used as a reference for evaluating subsequent power spectra taken at regular intervals throughout the machine life under similar operating conditions (Mathew, 1989). The comparison is usually done on a logarithmic amplitude scale. An increase of 6 - 8 dB above the baseline is considered significant while increase beyond 20 dB is considered as a serious problem (Randall, 1985).

2.2.4.2 Spectral trending

Spectral trending (Mingsian et al., 2005; Xiamin et al., 2012) gives an indication of the rate of fault progression. In its simplest form, spectral trending involves trending of the changes in the amplitude of all (or a number of selected) spectral lines over time. For complex machines, this can often involve a large number of data, resulting in information overload due to a large number of significant spectral lines. To simplify the detection process, several parameters based on the spectrum have been proposed which provide statistical measures of spectral differences. Such spectral parameters and their performance in detection and diagnosis of bearing faults are reported in the literature (Mechefske and Mathew, 1991). It is reported that a number of these parameters performed well in the detection of the faults. However, they are not of much significance vis-à-vis diagnostic information.

2.2.5 Fault diagnosis

The process of diagnosis is performed with spectral comparison and trending; typically, only the frequencies identified as having significant changes are analyzed in detail for diagnostic purposes. The vibration spectrum of even relatively simple machines can be quite complex due to the multiple harmonic structures of the vibration from various components. The expected spectral differences associated with various bearing faults are discussed by Su and Lin (1992). Faults such as large wear and unbalance, are distributed faults causing a significant change in the mean amplitude of the vibration at discrete frequencies; these can be diagnosed easily. These faults manifest themselves as changes in a few associated frequencies in the spectrum. Eccentricity and misalignment cause low-frequency sinusoidal modulation resulting in an increase of sidebands of certain frequencies (depends on the component to be diagnosed) and their harmonics. Localized faults create short impulsive vibration, which transforms into a large number of low amplitude frequencies in the spectrum; they are difficult to diagnose or even detect.

3 Fault diagnosis through vibration analysis

Condition monitoring (CM) is predictive maintenance process which monitors the condition of machinery. This can be achieved through an instrumentation technique such as machinery vibration analysis. Vibration analysis is the most commonly used method for diagnosing rotating machines. The frequency of the vibrations can also be mapped in order to identify failures; since certain frequencies will only be present when conditions that indicate an impending defect are present. Comparison of the vibration spectra of faulty condition signal versus good condition signal will provide the information required to make a decision when maintenance is required.

Vibrations can be measured using seismic or piezo-electric transducers and eddy-current transducers from the majority of critical machines. The measuring method of the vibration signal is a complex process that requires specialized training and experience. Exceptions are state-of-the-art technologies that do the vast majority of data analysis automatically and provide information instead of data. These frequencies correspond to certain mechanical components. The location and the problem nature can be identified by examining the individual frequencies present in the signal. These frequencies correspond to certain mechanical components. Most vibration analysis instruments today utilize a fast Fourier transform (FFT) which converts the vibration signal from its time domain representation to its equivalent frequency domain representation. However, frequency analysis (sometimes called Spectral Analysis or Vibration Signature Analysis) is only one aspect of interpreting the information contained in a vibration signal.

In many industrial applications, the vibration signal is used in many fault related studies. In a study for detecting faults in rotating machine, elements describe a device for detecting damage to rotators such as ball bearings (Noda, 1977). Mechanical oscillations are converted into an electrical signal and the peak value of this electric signal is detected. In another study, a roller bearing having a hairline fracture which will generate periodic vibrations each time the fracture contacts another machine element generating periodic vibrations were measured by using a vibration transducer attached to the machine (Hicho, 1992). The transducer converts vibration signal into an electrical signal which is filtered to obtain selected frequencies of the electrical signal. The filtered electrical signal is then converted into a frequency spectrum by a fast Fourier transform. Random or spurious components are eliminated leaving only frequency components that are representative of the machine at given running speed. Corresponding frequency components are averaged and the highest average amplitude value is used as a bearing condition indicator. In another study for finding faults in machines include a fault detection system for detecting mechanical faults of machines that have one or more rotating elements (Robinson et al., 1999). The system also includes a vibration sensor for sensing vibrations generated by at least targeted rotating machine elements during machine operation to produce a vibration signal.

Vibration signal was successfully used to monitor the tire pressure in an automobile. In this system, a MEMS-based triaxial accelerometer was used to acquire the vibration signal (Hemanth Mithun Praveen and Sugumaran, 2015). In another study, vibration signal was successfully implemented to monitor the single point cutting tool (Shalet, Elangovan, Jegadeeshwaran, Sugumaran., 2014). The application of vibration analysis was implemented in various conditions monitoring system such as centrifugal pump fault diagnosis (Sakthivel et al., 2010), tool wear monitoring (Sanidhya Painuli, Elangovan, Sugumaran, 2014), bearing fault diagnosis (Hemantha Kumar, Ranjit Kumar, Amarnath, Sugumaran, 2014) and brake fault diagnosis (Jegadeeshwaran and Sugumaran, 2013).

These vibration data can be analyzed as a frequency domain data or time domain data using above mentioned techniques. Moreover, the nature of the vibration signal arising from the brake system is periodic and random. Due to wear and tear, the vibration signals obtained from an automobile brake system will not be a stationary one. Data modeling through machine learning approach can solve such problems to a greater extent. Due to wear and tear, the vibration signals obtained from a machine component will not be a stationary one. Data modeling through machine learning approach can solve such problems to a greater extent (Jin et al., 2012).

4 Machine learning

In early 1975, the goal of fault diagnosis was to store the vibration spectrum and to provide graphical tools so that the analyst can quickly access the data and determine the problem with the machine. Due to the advancement in computer technology, acquisition, storage and processing of a large amount of data have become practical. Most of the data acquisition systems are capable of logging real-time data in digital form reliably. The technological development that goes into the memory devices makes it possible to reduce the cost and size required to store large data. Today’s technology provides the memory devices with much more reliability. The processors with high processing speed allow engineers to solve complex problems. Many of the machine learning methods are iterative in nature and they require such high-speed processors. The aforesaid developments accelerate the application of machine learning methods for solving problems in real time. Fault diagnosis is one of the application areas, where machine learning methods are widely used.

Machine learning approach can be implemented through the following sequential steps. Feature extraction, feature selection, and feature classification. There are many features available namely, statistical features (Sugumaran and Ramachandran, 2007;

Jegadeeshwaran and Sugumaran, 2015), histogram features (Sakthivel, Indira, Nair, Sugumaran, 2011; Sugumaran and Ramachandran, 2011) and wavelet features (Soman and Ramachandran, 2005; Muralidharan and Sugumaran, 2012). The required features were extracted from the vibration signals through feature extraction technique.

4.1 Feature extraction

Statistical analysis of vibration signals yields different parameters which provide the physical characteristics of time domain data. Research work reported by McFadden and Smith (1984). Statistical analysis of vibration signals with different parameter combinations was used to elicit information regarding bearing faults. Such procedures use allied logic often based on physical considerations. A fairly wide set of these statistical parameters was selected as a basis for the study. They are mean, standard error, median, standard deviation, sample variance, kurtosis, skewness, range, minimum, maximum, and sum. In this study, these parameters were extracted from vibration signals and used as features.

Observing the magnitude of the time domain signal, it is found that the range of vibration amplitude varies from class to class. A better graph to show the range of variation is the histogram plot. A histogram is a graphical representation of the distribution of numerical data. It is an estimate of the probability distribution of a continuous variable (quantitative variable) and was first introduced by Karl Pearson. To construct a histogram, the first step is to "bin" the range of values—that is, divide the entire range of values into a series of intervals—and then count how many values fall into each interval. The bins are usually specified as consecutive, non-overlapping intervals of a variable. The information derived from a histogram plot can be used as features in the fault diagnosis. A representative sample from each brake condition (class) is taken and the histogram is plotted.

Autoregressive-moving-average (ARMA) models are mathematical models of the autocorrelation in a time series. ARMA models can be used to predict the behavior of a time series of past values alone. Such a prediction can be used as a baseline to evaluate the possible importance of other variables to the system. ARMA models are widely used for prediction of economic and industrial time series (Kashyap and Rangasami, 1982).

Wavelets are mathematical functions that cut up data into different frequency components, and then study each component with a resolution matched to its scale. They have advantages over traditional Fourier methods in analyzing physical situations where the signal contains discontinuities and sharp spikes. As a mathematical tool, wavelets can be used to extract information from many different kinds of data, including – but certainly not limited to – audio signals, images, and vibration signals. Sets of wavelets are generally needed to analyze data fully. A set of "complementary" wavelets will decompose data without gaps or overlap so that the decomposition process is mathematically reversible. Thus, sets of complementary wavelets are useful in wavelet-based compression/decompression algorithms where it is desirable to recover the original information with minimal loss.

4.2 Feature selection

The process of selecting the best features from a pool of features is called 'feature selection'. The good feature will have feature values with minimum variation within a class and maximum variation between the classes. The main idea of feature selection is to choose a subset of input variables by eliminating features with little or no predictive information. Feature selection can significantly improve the comprehensibility of the resulting classifier models and often build a model that generalizes better to unseen points. The features can be any measure of data points or the signal; however, the relevance of them will depend on how well they help in the process of classification. Many techniques are used for feature selection. Some of them are principal component analysis (PCA) (Suykens, Van Gestel, Vandewalle, De Moor, 2003), genetic algorithm (GA) (Samanta, Al-balushi, Al-araim, 2003), decision tree (DT) (Sakthivel et al., 2010). Among them, Principle Component Analysis (PCA), decision tree and attribute evaluator are widely used.

Principal component analysis (PCA) is one of the widely used multidimensional features transformation tools. In PCA, the amount of information is measured in terms of variance. PCA is the preferred choice because it is a simple and non-parametric method for extracting relevant information from datasets. The goal of PCA is to reduce the dimensionality of the data while retaining as much as possible of the variation in the original dataset. PCA is a technique that can be used to simplify a dataset.

A decision tree is a graph that uses a branching method to illustrate every possible outcome of a decision. It is a tree-based knowledge methodology used to represent classification rules (Sugumaran and Ramachandran, 2011). Decision trees are typically built recursively, following a top-down approach. A standard decision tree consists of a number of branches. One branch is a chain of nodes from the root to a leaf, and each node involves one attribute. The incidence of an attribute in a tree provides the information about the importance of the associated. The c4.5 algorithm is a widely used one to construct decision trees.

4.3 Feature classification

The classifier is a function which maps a set of inputs from feature space to its corresponding classes. In the present study, the classifier maps the set of extracted features to the condition of the machine components such as bearings, pump impellers, gears, and brakes. In practice, pattern classification can be carried out using many classifiers. The following sections describe briefly about the commonly used classifiers.

The condition of the brake system (good or faulty) is basically fuzzy in nature. All the faults do not occur instantly. In that case, there is no threshold value (crisp data) based on which the decision on the condition of the brake component (whether it is in a good condition or a faulty condition) can be taken. The problems of this kind can be modeled using fuzzy logic more closely (Zeng and Wang, 1991; Huang et al., 1997; Huaqing Wang and Peng Chen, 2007). For brake fault diagnosis, fuzzy logic with statistical features provides better classification accuracy as 96.5 % (Jegadeeshwaran and Sugumaran, 2015).

If the training features are separated without errors by an optimal hyperplane, the expected error rate on a test sample is bounded by the ratio of the expectation of the support vectors to the number of training vectors. The smaller the size of the support vector set, more general the above result will be. Further, the generalization is independent of the dimension of the problem. In such case a hyperplane is not possible; the next best is to minimize the number of misclassifications whilst maximizing the margin with respect to the correctly classified features. Originally, support vector machines were designed for binary classification (Quinlan, J. Ross, 1986; Hsu and Lin, 2002). Currently, there are several methods that have been proposed for multi-class classification, such as ‘‘one-against-one’’, ‘‘one-against-all’’, and directed acyclic graph (DAG) (Wang, 1989; Sakthivel et al., 2010). A kernel function is an integral part of the SVM and contributes in obtaining an optimized and accurate classifier (Yang and Zhang, 2005). A kernel function serves as a separating function, a hypersurface which optimally separates input data into two classes involving minimal support vectors. The support vectors are data points in input space lying on the kernel function hypersurface. There is no formal way to decide, which kernel function is suited to a class of classifier problem (Qingbo, Ruqiang, Fanrang, Ruxu, 2009). Most commonly used kernels are Radial Basis Function (RBF), polynomial, linear, multilayer perceptrons and sigmoid. SVM with both statistical and histogram features gives 100% classification accuracy for the roller bearing fault diagnosis (Sakthivel et al., 2010). For centrifugal pump fault diagnosis, SVM generates 100 % accuracy (Sakthivel et al., 2009). The same is applied for brake fault diagnosis also. SVM with statistical achieves 98.36 % (Jegadeeshwaran and Sugumaran, 2015).

One implementation of artificial immune systems called Clonal selection algorithm (CLONALG) inspired by the clonal selection theory of acquired immunity. Given the background, theory and an application of CLONALG to engineering applications, a new clonal selection inspired classification algorithm called Clonal Selection Classification Algorithm (CSCA) has been designed. The clonal selection theory is a theory to describe the diversity of antibodies used defends the organism from invasion (William E. Paul, 1991). An antibody is a molecule produced by B lymphocyte cells that can neutralize a single antigen. Each B lymphocyte (white blood cell) creates unique or customized antibodies of a specific type. The theory, when originally proposed was a point of contention and competed with another model called template theory. Today, the Clonal selection theory is seen as fact given the overwhelming amount of empirical evidence. An artificial immune system technique that is inspired by the functioning of the Clonal selection theory of acquired immunity is CSCA (Clonal selection Classification Algorithm). CSCA performs better with Statistical a feature which gives 98.36 % classification accuracy for brake fault diagnosis (Jegadeeshwaran and Sugumaran, 2015).

Ensemble methods are often able to generate more accurate classifiers than the individual multiclass classifiers. They may be very slow or difficult to implement. As an alternative, it is common practice to transform multiclass problems into multiple two-class ones. The dataset is decomposed into several two-class problems, the algorithm is run on each one, and the outputs of the resulting classifiers are combined. Ensemble of Nested Dichotomy (END) is one such important technique which can be used as a learning algorithm to deal with multiclass problems directly. In multi-class problems, it is possible to obtain an ensemble, by combining binary classifiers. Lin Dong et al., developed a method to improve runtime for the multi-class problem using END (Lin Dong et al., 2005). Another study reported a method to improve the classification accuracy further using forests of nested dichotomies (Rodríguez et al., 2010). The nested dichotomy algorithm was successfully implemented for the brake fault diagnosis problem in 2015 (Jegadeeshwaran and Sugumaran, 2014).

Rough set theory deals with the analysis of this classificatory property of a set of objects. The main goal of the rough set is to synthesize an approximation of concepts from acquired data. Large data sets acquired from measurements or from human experts may represent vague knowledge, for instance, uncertain or incomplete knowledge. Rough set theory provides the means to discern and classify objects in data sets of this type when it is not possible to divide the objects into defined categories. In rough set theory, knowledge is represented as a dataset in information systems. If a new attribute which represents some classification of the objects is added in the information system, then the system is called as a decision system. In most of the cases, not all of the knowledge in an information system is necessary to divide the objects into a class. In these cases, it is possible to reduce the knowledge. Reducing the knowledge results in reducts. Hence a minimal set of attributes is called as Reduct. The discernibility function is a conjunction of all the entries in the discernibility matrix that are not the empty set. The results of conjunction or simplification are the possible reducts for the information system. It is also possible to generate a discernibility function from the discernibility matrix for one of the objects in the information system. The possible reducts for the particular object are obtained by simplifying this conjunction. From the reducts computed from the discernibility matrix, the decision rules are generated for classification of the objects. The objective is to generate a minimal number of possibly shortest rules or basic minimal covering rules for all the cases. For a monoblock centrifugal pump, bearing fault diagnosis, the rough set theory was successfully studied (Sakthivel et al., 2010).

The task of classifying objects in artificial intelligence is hard because often the data may be noisy or having irrelevant attributes. Many different approaches have been tried with varying success. Some well-known schemes and their representations include ID3, which uses decision trees (Quinlan; 1986; Quinlan, 1989) and the instance-based learners IB1 - IB5 (Aha et al., 1991; Aha., 1992). These schemes have demonstrated excellent classification accuracy over a large range of domains. However, these instance-based algorithms lack to handle real-valued attributes and attributes with missing values. Many schemes which handle real feature values are extended to cope with symbolic attributes in an ad-hoc manner. Thus, a unified approach is very much needed to handle both real attributes and symbolic attributes. Hence, an instance based K Star learner was used to perform the brake fault diagnosis study (Jegadeeshwaran and Sugumaran, 2014).

Best first decision tree learning tree produces good performance models. When building models, decision tree algorithms separate instances from the root node to the terminal nodes. While performing classification, the decision tree algorithms start at the root node, test the attribute, and then move down to the tree branch corresponding to the value of the attribute. This process is repeated until a terminal node is reached. The classification of the terminal node is the predicted value for the instance. The best-first decision tree learning expands the "best" node first. It generates fully expanded tree for a given set of data. Splitting criteria are designed to measure node impurity in order to find the best node. The goal of splitting is to find the maximal decrease of impurity at each node. The decrease of impurity is calculated by subtracting the impurity values of successor nodes from the impurity of the node. Information and Gini index are the two extensively used criteria in best-first decision tree learning (Quinlan, 1986). In information theory, the decrease in impurity is measured by the information gain. Similarly, the decrease in impurity is measured by the Gini gain in Gini index (Breiman et al., 1983).

There are many machine learning algorithms which have been studied for the various component fault diagnosis problems. Amongst them, Naïve Bayes (Muralidharan and Sugumaran, 2012), Bayes net (Hemantha Kumar, Ranjit Kumar, Amarnath and Sugumaran, 2014), rough set (Sakthivel, Sugumaran and Binoy B. Nair, 2012), random forest (Babu Devasenapati and Ramachandran, 2012), ripple down rule learner (Shirazi H.

and Sammut C. A., 2008), part classifier (Frank E. and Witten I. H., 1998), locally weighted learning (Englert Peter, 2012) have been tried for the various components fault diagnosis problems.

5 Future scope

There are some more machine learning approaches which have not been even tried for the fault diagnosis study. Artificial Immune Recognition System (AIRS), Variational Mode decomposition, ant minor.

Artificial Immune Systems (AIS) can be defined as computational systems motivated by theoretical immunology and observed immune functions, principles, and models, which are useful for problem-solving. Artificial immune recognition system [AIRS] is an area of study committed to the development of computational models based on the principles of the biological immune system (BIS). It is a budding area that explores and employs different immunological mechanisms to solve computational problems. BIS is a complex, adaptive, highly distributive learning system with several mechanisms for defense against pathogenic organisms. It employs several alternatives and complementary mechanisms for protection against foreign pathogens. The immune system learns, through adaptation, to distinguish between dangerous foreign antigens and the body's own cells or molecules. Clearly, nature has been very efficient in creating organisms that are capable of defending themselves against a wide variety of pathogens such as microbes, fungi, and parasites. The powerful information – processing capabilities of the immune system, such as feature extraction, pattern recognition, learning, memory, and its distributive nature provides rich metaphors for its artificial counterpart. Immune-inspired models have been developed in an attempt to solve complex real-world problems such as abnormality detection, pattern recognition, data analysis (clustering), function optimization, fault classification and computer security. Timmis and Neal (2001) experimented with resource limited artificial immune systems. In particular, they developed the concept of an artificial recognition ball (ARB), which has the same representation as a B cell, however, it may stand for any number of indistinguishable B cells. Each ARB represents a certain number of the B cells or resources, and the total number of resources of the system is bounded. AIRS (Artificial Immune Recognition System) is a novel immune inspired supervised learning algorithm using clonal selection, affinity maturation, and affinity recognition balls (ARBs) which gives a scope for a better machine learning algorithm for brake fault diagnosis study (Watkins and Timmis, 2004).

Variational Mode Decomposition (VMD) decomposes the signal into various modes or intrinsic mode functions using the calculus of variation. Each mode of the signal may have compact frequency support around a central frequency. VMD tries to find out these central frequencies and intrinsic mode functions centered on those frequencies concurrently using an optimization methodology called “alternating direction multiplier method (ADMM)” (Konstantin Dragomiretskiy and Dominique Zosso, 2014). The original formulation of the optimization problem is continuous in the time domain. Since the classification accuracy fully depends on the condition of the signal, the raw signal obtained from the brake setup cannot be used directly for analysis due to noise. It is essential to improve the condition of the signal through some preprocessing techniques. Hence a new preprocessing technique can be used to decompose the signal into various modes or intrinsic mode functions (IMFs) using calculus variations. The modes may have compact frequency support around the central frequency. ADMM is used as an optimization tool to find such central frequencies concurrently. The main purpose of decomposing a signal is to identify various components (descriptive statistical features) of the signal. This work may focus on a new algorithm - variational mode decomposition (VMD), which extracts different modes present in the signal. The extracted statistical feature modes can be classified using the various machine learning algorithm.

Referring to the literature study, the machine learning approach has been proposed for a brake system under static condition (constant brake force and constant speed) and the results are satisfactory ((Jegadeeshwaran and Sugumaran, 2013). Hence, this review suggested that the same machine learning approach can be implemented on the real-time model (Ex. Brake system of a car) considering all possible driving conditions. The fault diagnosis model can be implemented for a brake system by developing an onboard diagnostic model which states the condition of the vehicle directly. An onboard diagnostic model consists hardware and software for extracting features from the acquired vibration signals. The feature selection and feature classification, displaying the condition of the brake system are the ongoing research attempts which have been carried out in this automobile brake fault diagnosis study.

6 Conclusions

The machine learning approaches have been successfully studied for monitoring the machine components such as gears, tool condition, bearing faults, pump impeller faults, wind turbine blade faults, brake faults, etc. Based on the above review, there are many scopes for the machine learning approaches in the fault diagnosis field. The study proved that the suitably extracted statistical and histogram information can be used for diagnosing the brake faults. Based on this information, the decision about the action to be carried out can be scheduled. This will provide an effective fault diagnosis model which states the condition of the brake system. The application of machine learning can be extended for monitoring the brake condition on a real-time brake system which is an ongoing attempt.

References

Abdullah M.; Al-Ghamd; David Mba. (2006): A comparative experimental study on the use of acoustic emission and vibration analysis for bearing defect identification and estimation of defect size, *Mechanical systems and signal processing*, vol. 20, no. 7, pp. 1537-71.

Aha D. W.; Kibler D.; Albert M. K. (1991): Instance-based Learning Algorithms, *Machine Learning*, vol. 6, no. 1, pp 37 - 66.

Aha D.W. (1992): Tolerating Noisy, Irrelevant and Novel Attributes in Instance-based Learning Algorithms, *International Journal of Man-Machine Studies*, vol. 36, no. 2, pp. 267-287.

Andrew K. S.; Jardine Daming L.; Dragon B. (2006): A review of machinery diagnostics and prognostics implementing condition-based maintenance, *Mechanical systems and signal processing*, vol. 20, no. 7, pp. 1483 - 1510.

Arthur N.; Penman J. (1995): Inverter fed induction machine condition monitoring using the bispectrum, *Proceedings of the IEEE Signal Processing Workshop on Higher-Order Statistics, Banff*, 1995.

AES2-2012 standard; (2013): Annex B (Informative) Crest Factor, pp. 17-20.

Babu Devasenapati S.; Ramachandran K.I. (2012): Random Forest Based Misfire Detection Using Kononenko Discretiser, *ICTACT Journal on Soft Computing*, vol. 2, no. 2, pp. 270 – 275.

Badaoui E.; Cohouet V.; Guillet F.; Daniere J.; Velex P. (2001): Modeling and Dectection of Localized Tooth Defects in Geared Systems, *Transactions of the ASME*, vol. 123, pp. 422-430.

Bailllie D. C,; Mathew J. (1996): A comparison of autoregressive modeling techniques for fault diagnosis of roller element bearings, *Mechanical Systems and Signal Processing*, vol. 10, no. 10, pp. 1 – 17.

Boogert B. P.; Helay M. J. R.; Tukey J. W. (1963): The quefrency analysis of Time series for Echoes: cepstrum Pseudo-Auto covariance, cross-cepstrum and shape cracking, Proceedings of the Symposium on Time Series analysis, Wiley N.Y.

Breiman L.; Friedman J.; Olshen R.; Stone C. J. (1983): *Classification and Regression Trees, CRC Press.*

Butler, D. E. (1973): The Shock-pulse method for the detection of damaged rolling bearings, *Non-Destructive Testing*, vol. 6, no. 2, pp. 92-95.

Chow T. W. S.; Fei G. (1995): Three phase induction machines asymmetrical faults identification using bi-spectrum, *IEEE Transactions on Energy Conversion*, vol. 10, no. 4, pp. 688–693.

Colin, Mercer; (2001): Time Varying Overall Level Vibration (or Noise), Prosig Noise and Vibration Blog. 2001. [http://blog.prosig.com/2001/05/01/time-varying-overall-level/]

Cooley J.W.; Tukey J. W. (1965): An algorithm for the machine calculation of complex fourier series, *Mathematics of Computing*, vol .19, pp. 297-301.

Dadbin, A.; Wong J. C. H. (1991): Different vibration monitoring techniques and their applications to roller element bearings, *International Journal of Mechanical Engineering Education*, vol. 19, no. 4, pp. 295-304.

Dalpiaz G; Rivola A.; Rubini R. (2000): Effectiveness and sensitivity of vibration processing techniques for local fault detection in gears, *Mechanical Systems and Signal Processing*, vol. 14, no. 3, pp. 387-412.

Dejie Yu.; Junsheng Cheng.; Yu Yang. (2005): Application of EMD method and Hilbert spectrum to the fault diagnosis of roller bearings, *Mechanical systems and signal processing*, vol. 19, no. 2, pp. 259-70.

Dron J. P.; Rasolofondraibe L.; Couet C.; Pavan A. (1998). Fault detection and monitoring of a ball bearing bench test and a production machine via autoregressive spectrum analysis, *Journal of sound and vibration*, vol. 218, no. 3, pp. 501-525.

Englert Peter (2012): Locally Weighted Learning, *Seminar Class on Autonomous Systems*.

Fansen K.; Ruheng C. (2004): A combined method for triblex pump fault diagnosis, based on wavelet transform, fuzzy logic and neuro-networks, *Mechanical systems and signal processing*, vol. 18, no. 1, pp. 161 - 168.

Frank E.; Witten I. H. (1998): Generating accurate rule sets without global optimization, International Conference on Machine Learning, Sanfrancisco, USA.

Harris C. M, Piersol A. G. Shock and vibration handbook, *McGraw-Hill*, *New York*, 2002.

Hemantha Kumar; Ranjit Kumar T.A.; Amarnath M.; Sugumaran V. (2014): Fault diagnosis of bearings through vibration signal using Bayes classifiers, *International Journal of Computer Aided Engineering and Technology*, vol. 6, no. 1, pp. 14-28.

Hicho M. D. (1992): Method and apparatus for analyzing rotating machines, U.S. Patent US5109700,

Hsu C. W.; Lin C. J, A. (2002): Comparison of methods for multi-class support vector machines. *IEEE Transactions on Neural Networks*, vol. 13, pp. 415–425.

Huang Y. C.; Yang H. T.; Huang C. L. (1997): Developing a new transformer fault diagnosis system through evolutionary fuzzy logic. *IEEE Transactions on Power*

Delivery, vol. 12, no. 2, pp. 761-767.

Huaqing Wang; Peng Chen. (2007): Sequential condition diagnosis for centrifugal pump system using fuzzy neural network, *Neural Information Processing - Letters and Reviews,* vol. 2, no. 1, pp. 41–50.

Hunt Trevor M. (1993): Handbook of wear debris analysis and particle detection in liquids, *Springer Science & Business Media.*

Jegadeeshwaran R, Sugumaran V. (2013) Method and Apparatus for Fault Diagnosis of Automobile Brake System Using Vibration Signals, *Recent Patents on Signal Processing,* vol. 3, no. 1, pp. 2-11.

Jegadeeshwaran R.; Sugumaran V. (2015): Brake fault diagnosis using Clonal Selection Classification Algorithm (CSCA) - A statistical learning approach, *Engineering Science and Technology, an International Journal,* vol. 18, no. 1, pp. 14 – 23.

Jegadeeshwaran R.; Sugumaran V. (2014): Vibration Based Fault Diagnosis Study of an Automobile Brake System Using K STAR (K*) Algorithm – A Statistical Approach, *Recent Patents on Signal Processing,* vol. 4, no. 1, pp. 44-56.

Jegadeeshwaran R.; Sugumaran V. (2015): Fault diagnosis of automobile hydraulic brake system using statistical features and support vector machines, *Mechanical Systems and Signal Processing,* vol. 52–53, pp. 436-446.

Jegadeeshwaran R.; Sugumaran V. (2015): Health Monitoring of a Hydraulic Brake System Using Nested Dichotomy Classifier–A Machine Learning approach, *International Journal of Prognostics and Health Management,* vol. 10, no. 1, pp. 1-10.

Jegadeeshwaran R.; Sugumaran V. (2015); Fuzzy classifier with automatic rule generation for fault diagnosis of hydraulic brake system using statistical features, *International Journal of Fuzzy Computation and Modelling,* vol. 1, no. 3, pp. 333 – 350.

Jin Y, Liu X. X, Liu W. P. (2012) Design of hydraulic Fault diagnosis system based on LabVIEW, *Advanced Materials Research,* vol. 457, no. 1, pp. 257-260.

Juan J. Rodríguez; César García-Osorio; Jesús Maudes. (2010): Forests of nested dichotomies, *Pattern Recognition Letters,* vol. 31, pp. 125 - 132.

Julien L. B. J.; Ming X. (1995): Vibration monitoring of seal less pumps using spike energy, *sound and vibration,* vol. 29, pp. 12 - 20.

Konstantin Dragomiretskiy; Dominique Zosso. (2014): Variational Mode Decomposition, IEEE Transactions on Signal Processing. vol. 62, no. 3, pp. 531 - 544.

Kashyap; Rangasami L. (1982): Optimal choice of AR and MA parts in autoregressive moving average models, Pattern Analysis and Machine Intelligence, *IEEE Transactions on,* vol. 4, no. 2, pp. 99-104.

Liaw Andy; Matthew Wiener. (2002): Classification and regression by random forest, R news, vol. 2, no. 3, pp. 18-22.

Lin Dong; Eibe Frank; Stefan Kramer. (2005): Ensembles of Balanced Nested Dichotomies for Multi-Class Problems, *Knowledge Discovery in Databases*, pp. 84-95.

Luke J. E. (1969): Automatic speaker verification using cepstral measurements, *The Journal of the Acoustical Society of America*, vol. 46, no. 4B, pp. 1026 – 1032.

Luo G. Y.; Osypiw D.; Irle M. (2000): Real-time condition monitoring by significant and natural frequencies analysis of vibration signal with wavelet filter and autocorrelation enhancement, *Journal of Sound Vibration*, vol. 236, no. 1, pp. 413–430.

Lee G. (2015): What is shock pulse method?

(http://www.reliabilityweb.comindex.php/articles/what_is_shock_pulse_method).

Mathew J. Alfredson R. J. (1984): The condition monitoring of rolling element bearingsusing vibration analysis, Journal of vibration, Acoustics, stress and reliability in design, *Transaction of ASME*, vol. 106, pp. 447-453.

Mathew, J. (1989): Monitoring the vibrations of rotating machine elements - An overview, Machine Condition Monitoring Research Bulletin, *Monash University*, vol. 1, no. 1, pp. 2.1-2.13.

Mc Fadden P. D.; Smith J. D. (1984): Vibration Monitoring of Rolling Element Bearings by high frequency resonance Technique - A review, *Tribology International*, vol. 17, no. 1, pp. 3-10.

Mechefske C.K.; Mathew J. A. (1991): comparison of frequency domain trending and classification parameters when used to detect and diagnose faults in low speed rolling element bearings, *Machine Condition Monitoring Research Bulletin*, Monash University, vol. 3, no. 1, 4.1-4.7.

Mignano F. (1997): Envelop detection, *Shock and vibration digest*, vol. 29, no. 3, pp. 18-23.

Miller A. J. (2000): A new wavelet basis for the decomposition of gear motion error signals and its application to gearbox diagnostics, *M.Sc. Thesis, Graduate Program in Acoustics, The Pennsylvania State University*, 1999.

Muralidharan V,; Sugumaran V. (2012): A comparative study of Naïve Bayes classifier and Bayes net classifier for fault diagnosis of monoblock centrifugal pump using wavelet analysis, *Journal of Applied Soft Computing*, vol. 12, no. 8, pp. 1-7.

Noda B., (1977): Device for detecting damage on rotators, U.S. Patent US4007630.

Pal, Mahesh. (2005): Random forest classifier for remote sensing classification, *International Journal of Remote Sensing* vol. 26, no. 1, pp. 217-222.

Parker B. E.; Ware H. A.; Wipf D. P.; Tompkins W. R.; Clark B. R.; Larson E. C.; Poor H. V. (2000): Fault diagnostics using statistical change detection in the bi-spectral domain, *Mechanical Systems and Signal Processing,* vol. 14, no. 4, pp. 561 - 570.

Poyhonen S.; Jover P.; Hyotyniemi H. (2004): Signal processing of vibrations for condition monitoring of an induction motor, First International Symposium on Control, Communications and Signal Processing, New York.

Prem S. Mann. (1995): Introductory Statistics (2nd ed.). Wiley, 1995.

Qingbo H.; Ruqiang Y.; Fanrang K.; Ruxu D. (2009) Machine condition monitoring using principal component representations, *Mechanical systems and signal processing,* vol. 23, pp. 446-466.

Quinlan J. R. (1986): Induction of decision trees, *Machine learning,* vol. 1, no. 1, pp. 81-106.

Quinlan J. R. (1990): Learning Logical Definitions from Relations, *Machine Learning,* vol. 5, pp. 239 - 266.

Rades M. (2008): Dynamics of Machinery – III, Editura, PRINTECH.

Randall R. B. (1985): Computer aided vibration spectrum trend analysis for condition monitoring, *Maintenance Management International,* vol. 5, no. 1, pp. 161-167.

Randall, R. B. (1982): A new method of modeling gear faults, *Journal of Mechanical Design,* vol. 104, no. 1, pp. 259-267.

Roberts T.; Talebzadeh M. (2003): Acoustic emission monitoring of fatigue crack propagation. *Journal of Constructional Steel Research,* vol. 59, no. 6, pp. 695-712.

Robinson J. C, Vanvhooris B, Miller W, 1999.Machine fault detection using vibration signal peak detector, U.S. Patent US5895857,

Sakthivel N. R, Indira V, Nair B. B, Sugumaran V, (2011) Use of histogram features for decision tree based fault diagnosis of mono-block centrifugal pump, Inter-national Journal of Granular Computing, *Rough Sets and Intelligent Systems,* vol. 2, no. 1, pp. 23–36.

Sakthivel N. R.; Sugumaran V.; Babudevasenapati S. (2010): Vibration based fault diagnosis of monoblock centrifugal pump using decision tree, *Expert Systems with Applications,* vol. 37, no. 6, pp. 4040-4049.

Sakthivel N. R.; Sugumaran V.; Babudevasenapati S. (2010): Vibration based fault diagnosis of mono-block centrifugal pump using decision tree, *Expert Systems with Applications,* vol. 37, no. 6, pp. 4040-4049.

Sakthivel N. R.; Sugumaran V.; Nair B. B. (2010): Application of support vector machine (SVM) and proximal support vector machine (PSVM) for fault classification of mono-block centrifugal pump, *International Journal of Data Analysis Techniques and Strategies*, vol. 2, no. 1, pp. 38–61.

Sakthivel N. R.; Sugumaran V.; Nair B. B. (2010): Comparison of decision tree - fuzzy and roughset - fuzzy methods for fault categorization of mono-block centrifugal pump, *Mechanical Systems and Signal Processing*, vol. 24, no. 6, pp. 1887-1906.

Sakthivel N.R.; Sugumaran V.; Binoy B. Nair (2012): Automatic rule learning using roughset for fuzzy classifier in fault categorization of mono-block centrifugal pump, *Applied Soft Computing*, vol. 12, pp. 196–203.

Samanta B.; Al-balushi K. R.; Al-araim S. A. (2003) Artificial neural networks and support vector machines with genetic algorithm for bearing fault detection, *Engineering Applications of Artificial Intelligence*, vol. 16, pp. 657–665.

Sanidhya Painuli, Elangovan M, Sugumaran V, (2014) Tool condition monitoring using K-star algorithm, *Expert Systems with Applications*, vol. 41, no. 6, 2638-2643.

Shalet K. S., Sugumaran V., Jegadeeshwaran R., Elangovan M. (2015) Condition Monitoring Of Single Point Cutting Tool Using Arma Features And SVM Classifiers, International Journal of Applied Engineering Research, vol. 10, no. 3, pp. 8401-8416.

Sheam J. M.; Taylor J. K. (1990): Using spike energy for fault analysis and machine condition monitoring, *IRD Mech analysis Technical Report.*

Shirazi H.; Sammut C. A. (2008): Acquiring control knowledge from examples using ripple-down rules and machine learning, *Iranian Journal of Science and Technology*, vol. *32*, no. B3, pp. 295 - 304.

Soman, K. P.; Ramachandran K. I. (2005): Insight into wavelets from theory to practice. Prentice-Hall of India Private Limited.

Su Y.T, Lin S.J. (1992) On initial fault detection of a tapered roller bearing: Frequency domain analysis, *Journal of Sound and Vibration*, vol. 155, no. 3, pp. 75-84.

Sugumaran V.; Ramachandran K. I. (2011): Fault diagnosis of roller bearing using fuzzy classifier and histogram features with focus on automatic rule learning, *Expert Systems with Applications*, vol. 38, no. 5, pp. 4901-4907.

Suykens J. A. K.; Van Gestel T.; Vandewalle J.; De Moor B, A. (2003): Support vector machine formulation to PCA analysis and its Kernel version, ESAT-SCD-SISTA Technical Report.

Taylor J. I. (1980): Identification of bearing defects by spectral analysis, Journal of Mechanical design, *Transaction of ASME,* vol. 102, no. 2, pp. 199 - 204.

Timmis J.; Neal M. (2001): A resource limited artificial immune system for data analysis, Knowledge-based-systems, vol. 14, no. 1, pp. 121 – 130.

Upadhyaya B.; Kitamura M.; Kerlin T.; (1980): Signature monitoring of nuclear power plant dynamics – stochastic modeling and case studies, Proceedings of the IEEE conference on Decision and control, Albuquerque, NM.

Wang C. C.; Too G. P. G. (2002): Rotating machine fault detection based on HOS and artificial neural networks, *Journal of Intelligent Manufacturing,* vol. 13, no. 4, pp. 283-293.

Wang S. S. (1989): Artificial intelligence and expert systems for diagnostics. Proceedings of the Institute of International conference machining and diagnostics and exhibition, Las Vegas, Neveda, pp. 516–512.

Watkins A.; Timmis J. (2004): Artificial Immune Recognition System (AIRS): Revisions and Refinements, Proceedings of the AISB 2004 Symposium on the Immune System and Cognition, pp. 18-26.

William E. Paul. (1991): Immunology - Recognition and Response, *Scientific American Inc., W. H. Freeman and Company*.

Xiamin Z.; Tejas H. P.; Ming J. Z. (2012): Multivariate EMD and full spectrum based condition monitoring for rotating machinery, *Mechanical systems and signal processing*, vol. 27, no. 1, pp. 712 -728.

Xu M. (1995): Spike energy and its applications, Shock and vibration digest, vol. 27, no. 3, pp. 11-17.

Yang D. M.; Stronach A. F.; Macconnell P.; Penman J. (2002) Third-order spectral techniques for the diagnosis of motor bearing condition using artificial neural networks, *Mechanical Systems and Signal Processing,* vol. 16, no. 2, 391 - 411.

Yang J. Y.; Zhang Y. Y. (2005): Application research of support vector machines in condition trend prediction of mechanical equipment. *Lecture Notes in Computer Science,* vol. 3498, pp. 857–864.

Zeng L.; Wang Z. (1991): Machine-fault classification: A fuzzy approach, *International Journal of Advanced Manufacturing Technology,* vol. 6, pp. 83–94.

Zhen L.; Zhengjia H.; Yanyang Z.; Xuefeng C. (2008): Bearing condition monitoring based on shock pulse method and improved redundant lifting scheme, *Mathematics and computers in simulation*, vol. 79, no. 3, pp. 318-338.