Measurement 46 (2013) 3247–3260

Contents lists available at SciVerse ScienceDirect

Measurement

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

Comparative study of decision tree classifier and best first tree classifier for fault diagnosis of automobile hydraulic brake system using statistical features

R. Jegadeeshwaran *, V. Sugumaran

School of Mechanical and Building Sciences, Department of Mechanical Engineering, VIT University Chennai Campus, Chennai, Tamil Nadu, India

ARTICLE INFO

Article history:
Received 26 February 2013
Received in revised form 8 April 2013
Accepted 23 April 2013
Available online 27 May 2013

Keywords:
Decision tree
Statistical features
Feature extraction
Feature selection
Best first tree
Pruning

ABSTRACT

Hydraulic brakes are the most important components in automobile. It requires advanced supervision and fault diagnosis to improve the safety of passenger, reliability and economy. Condition monitoring is one of the major division through which the reliability of such components could be monitored. The condition of the brake components can be monitored by using the vibration characteristics which will reveal the condition of the brake systems. In this paper machine learning algorithm using vibration monitoring is proposed as a possible solution to this problem. From the hydraulic brake test set up, the vibration signals were acquired by using a piezoelectric transducer and data acquisition system. C4.5 decision tree algorithm was used to extract statistical features from vibration signals. Feature selection was also carried out. Since no much of methodologies are available to find the effective number of features for a given problem, a detailed study is needed to find the best possible number of features. Hence the effect of number of features was studied by using decision tree. The selected features were classified using C4.5 decision tree algorithm and Best first decision tree algorithm with pre pruning and post pruning techniques. The results are discussed and conclusions of the study are presented.

1. Introduction

Brakes are the most important control components responsible for the safety and stability of the vehicle. Every automobile should be equipped with an efficient brake system to bring the vehicle to rest within a reasonable distance even under the most adverse conditions. In addition to the main requirement that a vehicle shall stop within a reasonable distance, it is also desirable that the retardation should be smooth and free from shudder and the rate of retardation shall be proportional to the pedal effort. This means that whilst the effort required by the driver to operate the brakes shall not be excessive. The brake system should be very reliable to promote the highest degree of safety on the road. It is not that easy to maintain a brake system. There are many things that must be taken into account. The very important idea of maintenance is safety, not alone for the person driving but also for the others moving on the road. The motivation behind this study is to stop accidents due to faulty brakes. Since 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 should be monitored all the time and diagnosed when faults occur. Monitoring of brakes is a separate area of concern in the contemporary automotive world.

Condition based monitoring is the process of monitoring a parameter of condition in machinery. A failure will indicate some significant change in its physical structure. The use of condition monitoring allows maintenance or other actions to be scheduled, to avoid the consequences of failure. Condition monitoring systems can only measure the deterioration of the condition when failure occurs.

* Corresponding author. Tel.: +91 44 39931335; fax: +91 44 39932555.
E-mail addresses: krjegadeeshwaran@gmail.com (R. Jegadeeshwaran), v_sugu@yahoo.com (V. Sugumaran).

0263-2241/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.measurement.2013.04.068

Interference in the early stages of deterioration is usually much more cost effective and life saving than allowing the brakes to fail.

Machine fault diagnosis is a branch of study concerned with finding faults arising in machine components. To identify the most probable fault, many methods were used including vibration analysis, oil particle analysis, thermal imaging, etc. Amongst them vibration analysis is most commonly used one. Comparison of the vibration spectra of faulty signal conditions Vs good signal conditions will provide the information required to make a decision when intervention is required for maintenance. The vibration signals are processed and analyzed by using wavelet analysis, spectral analysis and waveform analysis. The results of such analysis are used to determine the original cause of the fault through root cause failure analysis.

Fault diagnosis involves three main steps namely, feature extraction, feature selection, and feature classification. Features can be mainly statistical features [1], histogram features [2] and wavelet features [3,4]. In this study statistical feature were used.

Many techniques were used for feature selection including principal component analysis (PCA) [5], genetic algorithm (GA) [6], decision tree (DT) [7], fuzzy and artificial neural network [8–10]. Principal component analysis (PCA) is one of the pattern identification techniques in data of high dimension. It reduces dimensionality of data by performing a covariance analysis betw1een factors. Hence PCA models have trouble with high dimensional data or large numbers of data points, and it is not obvious how to deal properly with incomplete data set, in which some of the data points are missing. In a study, Sakthivel et al. used decision tree (DT) to identify the best features from a given set of samples for classification [10]. Based on this technique, the most important features were identified. Decision trees are simple to understand and it can be represented more compactly as an influence diagram. Hence, in the present study decision tree was used for feature selection.

A number of classifiers are available for feature classification namely, Support vector machine (SVM), Proximal support vector machine (PSVM), Naïve Bayes (NB) and Bayes Net (BN), Artificial neural network (ANN), Fuzzy, Decision tree, etc. Sakthivel et al. developed a fault classification model for mono block centrifugal pump using support vector machine and proximal support vector machines [11]. Sugumaran et al. used support vector machine (SVM) and proximal support vector machine (PSVM) to classify the faults in roller bearing using statistical features [12]. Yuan and Chu developed a SVM based fault diagnosis model [13,14]. However, the size of the patterns increases, the training time increases and also the computational complexity increases in SVM based model.

In recent studies, machine learning approaches such as Bayes net and Naïve Bayes were reported for fault diagnosis of critical components [15]. Addin and Sapuan studied about damage detection technique in engineering materials [16]. Tool condition monitoring using Naïve Bayes and Bayes net algorithms were also studied [17]. Bayes net and Naïve Bayes were successfully applied for finding faults in mono block centrifugal pump, and was proved that the classification accuracy was more than the classification done by decision tree [18].

The robustness and effectiveness of fuzzy classifier depends on the fuzzy rules. Rajakarunakaran et al. developed a fault classification model for a centrifugal pump using artificial neural network approach [19]. Wang et al. also developed a neural network model for a centrifugal pump with frequency domain signal to detect faults at early stages [8]. Eventhough it gave very good result, training of an artificial neural network classifier was complex and time consuming process.

To overcome the above difficulties, researchers need to identify a classifier model which will give a better classification accuracy with simple training operation. It should do both feature selection and feature classification simultaneously. In many fault diagnosis applications, C4.5 decision tree algorithm was successfully used for both feature selection and feature classification. A novel hybrid system based on C4.5 algorithm was proposed by Polat and Gunes, to classify the multi-class problems [20]. Condition monitoring of roller bearing using decision tree was reported by Sugumaran et al. Statistical features were extracted from vibration signals [9].

Jin Yi et al. apply the virtual instrumentation technology (Lab VIEW and MATLAB) to examine the hydraulic brake system and put forward the fault diagnosis system to vehicle hydraulic system. The diagnosis system thus effectively collects the characteristic signals of hydraulic system to analyze and compare the fault signals [21].

Many classification algorithms have been used for fault classification in various elements. In order to find a better algorithm, a detailed study is needed. This study particularly focuses the performances of best first tree classifier algorithms in the fault classification of automobile hydraulic brake system. The best first tree is highly sensitive classifier algorithm. Hence it can be used for feature classification effectively. However, classification of faults in automobile hydraulic brake system using best first tree and decision tree algorithm has not been attempted. Hence an effort was made in the present study to classify the faults in hydraulic brake using decision tree and best first tree. The flow chart of the fault diagnostic system is shown in Fig. 1.

Contributions in the present work are the following:

(1) The procedure of fault diagnosis of hydraulic brake system was illustrated. From the brake fault simulator experimental setup the following fault conditions were simulated and the vibration signals were recorded. Air in the brake fluid, brake oil spill on disc brake, drum brake pad wear, disc brake pad wear (even) – inner, disc brake pad wear (even) – inner and outer, disc brake pad wear (uneven) – inner, disc brake pad wear (uneven) – inner and outer, reservoir leak, drum brake mechanical fade.
(2) From the time domain signals, a set of statistical features were extracted and the order of importance and effective number of features were found using decision tree.

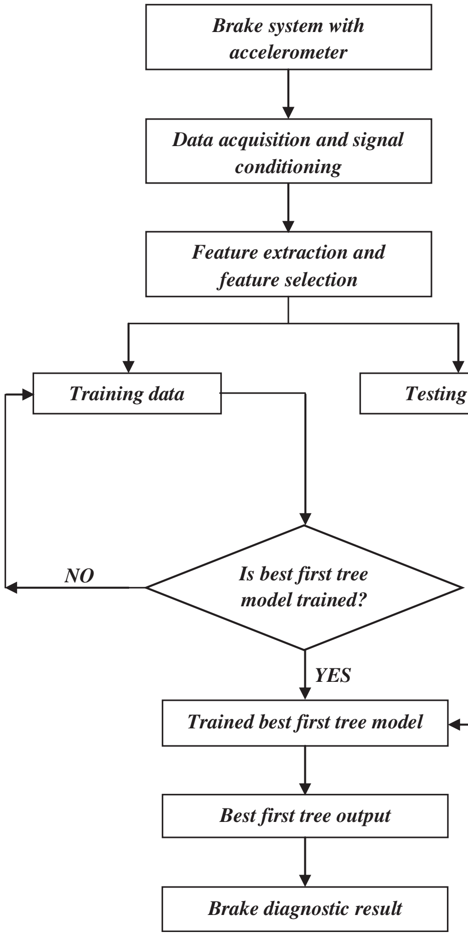

Fig. 1. Flowchart of decision tree algorithm.

(3) Best first tree (with pre pruning and post pruning) was used as classifier. Best first tree was trained and the results were compared. The results show the effectiveness of the features that were extracted features from the acquired vibration signals.

2. Experimental studies

Referring to Fig. 1, the first two blocks are described in the following sub sections, namely experimental setup and experimental procedure. The study was conducted on a test rig resting on a stationary stand.

2.1. Experimental set up

A commercial passenger car's (Maruti Swift) hydraulic brake system (Fig. 2) was used to fabricate the brake test rig. The test rig consists of disc (2 of Fig. 2.) and rear drum (1 of Fig. 2) brake coupled together by a shaft. a DC motor (1HP) (4 of Fig. 2) was used to run the shaft by means of a belt drive system (3 of Fig. 2). DC motor consists of an inbuilt drive. A lever is placed at the top of the motor which is connected to the accelerator pedal (5 of Fig. 2) providing variable speeds up to 2500 rpm. Brake pedal (6 of Fig. 2) is provided to the left side of the accelerator pedal. It is attached to the piston in the master cylinder (8 of Fig. 2) via a push rod. Master cylinder, the most important part of hydraulic brake is provided with pistons to move along the bore. Since hydraulic brakes are prominent brake system in medium motor vehicle like cars, in order to experiment with the components used in real world, branded vehicles (cars) parts were considered. The dimension of test rig is 80 cm $\times$ 80 cm $\times$ 40 cm.

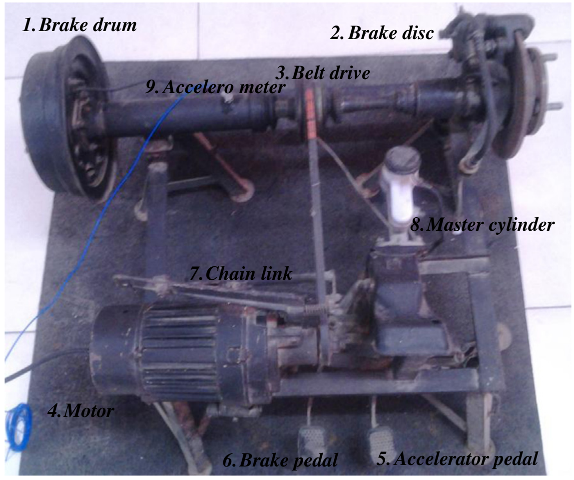

Fig. 2. Experimental setup – brake fault diagnosis.

Piezoelectric type accelerometer was used as transducer for acquiring vibration signals. It has high-frequency sensitivity for detecting faults. Hence accelerometers are widely used in condition monitoring. In this case, an uniaxial accelerometer of 50 g range, 100 mV/g sensitivity, and resonant frequency around 40 Hz was used. The accelerometer was mounted near the brake drum (and brake disc) using an adhesive mounting technique. It was connected to the DAQ system through a cable. The DAQ system used was NI USB 4432 model. The card has five analog input channels with a sampling rate and resolution of 102.4 kilo samples per second and 24-bit respectively. The accelerometer is coupled to a signal conditioning unit which consist an inbuilt charge amplifier and an analogue-to digital converter (ADC). From the ADC, the vibration signal was taken. These vibration signals were used to extract features through feature extraction technique. One end of the cable is plugged to the accelerometer and the other end to the AIO port of DAQ system. NI – LabVIEW was used to interface the transducer signal and the system (PC).

2.2. Experimental procedure

Initially the test rig was assumed to be in good condition. The vibration signals were measured from the hydraulic brake system set up under the braking condition (original speed: 40 km/h, brake load: 7 kg). From the accelerometer the vibration signals were taken with the following settings:

(1) Sample length: If the number of samples are more, then the statistical measures are meaningful. If the number of samples increases, the computation time will also increases. To strike a balance, sample length was arbitrarily chosen as 1024.
(2) Sampling frequency: Using Nyquist sampling theorem, sampling frequency was selected as 24 kHz.
(3) Number of samples: Minimum of 55 trials was taken for each conditions of the hydraulic brake system. For each trial the vibration signals were recorded by using NI LabVIEW graphical programming (Fig. 3).

Data acquisition is the process of converting analog sampling signals to digital numeric values that can be manipulated by a computer. DAC hardware is used here to interface between the sensor signal and a PC.

The following faults were simulated one at a time while all other components remain in good condition and the corresponding vibration signals were acquired.

(1) Air in the brake fluid (AE): The atmospheric air is being sucked in through the valve on the upstroke using pump and hold method.
(2) Brake Oil Spill on disc brake (BO): Using Pasteur pipette 5 ml of brake oil was applied on the brake disk.
(3) Drum Brake Pad Wear (DRPW): Drum brake pad made of asbestos with thickness 7.50 mm was chosen for the study. Using cylindrical grinding machine (silicon carbide wheel) thickness of both the brake pads (left and right) was reduced from 7.50 mm to 5.70 mm (Fig. 4a).
(4) Disc Brake Pad Wear (Even) – Inner (DWI): Asbestos of thickness 16.50 mm was selected for the study. Using surface grinding machine (silicon carbide wheel) thickness of inner brake pad was reduced from 16.50 mm to 12.40 mm (Fig. 4b).

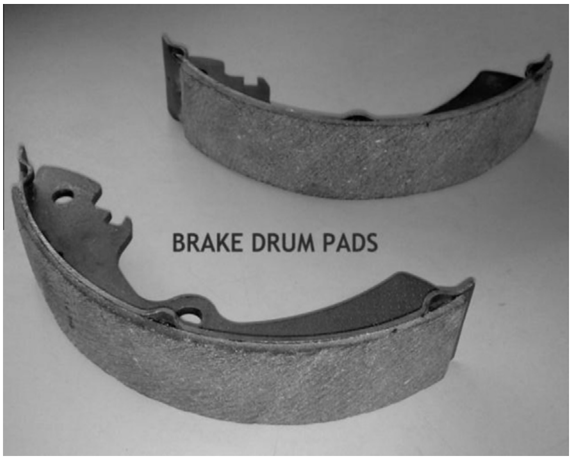

Fig. 4a. Drum brake pads wear condition.

(5) Disc Brake Pad Wear (Even) – Inner and Outer (DWIO): Disc brake pad wear under study was made of asbestos and of thickness 16.50 mm. Using surface grinding machine 16.5 mm thickness of inner brake pad made of asbestos was reduced to 11.60 mm and outer brake pad was reduced to 12.40 mm. Inner brake pad is prone to more wear than the outer brake pad, hence former was worn out more.
(6) Disc Brake Pad Wear (Uneven) (UDWI)– Inner: Disc brake pad made of asbestos with thickness 16.50 mm was machined with a downward gradient (0.6°) 15.12 mm (big radius) – 14.72 mm (small radius), using shaper machine inner brake pad.
(7) Disc Brake Pad Wear (Uneven) – Inner and Outer (UDWIO): Disc brake pad wear under study was made of asbestos and of thickness 16.50 mm. Using shaper machine inner brake pad was machined with

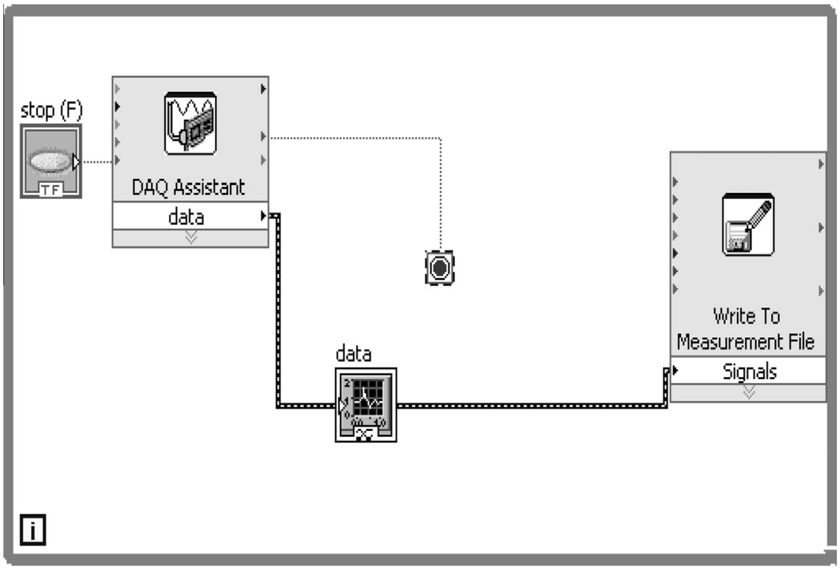

Fig. 3. LabVIEW graphical programming.

Fig. 4b. Disc brake pads wear condition.

a downward gradient (1.6) 14.76 mm (big radius) – 14.12 mm (small radii) and outer brake pad was machined with a downward gradient (0.6°) 15.12 mm (big radius) – 14.72 mm (small radius).

(8) Reservoir Leak (RL): The air tight seal of reservoir is opened. Using pipette 50 ml of brake fluid was siphoned out of reservoir and seal was left open.

(9) Drum brake Mechanical Fade (DRMF): Brake fade, is the reduction in stopping power that can occur after repeated or sustained application of the brakes, especially in high load or high speed conditions. Drum brake under study made of cast iron and has volumetric coefficient of expansion $33.3 \times 10^{-6}\ \mathrm{m}^3$ -per °C. Already used brake drum was used to simulate this condition (Fig. 4c).

Figs. 5a–5j show the time domain signals taken from the brake setup for various fault condition. This gives some basic idea about how the magnitude of the acquired vibration signal varies with respect to the faults that were simulated.

3. Feature extraction and feature selection

Feature extraction is defined as the process of computing some measures which will represent the signal. A quite ample set of statistical parameters namely standard error, kurtosis, sample variance, skewness, minimum, standard deviation, maximum, count, mean, median and mode were selected for the study (Table 1). These statistical features were extracted from vibration signals by using feature extraction technique [9]. Fig. 6 shows the decision tree from which the good features that will contribute for feature classification were selected. In the present study, minimum, sample variance, standard error, skewness and kurtosis were selected as good features and the same were classified by using decision tree classifier and best first tree classifier.

Fig. 4c. Drum brake mechanical fade.

4. Feature classification based on decision tree and best first tree classifier

4.1. C4.5 Decision tree learning

C4.5 is a statistical classifier algorithm which can be used for classification. It is a tree based knowledge representation methodology which consists of branches, root, nodes and leaves to define classification rules [13,14]. Decision tree algorithm (C4.5) has two phases namely building phase and pruning phase. In the building phase, C4.5 builds decision tree by using the concept ofi nformation theory. The tree has a single root node for the entire training set. A new node is added to the decision tree for every partition. For a set of samples in a partition S, a test attribute X is selected for further partitioning the set into $S_1, S_2, \ldots, S_L$. New nodes for S are created and these are added to the decision tree as children. The construction of decision tree depends on a test attribute X. C4.5 uses entropy based information gain as the selection criteria. The entropy is calculated as follows.

4.1.1. Information gain and entropy reduction

In information theory, entropy is a measure of the uncertainty in a random variable. Information gain is the expected reduction in entropy caused by partitioning the examples according to the given feature. It measures how well a given attribute separates the training examples according to their target function.

Information gain (S, A) of a feature A relative to a collection of examples S, is defined as

$$\text{Gain}\ (S,A) = \text{Entropy}(S) - \sum_{v \in \text{Values}(A)} \frac{|S_v|}{|S|}\text{Entropy}(S_v) \qquad (1)$$

where $S_v = (\{s \epsilon\ S|\ A(s) = v\})$.

Entropy is a measure of homogeneity of the set of examples and it is given by

$$\text{Entropy}\ (S) = \sum_{i=1}^{c} - P_i \log_2 P_i \qquad (2)$$

where 'P_i' is the proportion of 'S' belonging to the class 'i' and 'c' is the number of classes.

The second term in the equation is the expected value of the entropy after S is partitioned, using feature A. The expected entropy described by the second term is simply the sum of the entropies of each subset S_v, weighted by the fraction of examples $|S_v|/|S|$ that belongs to S_v. Hence

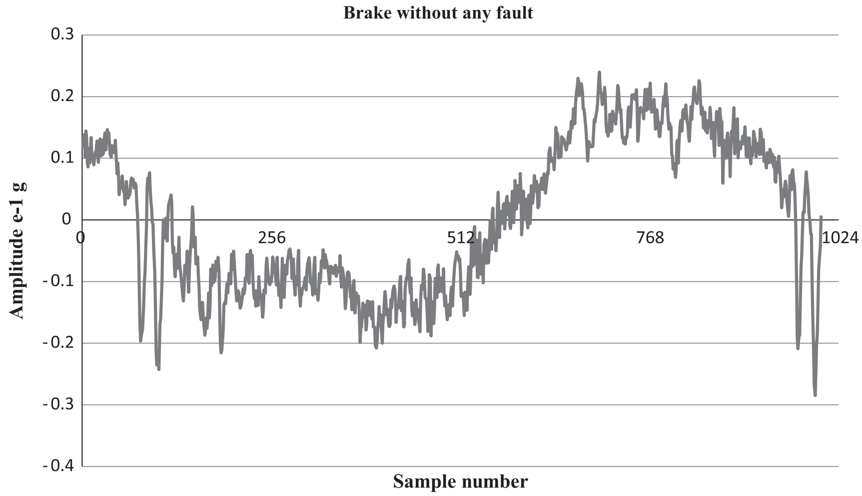

Fig. 5a. Vibration signal – brake without any fault.

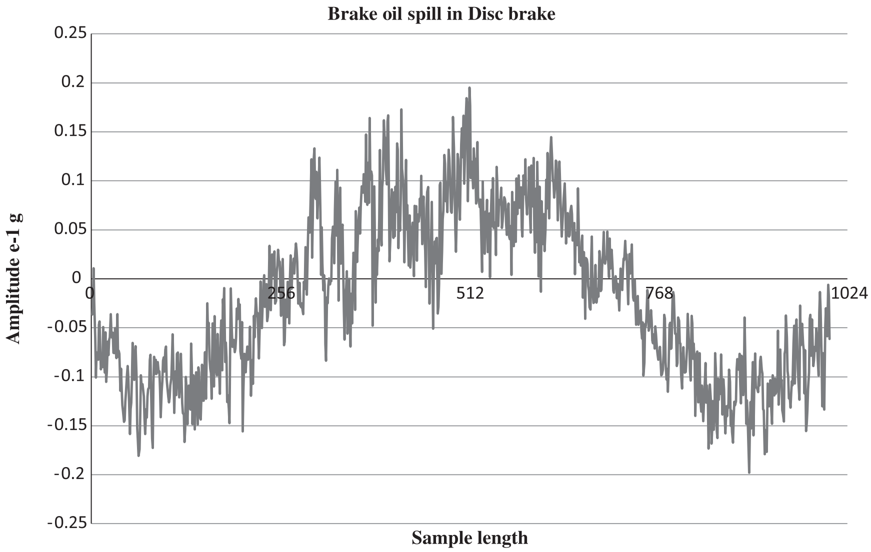

Fig. 5b. Vibration signal – brake oil spill.

Gain (*S, A*) is the expected reduction in entropy caused by knowing the value of feature *A*.

A large decision tree may have less accuracy due to over-training or under fitting. Hence a fully grown decision tree needs to be pruned to obtain better classification performance by removing the less reliable branches. In C4.5 decision tree algorithm, an error-based post pruning strategy is used to calculate the predicted error rate based on the total aggregate of misclassifications at that particular node.

4.2. Best-first tree learning

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.

4.2.1. Splitting criteria

Splitting criteria are designed to measure node impurity in order to find the best node [22]. 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 [23]. In information, the decrease in impurity is measured by the

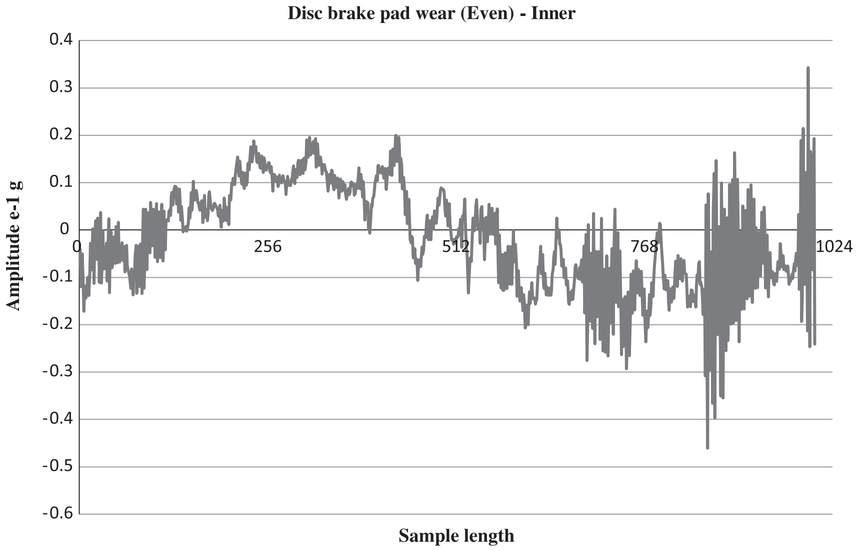

Fig. 5c. Vibration signal – disc brake pad wear (even) – inner.

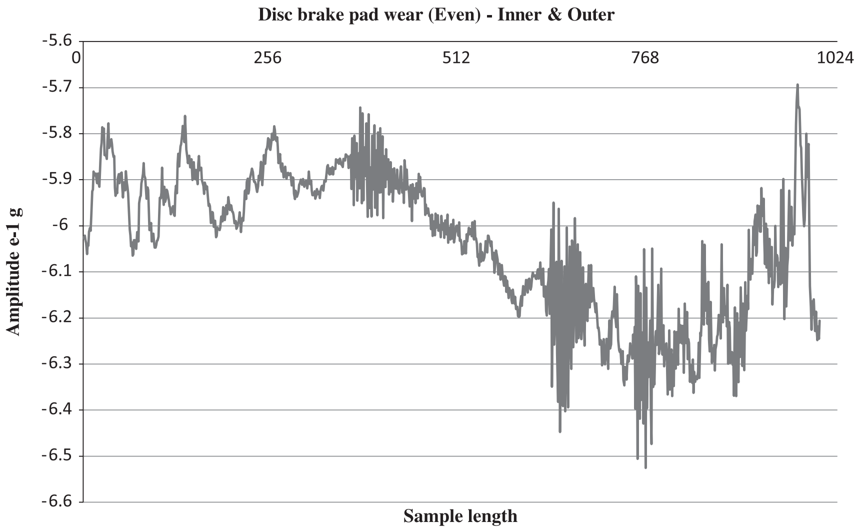

Fig. 5d. Vibration signal – disc brake pad wear (Even) – inner and outer.

information gain. Similarly, the decrease in impurity is measured by the Gini gain in Gini index (Fig. 7) [24].

The following formulas were used to find out Gini gain in Gini index.

$$\mathrm{gini}(p1, p2, \ldots, pn) = \sum_{j \neq i} p_i p_j \tag{3}$$

$$\mathrm{gini}(p1, p2, \ldots, pn) = \sum_{j \neq i} p_j(1 - p_j) = 1 - \sum_{j} p_j^2 \tag{4}$$

where p_i probability of an instance in class i, p_j probability of an instance in class j.

4.2.2. Splitting rules

The splitting rules are used to find the split which maximally reduces the impurity. In splitting criterion, the splitting rules find the split which leads to maximal information gain or Gini gain. The goal of the splitting rules is to find the minimal values of the weighted sum of the information values or the Gini index values of its successor nodes.

4.2.3. Construction of best first decision trees

Like standard decision trees, best-first decision trees are constructed based on the following procedural steps. (i) To find the best attribute to split, (ii) to find which node is to be expanded next and (iii) to make the decision when to stop growing trees. Best-first decision tree learning chooses the best node to split at each step. In order to find the best node, sort all node in the list in descending order according to Gini gain or information gain. After sorting, the first node is to be expanded next. If the reduction of impurity of the first node is zero, then the reduction of

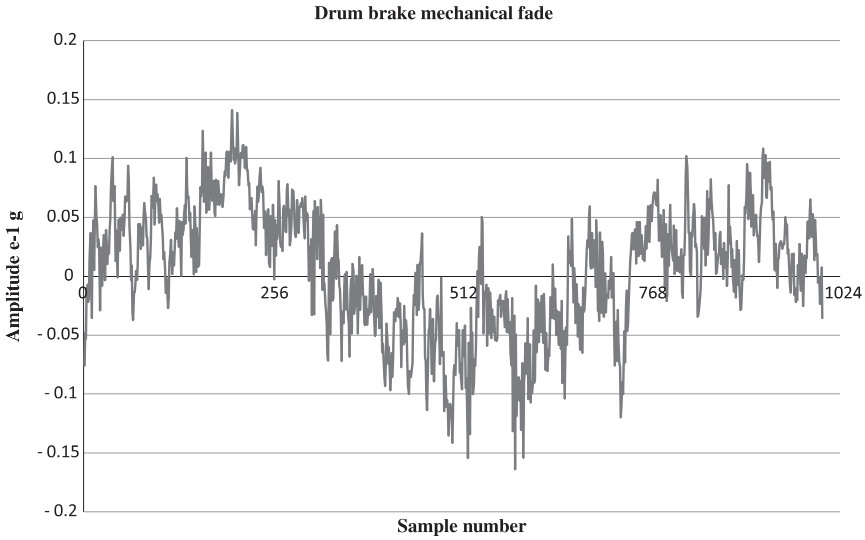

Fig. 5e. Vibration signal – drum brake mechanical fade.

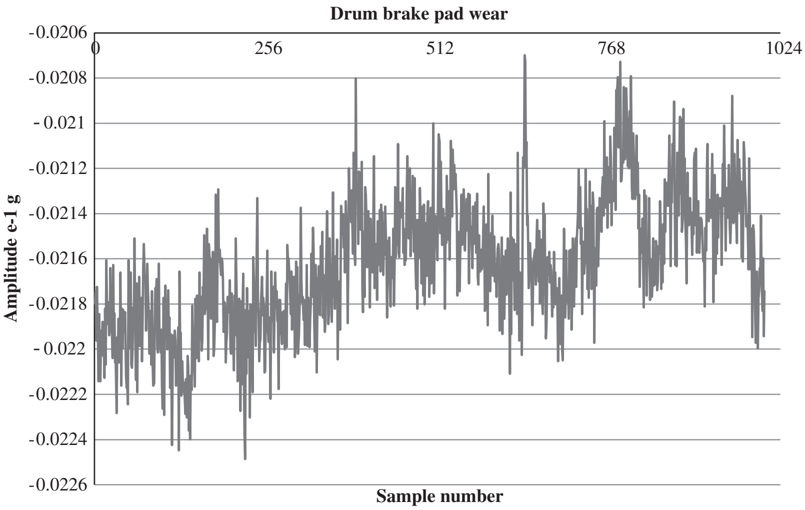

Fig. 5f. Vibration signal – drum brake pad wear.

all nodes is also zero. Thereafter further split cannot be possible. Hence the stopping criteria, stops expanding a tree the impurity of all nodes cannot be reduced by further splitting. If a fixed number is specified in the best-first decision tree learning, the tree stops expanding. Hence the stopping criterion enables us to investigate new pre-pruning and post-pruning methods by choosing the fixed number of expansions based on cross-validation. Then all training folds are created parallel in post pruning. During this process the average error estimate is calculated.

This step is repeated until the tree cannot be expanded any more. Then a sequence of the number of expansions and their corresponding error estimates based on the cross-validation can be calculated. The number at which the error estimate is minimal will be chosen as the final number of expansions. According to this number of expansions the final tree is built.

4.2.4. Kappa statistics

Kappa statistic is used in assessing the degree of agreement between two or more raters, while examining the same data. Let each object in a group of M objects is assigned to one of n categories. The categories are at nominal scale. For each object, such assignments are done by k raters. Then the kappa measure of agreement is given by the ratio:

$$k = \frac{P(A) - P(E)}{1 - P(E)} \tag{5}$$

where $P(A)$ is the proportion of times the k raters agree, $P(E)$ is the proportion of times the k raters are expected to agree by chance alone. If $k = 1$, then it is called as complete agreement; If $k = 0$, then it called as lack of agreement. Negative values of 'k' would mean negative agreement which gives the propensity of raters to avoid assignments made by other raters.

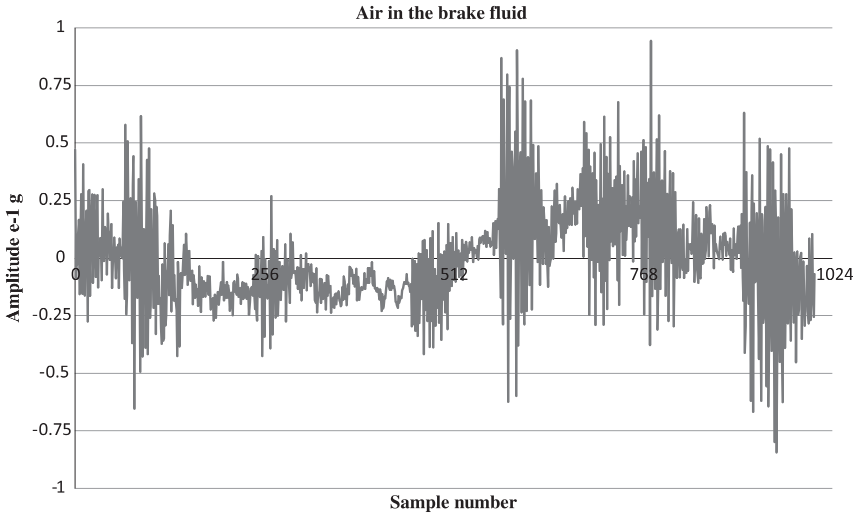

Fig. 5g. Vibration signal – air in the brake fluid.

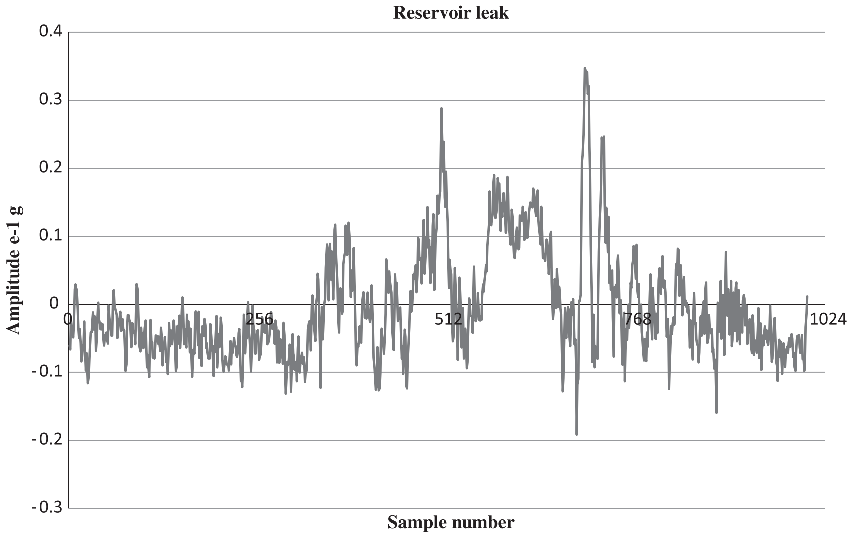

Fig. 5h. Vibration signal – reservoir leak.

5. Results and discussion

The fault diagnosis of hydraulic brake system was taken up. Machine learning approach was used with statistical feature, decision tree and best first tree. The results are discussed below.

Twelve set of statistical parameters namely standard error, kurtosis, sample variance, skewness, standard deviation, minimum, maximum, count, mean, median, range and sum were extracted for the study.

5.1. Effect of number of features

(1) Decision tree shows their relative importance of each feature in classification. The level of contribution by individual feature is given within the parenthesis in the decision tree (Fig. 6). The first number in the parenthesis indicates the number of data points that can be classified correctly using that feature set. The second number indicates the number of samples against this action. The features can be ignored, if the first number is very small compared to the total number of samples.

(2) Five contributors namely, minimum, sample variance, standard error, kurtosis and skewness were identified as leading contributors amongst the features that were extracted.

(3) Referring to decision tree (Fig. 6), 'minimum' is the best feature. The reason is that the features were selected based on entropy reduction and information gain. The information gain is a measure of discriminating capability of a feature for the given data set.

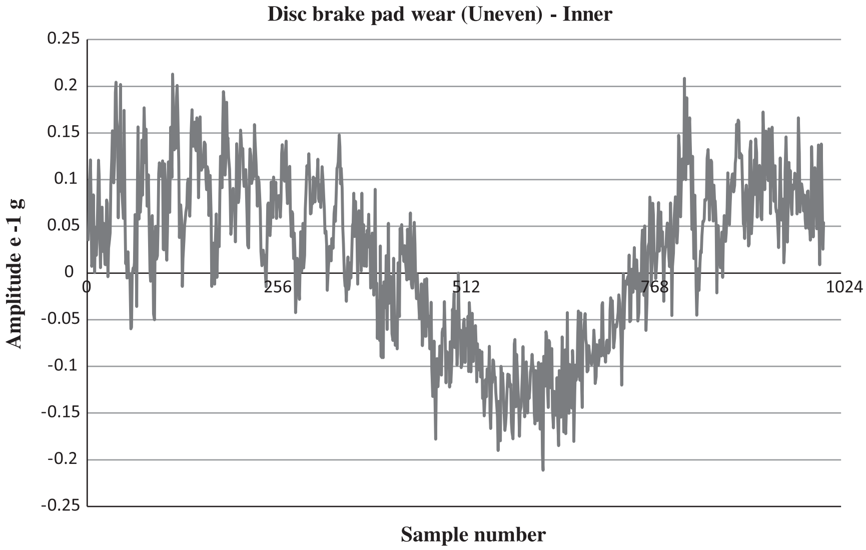

Fig. 5i. Vibration signal – disc brake pad wear (uneven) – inner.

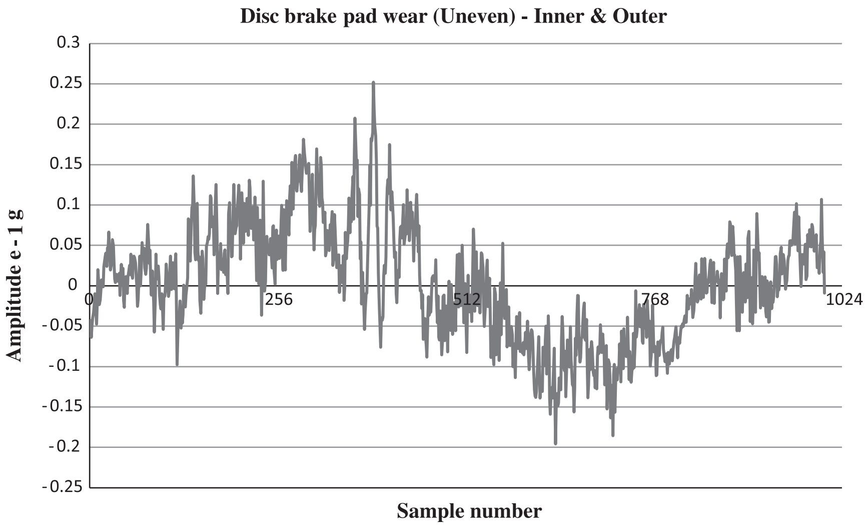

Fig. 5j. Vibration signal – disc brake pad wear (uneven) – inner and outer.

(4) The classification accuracy of decision tree and best first tree with pre pruning and post pruning process for statistical features is presented in Table 2. Fig. 8 shows the graph of classification accuracy Vs number of features. It indicates that when the number of features is five, the classifier gives good accuracy.
(5) Referring Table 2, the accuracy falls down beyond certain number of features due to the following reasons:
 (i) Unnecessary confusion during training.
 (ii) They increase the complexity of the problem.

5.2. Statistical features classification using C4.5 decision tree algorithm

1. From the decision tree, five features that are contributed for classification were only selected for training and testing. They are: (1) minimum (2) standard error, (3) sample variance, (4) kurtosis and (5) skewness.
2. The effective number of features required to obtain maximum classification accuracy is given in Table 2. The corresponding classification accuracy for C4.5 decision tree, best first tree with pre pruning and best first tree with post pruning is presented. The classifier gives

Table 1
Statistical features definition.

Name of the Statistical features	Formula/description
Standard error	$\sqrt{\frac{1}{n-2}\left[\sum(y-\bar{y})^2 - \frac{\sum[(x-\bar{x})(y-\bar{y})^2]}{\sum(x-\bar{x})^2}\right]}$
Standard deviation	$\sqrt{\frac{n\sum x^2-(\sum x)^2}{n(n-1)}}$
Sample variance	$\sqrt{\frac{n\sum x^2-(\sum x)^2}{n(n-1)}}$
Kurtosis	$\left\{\frac{n(n+1)}{(n-1)(n-2)(n-3)}\sum\left(\frac{x_i-\bar{x}}{S_d}\right)^4\right\} - \frac{3(n-1)^2}{(n-2)(n-3)}$
Skewness	$\frac{n}{(n-1)(n-2)}\sum\left(\frac{x_i-\bar{x}}{S_d}\right)^3$
Maximum value	Maximum signal point value in a given signal
Minimum value	Minimum signal point value in a given signal
Range	Difference in maximum and minimum signal point values for a given signal
Sum	Sum of all feature values for each sample
Mean	The arithmetic average of a set of values or distribution
Median	Middle value separating the greater and lesser halves of a data set

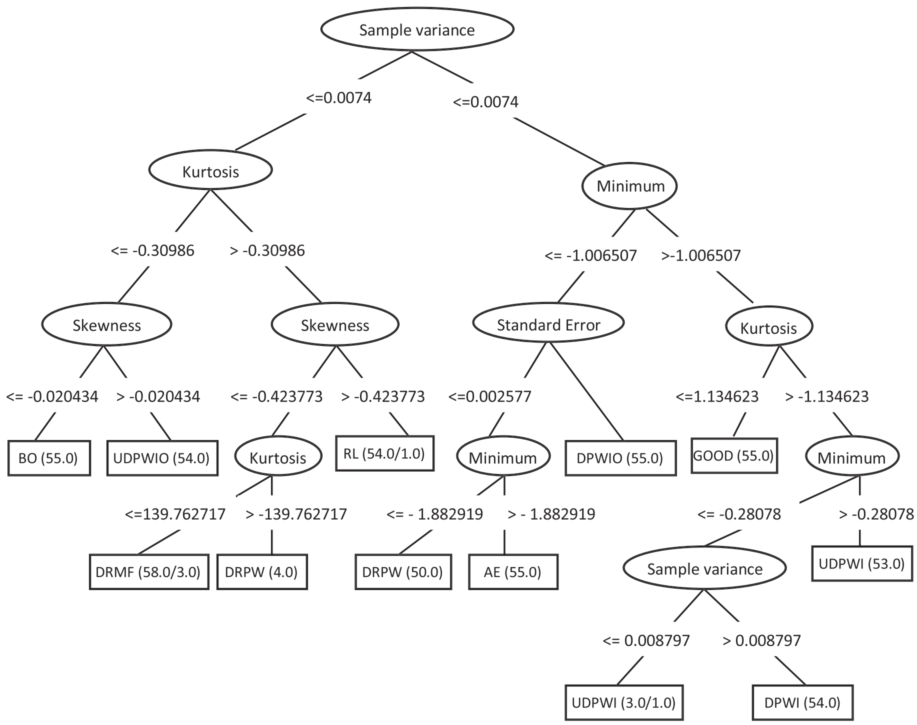

Fig. 6. Decision tree.

good result if the number of features is five in each class. The strategy is to select the number of features which gives less computation time.

3. Mis-classification details for best first decision tree algorithm with post pruning and pre-pruning processes are presented in Tables 4 and 5 respectively. Summary of stratified cross validation is given below. C4.5 decision tree gives gives better classification accuracy (97.45%) for a selected number of features.

Total number of instances	550	
Correctly classified instances	536	97.45%
Incorrectly classified instances	14	2.55%
Kappa statistic	0.9717	

4. Misclassification details are presented as confusion matrix in Table 3. In confusion matrix the first row represents the total number of data points corresponding

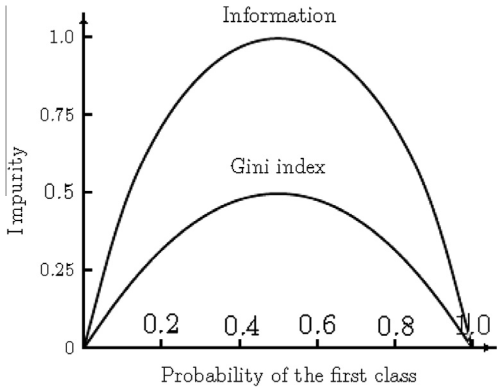

Fig. 7. Gini index.

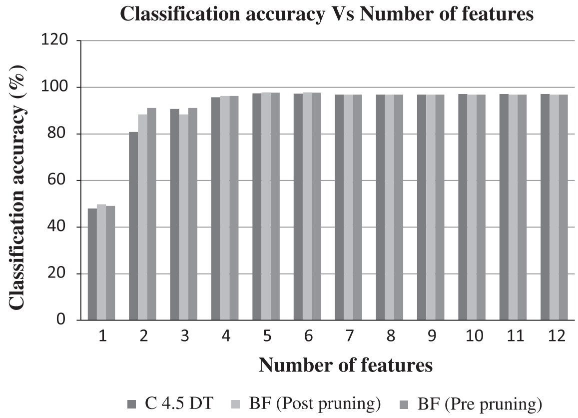

Fig. 8. Classification accuracy Vs number of features.

to "GOOD" condition of the brake system. The first column in first row represents how many is correctly classified as "GOOD" condition. The total in the first row is 55 out of that 54 is correctly classified and one is misclassified as "disc brake pad wear (even) – inner (DPWI)" condition. Similarly, the other elements in the first row are zero. It means no other data sets in "GOOD" condition are misclassified as other faulty conditions.

5. The second row in the confusion matrix represents the total number of data points related to "brake oil spill" condition; the first column represents misclassification of those data points as "GOOD" condition. In second row, only one data point among 55 data sets is misclassified as "drum rake mechanical fade (DRMF)". Second row second column in the confusion matrix represents how many of "brake oil spill (BO)" data points have been correctly classified as "brake oil spill" condition. In a "disc brake pad wear inner and outer (DPWIO)" condition, out of 55 data points all are correctly classified and there is no misclassification.

5.3. Statistical features classification using best first tree decision tree algorithm

1. From the decision tree, five features that are contributed for classification were only selected for training and testing. They are: (1) sample variance (2) standard error, (3) minimum value, (4) kurtosis and (5) skewness.
2. The selected features were trained and tested by using best first tree algorithm with pre pruning and post pruning processes. Summary of stratified cross validation is given below. Best first decision tree gives better classification accuracy (97.81%) for a selected number of features.

Pre pruning operation		
Total number of instances	550	
Correctly classified Instances	538	97.81%
Incorrectly classified instances	12	3.37%
Kappa statistic	0.9758	
Post pruning operation		
Total number of instances	550	
Correctly classified instances	537	97.63%
Incorrectly classified instances	13	3.37%
Kappa statistic	0.9737	

3. Referring the above summary best first tree with pre pruning process gives a better.
4. Efficiency (97.81%) than that of the post pruning processes.

Table 2
Classification accuracy of statistical features using C4.5 decision tree and best first decision tree algorithm.

No. of features	Classification accuracy		
	C4.5 decision tree	Best first tree (Pre pruning)	Best first tree (Pre pruning)
1	48.00	49.81	49.09
2	80.90	88.36	91.09
3	90.72	88.36	91.09
4	95.81	96.36	96.36
5	**97.45**	**97.81**	**97.63**
6	97.27	97.81	97.63
7	96.90	96.90	96.90
8	96.90	96.90	96.90
9	96.90	96.90	96.90
10	97.09	96.90	96.90
11	97.09	96.90	96.90
12	97.09	96.90	96.90

Table 3
Confusion matrix using C4.5 decision tree algorithm time taken to build model: 0.01 s.

Category	GOOD	BO	DPWI	DPWIO	UDPWI	UDPWIO	DRMF	DRPW	AE	RL
GOOD	54	0	1	0	0	0	0	0	0	0
BO	0	54	0	0	0	0	1	0	0	0
DPWI	0	0	54	0	0	0	0	0	1	0
DPWIO	0	0	0	55	0	0	0	0	0	0
UDPWI	0	0	0	0	53	0	0	1	0	1
UDPWIO	0	0	0	0	1	54	0	1	0	0
DRMF	0	0	0	0	0	0	52	1	0	2
DRPW	0	0	0	0	0	0	1	54	0	0
AE	0	0	1	1	0	0	0	0	53	0
RL	0	0	0	0	0	0	2	0	0	53

GOOD: Brake without any fault; BO: Brake oil spill; DPWI: Disc brake pad wear – Inner; DPWIO: Disc brake pad wear Inner and outer; UDPWI: Uneven disc pad wear (Inner) UDPWIO: Uneven disc pad wear (Inner and Outer); DRMF: Drum brake mechanical fade; DRPW: Drum brake pad wear; AE: Air in brake fluid; RL: Reservoir leak.

Table 4
Confusion matrix using best first decision tree algorithm with post pruning processes Time taken to build model: 0.01 s.

Category	GOOD	BO	DPWI	DPWIO	UDPWI	UDPWIO	DRMF	DRPW	AE	RL
GOOD	55	0	0	0	0	0	0	0	0	0
BO	0	55	0	0	0	0	0	0	0	0
DPWI	0	0	55	0	0	0	0	0	0	0
DPWIO	0	0	0	54	0	0	0	0	1	0
UDPWI	0	0	2	0	53	0	0	1	0	1
UDPWIO	0	0	0	0	0	53	2	0	0	0
DRMF	0	0	0	0	0	1	52	0	0	1
DRPW	0	0	0	0	0	0	0	54	0	1
AE	0	0	0	0	0	0	0	0	55	0
RL	0	0	0	0	0	0	3	0	0	52

Table 5
Confusion matrix using best first decision tree algorithm with pre pruning processes time taken to build model: 0.11 s.

Category	GOOD	BO	DPWI	DPWIO	UDPWI	UDPWIO	DRMF	DRPW	AE	RL
GOOD	55	0	0	0	0	0	0	0	0	0
BO	0	55	0	0	0	0	0	0	0	0
DPWI	0	0	55	0	0	0	0	0	0	0
DPWIO	0	0	0	54	0	0	0	0	1	0
UDPWI	0	0	3	0	52	0	0	0	0	0
UDPWIO	0	0	0	0	0	53	2	0	0	0
DRMF	0	0	0	0	0	2	52	0	0	1
DRPW	0	0	0	0	0	0	0	54	0	1
AE	0	0	0	0	0	0	0	0	55	0
RL	0	0	0	0	0	0	3	0	0	52

6. Conclusion

In this study, C4.5 decision tree algorithm and best first decision tree algorithm were used as classifiers for classifying brake faults using the statistical features extracted from the vibration signals of a brake test rig. The accuracy of C4.5 decision tree classifier was found to be 97.45% while Best first tree classifier system with post pruning operation gave an accuracy of 97.63%. Best first tree classifier with pre pruning yielded an accuracy of 97.81%. The best first tree classifier with post pruning is more accurate when compared to the best first decision tree classifier with post pruning and C4.5 decision tree classifier. Hence, best first tree classifier with statistical features can be used to monitor the condition of a hydraulic brake system of an automobile. The proposed system in the present study gives a possible application to automate the condition monitoring system in online. It informs the driver of any faults. This will significantly reduces the cause of major or minor accidents and thus increasing the safety of human life.

References

[1] V. Sugumaran, K.I. Ramachandran, Effect of number of features on classification of roller bearing faults using SVM and PSVM, Expert Systems with Applications 38 (2011) 4088–4096.
[2] N.R. Sakthivel, V. Indira, B.B. Nair, V. Sugumaran, Use of histogram features for decision tree based fault diagnosis of monoblock centrifugal pump, International Journal of Granular Computing, Rough Sets and Intelligent Systems (IJGCRSIS) 2 (2011) 23–36.
[3] F. Konga, R. Chen, A combined method for triplex pump fault diagnosis based on wavelet transform, fuzzy logic and neuro-networks, Mechanical Systems and Signal Processing 18 (2004) 161–168.

[4] K.P. Soman, K.I. Ramachandran, Insight into Wavelets from Theory to Practice, Prentice-Hall of India Private Limited, 2003.
[5] J.A.K. Suykens, T. Van Gestel, , Vandewalle, J., & De Moor, B., A support vector machine formulation to PCA analysis and its Kernel version, ESAT-SCD-SISTA Technical Report, 2003.
[6] B. Samanta, K.R. Al-balushi, S.A. Al-araim, 'Artificial neural networks and support vector machines with genetic algorithm for bearing fault detection', Engineering Applications of Artificial Intelligence 16 (2003) 657–665.
[7] N.R. Sakthivel, V. Sugumaran, S. Babudevasenapati, Vibration based fault diagnosis of monoblock centrifugal pump using decision tree, International Journal of Expert Systems with Applications 2 (2010) 38–61.
[8] H.Q. Wang, P. Chen, Sequential condition diagnosis for centrifugal pump sys-tem using fuzzy neural network, Neural Information Processing: Letters and Reviews 2 (2007) 41–50.
[9] V. Sugumaran, K.I. Ramachandran, Automatic rule learning using decision tree for fuzzy classifier in fault diagnosis of roller bearing, Mechanical Systems and Signal Processing 21 (2007) 2237–2247.
[10] N.R. Sakthivel, V. Sugumaran, B.B. Nair, Automatic rule learning using roughset for fuzzy classifier in fault categorization of centrifugal pump, International Journal of Applied soft computing 12 (2012) 196–203.
[11] N.R. Sakthivel, V. Sugumaran, B.B. Nair, 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 2 (2010) 38–61.
[12] V. Sugumaran, V. Muralidharan, K.I. Ramachandran, Feature selection using decision tree and classification through proximal support vector machine for fault diagnostics of roller bearing, Mechanical Systems and Signal Processing 21 (2007) 930–942.
[13] S.F. Yuan, F.-L. Chu, Support vector machines-based fault diagnosis for turbo-pump rotor, Mechanical Systems and Signal Processing 20 (2006) 939–952.
[14] S.-F. Yuan, F.-L. Chu, Fault diagnostics based on particle swarm optimization and support vector machines, Mechanical Systems and Signal Processing 21 (2007) 1787–1798.
[15] J. Chen, H. Huan, et al., Feature selection for text classification with Naïve Bayes, Expert Systems with Applications 36 (2009) 5432–5435.
[16] O. Addin, S.M. Sapuan, et al., A Naïve-Bayes classifier for damage detection in engineering materials, Materials and Design 28 (2008) 2379–2386.
[17] M. Elangovan et al., Studies on Bayes classifier for condition monitoring of single point carbide tipped tool based on statistical and histogram features, Expert Systems with Applications 37 (2010) 2059–2065.
[18] V. Muralidharan, V. Sugumaran, A comparative study of Naïve Bayes classifier and Bayesnet classifier for fault diagnosis of monoblock centrifugal pump using wavelet analysis, Journal of Applied Soft Computing (2012) 1–7.
[19] S. Rajakarunakaran, P. Venkumar, D. Devaraj, K. Surya Prakasa Rao, Artificial neural network approach for fault detection in rotary system, Applied Soft Computing 8 (2008) 740–748.
[20] K. Polat, S. Gunes, A novel hybrid intelligent method based on C4.5 decision tree classifier and one-against-all approach for multi-class classification problems, Expert Systems with Applications 36 (2009) 1587–1592.
[21] Yi Jin, Xi Xia Liu, Wei Ping Liu, Design of hydraulic Fault diagnosis system based on Lab VIEW, Advanced Materials research 458–457 (2012).
[22] J.R. Quinlan, Induction of Decision Trees, 1986.
[23] J.R. Quinlan, C4.5: Programs for, Machine Learning, 1993.
[24] L. Breiman, J. Friedman, R. Olshen, C.J. Stone, Classification and Regression Trees, 1983.