

Health Monitoring of a Hydraulic Brake System Using Nested Dichotomy Classifier – A Machine Learning approach

R. Jegadeeshwaran¹, V. Sugumaran²

^{1,2}*School of Mechanical and Building Sciences, VIT University, Chennai Campus, Chennai, Tamil Nadu, 600 127, India*

¹*jegadeeshwaran.r@vit.ac.in*
²*sugumaran.v@vit.ac.in*

ABSTRACT

Hydraulic brakes in automobiles play a vital role for the safety on the road; therefore vital components in the brake system should be monitored through condition monitoring techniques. Condition monitoring of brake components can be carried out by using the vibration characteristics. The vibration signals for the different fault conditions of the brake were acquired from the fabricated hydraulic brake test setup using a piezoelectric accelerometer and a data acquisition system. Condition monitoring of brakes was studied using machine learning approaches. Through a feature extraction technique, descriptive statistical features were extracted from the acquired vibration signals. Feature classification was carried out using nested dichotomy, data near balanced nested dichotomy and class balanced nested dichotomy classifiers. A Random forest tree algorithm was used as a base classifier for the nested dichotomy (ND) classifiers. The effectiveness of the suggested techniques was studied and compared. Amongst them, class balanced nested dichotomy (CBND) with the statistical features gives better accuracy of 98.91% for the problem concerned.

1. INTRODUCTION

Fault diagnosis is an important process in preventive maintenance, as it avoids serious damage during operation. Early detection of the defects can prevent the system from malfunction which leads to damage of the entire system or accident. Therefore, a condition monitoring system can be effectively used as a decision support tool to identify failures. The brake system in an automobile is such an essential component which must be monitored continuously to avoid serious damage. Hence the malfunction of the brake system can be identified through its symptoms or some warning sign. Miller et al., proposed a method for

monitoring the applications of the brakes in aircrafts. This device comprises a chart recorder with traces driven by a transducer for measuring the brake force (Miller, Marshall, Aexander Baiey & Griffin, 2004). Reinecke (1988) developed an apparatus for measuring and / or regulating a braking force using sensors. In both the cases, the sensors have been used to measure some parameters like brake temperature, friction force and braking force. No such system has been proposed to measure brake pad wear, mechanical fade, reservoir leak, etc. The brake faults like brake pad wear, mechanical fade of drum brake, reservoir leak can alter the characteristics of the vibration signals. It may not be possible to accurately measure the amount of wear / fade; however, one can capture some identification of them for the purpose of fault diagnosis. Hence the vibration based fault diagnosis approach has been attempted in the present study to monitor the condition of a brake system. In condition monitoring of rotating machine components, vibration and acoustic emission (AE) signals are widely used to identify faults (Nowicki, Slowinski & Stefanowski, 1992). The AE signal is complex and stochastic in nature. It can be used for applications with a very high frequency range (20 kHz – 100 MHz). Because of its high frequency content, AE techniques require much higher sampling rates than vibration based techniques. It needs more time to acquire the signal. Hence the cost of the sensors and the data acquisition will be more. Moreover the vibration signals are considerably less complex in nature, more comprehensive, and convenient to analyze. Hence it was used for identifying brake faults. 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 (Shen Yin, Ding, Haghani, Hao & Zhang, 2012).

The machine learning approach can be implemented through the following sequential steps. Feature extraction, feature selection, and feature classification. There are many features

Jegadeeshwaran et al. This is an open-access article distributed under the terms of the Creative Commons Attribution 3.0 United States License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

available, namely, statistical features (Sugumaran & Ramachandran 2007; Jegadeeshwaran & Sugumaran 2013, 2014), histogram features (Sakthivel, Indira, Nair, & Sugumaran, 2011) and wavelet features (Muralidharan & Sugumaran, 2012) (Soman & Ramachandran 2005). The present study focuses on statistical features. The required statistical features were extracted from the vibration signals through feature extraction technique.

For feature selection, many techniques are available. 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, Sugumaran & Babudevasenapathy 2010) (Sugumaran & Ramachandran 2011). In the present study attribute evaluator has been used for feature selection.

Number of classifiers has been already reported in literature for solving problems like bearing fault diagnosis, tool condition monitoring, condition monitoring of centrifugal pumps, etc. The condition monitoring problem in rotating systems has been treated as a classification problem using artificial neural network (ANN) based on the training pattern (Rajakarunakaran, Venkumar, Devaraj & Surya Prakasa Rao, 2008). However the traditional ANN has limitations on convergence. Support vector machines (SVM) and Proximal support vector machines (PSVM) were reported for the fault classification of bearings (Sugumaran, Muralidharan & Ramachandran 2007). As the size of the patterns increases, the training time increases, thereby, the computational complexity also increases in PSVM. SVM has high classification accuracy and good generalization capabilities only for crisp data (Burgess, 1998) (Jack & Nandi, 2000). Fault classification of mono block centrifugal pump using Naïve Bayes (NB) (Addin & sapuan, 2008) (Huang, 2009) and Bayes Net (BN) (Muralidharan & Sugumaran, 2012) were reported. However, NB and BN may not be suitable for many applications; since, it often fails to produce a good estimate of correct class probabilities.

Fuzzy classifier was used to classify the statistical features extracted from the vibration signals of the faulty gear box (Saravanan, Cholaairajan & Ramachandran 2009). Moreover, the computation time involved in fuzzy classifier is less compared to ANN and SVM (Sakthivel, Sugumaran & Nair 2010). C4.5 decision tree and best first tree algorithm have been reported for the brake related study (Jegadeeshwaran & Sugumaran 2013). Each multiclass classifier has its own demerits on the fault classification problems. Hence it is very important to find a suitable method to classify faults with high accuracy.

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 multi class problems using Ensembles of Balanced Nested Dichotomies (Dong, Frank & Kramer 2005). Another study reported a method to improve the classification accuracy further using forests of nested dichotomies (Rodríguez, Osorio & Maudes 2010).

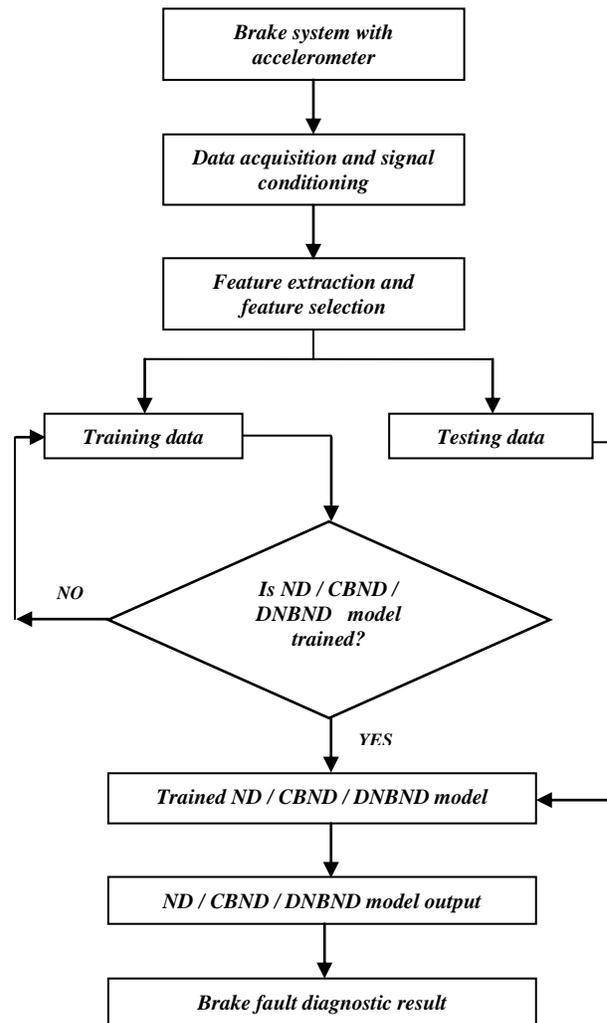


Figure 1. Flow chart of brake fault diagnosis using ND / CBND / DNBND algorithm

Nested dichotomy (ND) is a method for dealing with multiclass classification problems using binary classifiers. A nested dichotomy organizes the classes in a tree; each internal node has a binary classifier. A set of classes can be

organized in different ways in a nested dichotomy. An END is formed by several nested dichotomies. This paper studies the use of the method, in conjunction with ensembles of decision trees (random forests). Empirical experiments show that this approach yields accurate multiclass classifiers and is able to improve predictive performance even in the case of classifiers, such as decision trees, that can deal with multiclass problems directly. However, classification of faults in hydraulic brake system using nested dichotomy classifiers has not been studied. Hence an attempt has been made, in the present study to classify the faults in hydraulic brake system using ensembles of ND classifier. The flow chart of the fault diagnostic system is shown in Figure 1.

Contributions in the present work are as follows:

1. Different fault conditions were simulated. Vibration signals for different simulated fault conditions were acquired
2. Relevant statistical features were extracted from the vibration signal.
3. The required number of features was selected to get good classification accuracy.
4. Feature classification using different types of Nested dichotomy classifier.

2. EXPERIMENTAL STUDIES

The main objective of the study is to monitor the condition of the brake system. This paper focuses on the use of ND classifier for fault classification of the hydraulic brake set up. Referring to Figure 2, the hydraulic brake system test rig and data acquisition system is discussed in the following topics under experimental setup and experimental procedure, respectively.

2.1. Experimental set up

A commercial passenger car's (Maruti Swift) hydraulic brake system (Figuer. 2) was used to fabricate the brake test setup. The test setup consists of disc and rear drum brake coupled together by a shaft. The shaft is in turn run by a DC motor (1HP) coupled to a belt drive system. A lever is placed at the top of the motor which is connected to the accelerator pedal providing variable speeds up to 2500rpm. The brake pedal is provided in the left side of the accelerator pedal. It is attached to the piston in the master cylinder 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 the real world, branded vehicles (cars) parts were considered. The dimension of test rig is 80cm x 80cm x 40cm.

Vibration signal was acquired using piezoelectric transducers (accelerometer). Accelerometers have the large frequency response and it can detect very small vibrations

without being damaged by large vibrations; output is proportional to forces which are the cause of internal damage, and high-frequency sensitivity for detecting faults. Due to these reasons the accelerometers are widely used for fault classification. An uni-axial accelerometer (50g range, 100mV/g sensitivity, and resonant frequency 40 Hz) was mounted on the drive shaft cover near the brake drum (and brake disc) using a direct adhesive mounting technique. It was connected to the DAQ (Model NI USB 4432) system through a cable. The card has 5 analog input channels with a sampling rate and resolution of 102.4 kilo samples per second and 24-bit respectively. The accelerometer is connected to the signal-conditioning unit, where the signal goes through the charge amplifier and an analogue-to digital converter (ADC). The vibration signal in digital form is input to the computer through an USB port. NI – Lab VIEW was used to interface between transducer signal and the system (PC).

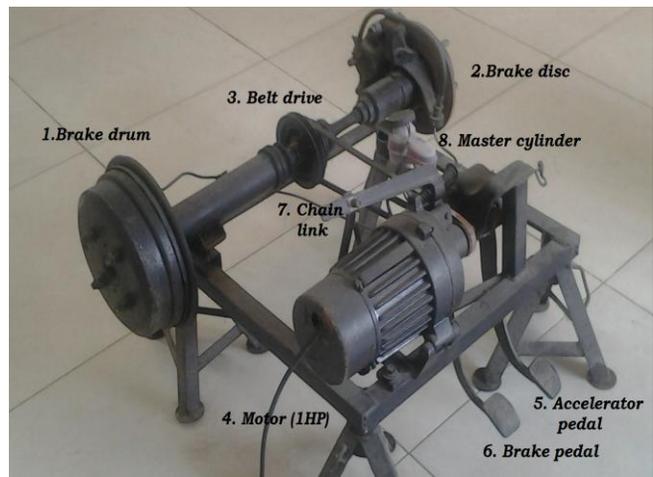


Figure 2. Experimental setup – Brake fault diagnosis

2.2. Experimental Procedure

Initially the test rig was assumed to be in good condition. (All components were brand new). The vibration signals were measured from the hydraulic brake system working under braking conditions. (Original Speed 667 RPM, Brake load 68.9 N). The vibration signal from an accelerometer mounted on the brake shaft was taken with the following settings:

1. Sample length: The sample length was chosen arbitrarily as 1024.
2. Sampling frequency: As per Nyquist sampling theorem, the sampling frequency should be at least twice the highest frequency contained in the signal. By using this theorem sampling frequency was calculated as 24 kHz.
3. Number of samples: Minimum of 55 trials was taken for each condition of the hydraulic braking system, and vibration signals were stored in the data files.

Data acquisition is the process of sampling signals that measure real world physical conditions and converting the resulting samples into digital numeric values that can be manipulated by a computer. Data Acquisition Card (DAC) hardware is used here to interface between the sensor signal and a Personal computer (PC). The faults were simulated one at a time while all other components remain in good condition and the corresponding vibration signals were acquired. The simulated faults are as follows. air in the brake fluid (AE) , brake oil spill on disc brake (BO), drum brake pad wear (DRPW), disc brake pad wear (even) – inner (DPWI), disc brake pad wear (even) – inner and outer (DPWIO), disc brake pad wear (uneven) (UDPWI) – inner, disc brake pad wear (uneven) – inner and outer (UDPWIO), reservoir leak (RL), drum brake mechanical fade (DRMF) (Jegadeeshwaran & Sugumaran, 2013). Once the faults were simulated, the vibration signals were recorded and feature extraction and feature selection were carried out using these vibration signals.

3. FEATURE EXTRACTION AND FEATURE SELECTION

In pattern recognition feature extraction is a special form of dimensionality reduction. When the input data is expected to be large, then the input data will be transformed into a reduced representation set of features. Transforming the input data into the set of features is called feature extraction. Feature extraction involves simplifying the amount of resources required to describe a large set of data accurately. Feature extraction is a general term for methods of constructing combinations of the variables to get around these problems while still describing the data with sufficient accuracy. The definition and process of extracting statistical features were described for brake by Reference (Jegadeeshwaran & Sugumaran 2013). The following set of statistical parameters, namely, minimum, skewness, mean, standard error, maximum, sample variance, standard deviation, count, kurtosis, mode, sum and median were extracted as the basic features for the study. In machine learning, feature selection (attribute selection) is the process of selecting a subset of relevant features that are used in model construction. Feature selection is carried out when the data set contains many redundant or irrelevant features. Redundant features do not provide any information, and irrelevant features provide no useful information in any context. The best number of features was selected using attribute evaluator (Jegadeeshwaran & Sugumaran, 2015). The classification accuracy with selected number of features were given in Tale 1. Referring Table 1, the number of features required for classification, was found.

4. FEATURE CLASSIFICATION

Next step after feature selection is feature classification. The selected features are to be classified using the Nested Dichotomy classifier.

4.1. Nested dichotomy (ND)

A simple way to improve the classification accuracy for problems with a small number of classes is to cache two-class models and re-use them in different members of an ensemble of nested dichotomies. A system of nested dichotomies is a statistical model that is used to decompose a multi-class problem into multiple two-class problems. The decomposition can be represented as a binary tree. Each node of the tree stores a set of class labels, the corresponding training data and a binary classifier. At the very beginning, the root node contains the whole set of the original class labels corresponding to the multi-class classification problem. This set is then split into two subsets. These two subsets of class labels are treated as two “meta” classes and a binary classifier is learned for predicting them (Dong, Frank & Kramer 2005).

The training dataset is split into two subsets corresponding to the two meta classes and one subset of training data is regarded as the positive examples while the other subset of testing data is regarded as the negative examples. The two successor nodes of the root inherit the two subsets of the original class labels with their corresponding training datasets and a tree is built by applying this process recursively. The process finally reaches a leaf node if the node contains only one class label. It is obvious that for any given c -class problem, the tree contains c leaf nodes (one for each class) and $(c-1)$ internal nodes. Each internal node contains a binary classifier. A nice feature of using a system of nested dichotomies for multi-class problems is that it yields class probability estimates in a straightforward fashion Frank and Kramer (2004) sampled randomly from the space of all possible trees by considering each tree with equal probability. The selection of the tree structure will influence the classification results. It makes sense to use all possible nested dichotomies for a given problem and average their probability estimates to yield accurate predictions. For a c -class problem, the number of possible systems of nested dichotomies is $(2c - 3)!!$ (Frank & Kramer, 2004). Hence, using all possible nested dichotomies is infeasible.

In the absence of prior knowledge, any sampling scheme that does not give preferential treatment to a particular class can be considered for a suitable candidate. The problem with random sampling based on a uniform distribution over trees is that the tree depth is only limited by the number of classes, and deep trees can take a long time to build (Dong, Frank & Kramer 2005).

4.2. Class-balanced nested dichotomies (CBND)

The Class-balanced nested dichotomy (CBND) method is based on balancing the number of classes at each node. Instead of sampling from the space of all possible trees (as in nested dichotomy), it is sampled from the space of all balanced trees. The advantage of this method is that the

depth of the tree is guaranteed to be logarithmic in the number of classes. The number of possible class-balanced nested dichotomies is obviously smaller than the total number of nested dichotomies. The following recurrence relation defines the number of possible class-balanced trees:

$$T(c) = \begin{cases} \frac{1}{2} \binom{c}{c/2} T(c/2) T(c/2) & : \text{if } c \text{ is even} \\ \binom{c}{(c+1)/2} T\left(\frac{c+1}{2}\right) T\left(\frac{c-1}{2}\right) & : \text{if } c \text{ is odd} \end{cases} \quad (1)$$

Table 1 shows the number of possible systems of nested dichotomies for up to 12 classes for the class-balanced (CBND) and the unconstrained case (ND). At each node the set of classes is split into equal size subsets (of course, if the number of classes is odd, the size will not be exactly equal), and the base learning algorithm is applied to the data corresponding to these two subsets. The algorithm then recurses until only one class is left. It is applied repeatedly with different random number seeds to generate a committee of trees. It shows that a non-trivial number of CBNDs can be generated for classification problems with five or more classes. There is a further opportunity to improve the training time for ensembles of nested dichotomies. This is an added advantages of the class balanced nested dichotomy.

There is a drawback with the class balanced approach. Some multi-class problems are very unbalanced and some classes are much more populous than others. In that case a class-balanced tree does not imply that it is also data balanced. This can negatively affect runtime if the base learning algorithm has time complexity worse than linear in the number of instances. The pseudo code for the class balanced nested dichotomy algorithm has been given below. (Dong, Frank & Kramer, 2005).

```

if |C| = 1 then return
    P = subset of C, randomly chosen from all subsets
    of size [C/2]
N = C \ P
Dp = all instances in D apart from those pertaining to classes
in P
    buildClassBalancedNestedDichotomies (Dp, P)
Dn = all instances in D apart from those pertaining to classes
in N
    buildClassBalancedNestedDichotomies (Dn, N)
D = a two-class version of D created based on N and P
classifierForNode = buildClassifier (D')

```

4.3. Data Near Balanced Nested Dichotomy (DNBND)

A simple algorithm called data near balanced nested dichotomy (NDBND) can be used as a alternate for the class balanced nested dichotomy. Since this method violates the condition that the sampling scheme should not be biased towards a particular class. It randomly assigns classes to two subsets until the size of the training data in one of the subsets exceeds half the total amount of training data at the

node. It is very essential to maintain a degree of randomness in the assignment of classes to subsets in order to preserve diversity in the committee of randomly generated systems of nested dichotomies. In the case of a skewed class distribution the base nested dichotomy algorithm's runtime is worse than linear. In that case, the number of instances is divided as evenly as possible at each node, so as to reduce the maximum amount of data considered at a node as quickly as possible. The pseudo code for the data near balanced nested dichotomy algorithm has been given below (Dong, Frank and Kramer 2005).

```

if |C| = 1 then return
    C = random permutation of C
    Dp = ∅, Dn = ∅
do
    if (|C| > 1) then
add all instances from D pertaining to first class in C to Dp
add all instances from D pertaining to last class in C to Dn
    remove first and last class from C
else
add all instances from D pertaining to remaining class in C
to Dp
    remove remaining class from C
while (|Dp| < [D/2]) and (|Dn| < [D/2])
if ((|Dp| ≥ [D/2]) then
add instances from D pertaining to remaining classes in C to
Dn
else
add instances from D pertaining to remaining classes in C to
Dp
    P = all classes present in Dp,
    N = all classes present in Dn
buildDataBalancedNestedDichotomies(Dp, P)
buildDataBalancedNestedDichotomies(Dn, N)
    D' = a two-class version of D created based on N
    and P
classifierForNode = classifier learned by base learner from
D'

```

5. RESULTS AND DISCUSSION

The experiment was carried on the brake setup for the different simulated fault conditions. Machine learning approach was used with statistical feature for fault classification. The obtained results are discussed below.

5.1. Effect of number of features on classification accuracy

Twelve statistical features, namely, mean, standard error, median, standard deviation, variance, kurtosis, skewness, range, minimum, maximum, sum and count were extracted from the acquired vibration signal. The effect of number of features on classification accuracy was found and is given in Table 1. It shows that when the number of features is 8 in each class, the classifier gives good accuracy. In the present study, minimal computation time strategy was used because

the on board processors on vehicle have limited computational resources.

No. of features	Classification accuracy (%)		
	ND ^a	CBND ^b	DNBND ^c
1	44.55	43.94	43.27
2	90.00	88.91	89.45
3	89.45	90.00	90.18
4	96.18	95.64	95.82
5	97.27	96.73	96.55
6	97.45	97.82	97.64
7	97.45	98.36	98.18
8	98.00	98.91	98.73
9	98.00	97.82	98.18
10	97.82	97.27	97.63
11	98.36	98.73	98
12	98.00	97.82	97.82

^aGND = Nested dichotomy, ^bCBND = Class balanced nested dichotomy, ^cDNBND = Data near balanced nested dichotomy

Table 1. Effect of number of features on classification accuracy

5.2. Statistical features using Nested Dichotomy (ND) algorithm

Referring Table 1, among the twelve extracted features, top eight features were selected for classification.

The general procedure for reading and understanding the confusion matrix is as follows. It looks in the form of a square matrix. Referring to Table 2, the first element in the first row represents the total number of data points corresponding to air in the brake fluid (AE). In the first row, out of 55 data sets, all are correctly classified as (AE). There is no misclassification. Hence all other elements in the first row are zero. In second row, second element represents the total number of data points related to brake oil spill (BO) condition. The total in the second row elements (out of 55 data sets), 54 data sets was correctly classified as brake oil spill (BO). One data set was misclassified as disc brake pad wear uneven inner & outer (UDPWIO).

As discussed above, misclassification details of classifier with the statistical features can be illustrated in a better way using the confusion matrix. From the confusion matrix, one can understand that 55 samples were considered for each condition of the brake system. All the diagonal elements of the confusion matrix represent the number of correctly

classified data points and the non-diagonal elements represent the incorrectly classified data points. In this fashion, the classification accuracies were found and compared.

Total number of instances	550
Correctly classified instances	539 (98 %)
Incorrectly classified instances	11 (2 %)
Kappa statistic	0.9778
Mean absolute error	0.0075
Root mean squared error	0.0547

Category	1	2	3	4	5	6	7	8	9	10
1	55	0	0	0	0	0	0	0	0	0
2	0	54	0	0	0	1	0	0	0	0
3	0	0	52	0	1	1	0	0	1	0
4	0	0	0	55	0	0	0	0	0	0
5	0	0	0	0	55	0	0	0	0	0
6	0	0	0	0	0	55	0	0	0	0
7	0	0	0	0	0	1	53	0	0	1
8	1	0	0	0	0	1	1	52	0	0
9	0	0	0	0	0	0	0	0	55	0
10	0	0	0	0	0	0	2	0	0	53

1 = AE : Air in brake fluid; 2 = BO : Brake oil spill; 3 = DPWI : Disc brake pad wear – Inner; 4 = DPWIO : Disc brake pad wear Inner & outer; 5 = UDPWI : Uneven disc pad wear (Inner); 6 = UDPWIO : Uneven disc pad wear (Inner & Outer); 7 = DRMF : Drum brake mechanical fade; 8 = DRPW : Drum brake pad wear; 9 = GOOD : Brake without any fault; 10 = RL : Reservoir leak.

Table 2. Confusion matrix for ND classifier

In this case, none of the good condition data points have been misclassified and hence the fault detection accuracy is close to 100%. None of the fault conditions were also misclassified as a GOOD condition. However, there were some misclassification amongst the fault conditions and hence classification accuracy for decision tree was found to be 98 %.

Table 3 shows the detailed accuracy by class of nested dichotomy classifier. True positive rate (TP rate) should be ideally 1. TP (True positive) rate means the number of items correctly labeled as belonging to the positive class. FP (false positive) is a result that indicates a given condition has been fulfilled, when it actually has not been fulfilled.

In the first row, TP rate for the condition “AE” is 1 and FP rate is 0.002; precision which means the fraction of retrieved instances that are relevant is 0.981. One data point relevant to DRPW has been misclassified as AE. Hence the precision value is less. The recall (fraction of relevant instances that are retrieved) is 1. Since none of the AE data points were misclassified as other fault condition. Both recall and precision value should be ideal. As discussed above the detailed accuracy by class were calculated for all

classes. Twelve statistical features namely, mean, standard error, median, standard deviation, variance, kurtosis, skewness, range, minimum, maximum and sum, count were extracted from the acquired vibration signal.

<i>TP Rate</i>	<i>FP Rate</i>	<i>Precision</i>	<i>Recall</i>	<i>F-Measure</i>	<i>ROC Area</i>	<i>Class</i>
1	0.002	0.982	1	0.991	1	1
0.982	0	1	0.982	0.991	1	2
0.945	0	1	0.945	0.972	1	3
1	0	1	1	1	1	4
1	0.002	0.982	1	0.991	1	5
1	0.008	0.932	1	0.965	1	6
0.964	0.006	0.946	0.964	0.955	0.999	7
0.945	0	1	0.945	0.972	1	8
1		0.002	0.982	1	0.991	9
0.964	0.002	0.981	0.964	0.972	1	10

1 = AE : Air in brake fluid; 2 = BO : Brake oil spill; 3 = DPWI : Disc brake pad wear – Inner; 4 = DPWIO : Disc brake pad wear Inner & outer; 5 = UDPWI : Uneven disc pad wear (Inner); 6 = UDPWIO : Uneven disc pad wear (Inner & Outer); 7 = DRMF : Drum brake mechanical fade; 8 = DRPW : Drum brake pad wear; 9 = GOOD : Brake without any fault;10 = RL : Reservoir leak.

Table 3. Detailed accuracy by class for ND Classifier

5.3. Statistical features using Class balanced nested dichotomy (CBND) algorithm

The effect of number of features on classification accuracy classification accuracy is given in Table 1. The classifier gives maximum accuracy when the number of features is equal to 8. Referring Table 1, the classification accuracy of class balanced nested dichotomy classifier algorithm was calculated as 98.91 %. The misclassification details are presented in Table 4.

<i>Category</i>	<i>1</i>	<i>2</i>	<i>3</i>	<i>4</i>	<i>5</i>	<i>6</i>	<i>7</i>	<i>8</i>	<i>9</i>	<i>10</i>
1	55	0	0	0	0	0	0	0	0	0
2	0	55	0	0	0	0	0	0	0	0
3	0	0	55	0	0	0	0	0	0	0
4	0	0	0	55	0	0	0	0	0	0
5	0	0	0	0	55	0	0	0	0	0
6	0	0	0	0	0	55	0	0	0	0
7	0	1	0	0	0	0	53	0	0	1
8	1	0	0	0	0	0	0	53	0	1
9	0	0	0	0	0	0	0	0	55	0
10	0	0	0	0	0	0	2	0	0	53

Table 4. Confusion matrix for CBND

<i>TP Rate</i>	<i>FP Rate</i>	<i>Precision</i>	<i>Recall</i>	<i>F-Measure</i>	<i>ROC Area</i>	<i>Class</i>
1	0.002	0.982	1	0.991	0.999	1
1	0.002	0.982	1	0.991	1	2
1	0	1	1	1	1	3
1	0	1	1	1	1	4
1	0	1	1	1	1	5
1	0	1	1	1	1	6
0.964	0.004	0.964	0.964	0.999	0.99	7
0.964	0	1	0.964	0.981	0.99	8
1	0	1	1	1	1	9
0.964	0.004	0.964	0.964	0.964	1	10

1 = AE : Air in brake fluid; 2 = BO : Brake oil spill; 3 = DPWI : Disc brake pad wear – Inner; 4 = DPWIO : Disc brake pad wear Inner & outer; 5 = UDPWI : Uneven disc pad wear (Inner); 6 = UDPWIO : Uneven disc pad wear (Inner & Outer); 7 = DRMF : Drum brake mechanical fade; 8 = DRPW : Drum brake pad wear; 9 = GOOD : Brake without any fault;10 = RL : Reservoir leak.

Table 5. Detailed accuracy by class for CBND

None of the faulty conditions in CBND, was misclassified as ‘GOOD’. Hence the overall classification accuracy is 98.91 %, which is higher than ND, and hence CBND stands one step forward. Table 5. shows the detailed accuracy by class for CBND classifier. The TP rate, precision and F-measure values for the data sets belongs to DPWI, DPWIO, DRMF, DRPW and GOOD condition, has the unit value 1. Hence it is having maximum classification accuracy.

5.4. Feature classification using Data Near balanced nested dichotomy (DNBND) algorithm

The effect of number of features on classification accuracy was found using Data near balanced nested dichotomy (DNBND) algorithm is given in Table 1. In the present study, eight statistical features, namely, minimum value, standard error, sample variance, kurtosis, skewness, mean, median and mode were used for classification to reduce the computation time. The Data near balanced nested dichotomy (DNBND) algorithm was trained using a selected number of statistical features of vibration signals and the classification results are tabulated as a confusion matrix shown in table 6.

Total number of instances	550
Correctly classified instances	543 (98.73 %)
Incorrectly classified instances	6 (1.27 %)
Kappa statistic	0.9859
Mean absolute error	0.0075
Root mean squared error	0.0538

The Data near balanced nested dichotomy (DNBND) algorithm was trained using selected number of statistical features of vibration signals and the classification results are tabulated Referring the above discussion, the maximum

classification accuracy for DNBND algorithm for selected number of features was calculated as 98.73 %. As in CBND algorithm, none of the ‘GOOD’ condition data points were misclassified as other fault conditions; and none of the fault condition data points were misclassified as ‘GOOD’ condition. The overall classification accuracy was found to be 98.73 %. Table 7. Shows the detailed accuracy by class for DNBND classifier. Except few conditions, all other fault conditions are having a small misclassification detail. Compared to CBND, the DNBND has less classification accuracy.

Category	1	2	3	4	5	6	7	8	9	10
1	55	0	0	0	0	0	0	0	0	0
2	0	55	0	0	0	0	0	0	0	0
3	0	0	55	0	0	0	0	0	0	0
4	0	0	0	55	0	0	0	0	0	0
5	0	0	0	0	55	0	0	0	0	0
6	0	0	0	0	0	55	0	0	0	0
7	0	1	0	0	0	0	52	0	0	1
8	1	0	0	0	0	0	0	53	0	1
9	0	0	0	0	0	0	0	0	55	0
10	0	0	0	0	0	0	2	0	0	53

Table 6: Confusion matrix for DNBND

TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
1	0.002	0.982	1	0.991	1	1
1	0.002	0.982	1	0.991	1	2
1	0.002	0.982	1	0.991	1	3
1	0	1	1	1	1	4
0.982	0	1	0.982	0.991	1	5
1	0	1	1	1	1	6
0.945	0.004	0.963	0.945	0.954	0.99	7
0.964	0	1	0.964	0.981	0.991	8
1	0	1	1	1	1	9
0.982	0.004	0.964	0.982	0.973	0.999	10

1 = AE : Air in brake fluid; 2 = BO : Brake oil spill; 3 = DPWI : Disc brake pad wear – Inner; 4 = DPWIO : Disc brake pad wear Inner & outer; 5 = UDPWI : Uneven disc pad wear (Inner); 6 = UDPWIO : Uneven disc pad wear (Inner & Outer); 7 = DRMF : Drum brake mechanical fade; 8 = DRPW : Drum brake pad wear; 9 = GOOD : Brake without any fault; 10 = RL : Reservoir leak.

Table 7. Detailed accuracy by class for DNBND

5.5. Comparative Study

Table 1 shows the estimated accuracy for NDs, CBNDs, and DNBNDs on the given datasets. It is evident that there is no dataset with a significant difference in accuracy for DNBNDs and CBNDs. This is the desired outcome. For

NDs, there is some dataset where the accuracy is significantly reduced compared to CBNDs. This is due to the diversity of the ensembles on the data set. The diversity of ensembles on this dataset can be measured using the kappa statistic. This statistic can be used to measure agreement between pairs of ensemble members (Frank & Kramer, 2004). For CBNDs, the mean kappa value over all pairs, measured on the training data, was 0.9879, which was indeed higher than the mean kappa values for DNBNDs and NDs (0.9859 and 0.9776 respectively). This indicates that reduction in diversity is the reason for the drop in performance. The mean absolute error (MAE) measures the average magnitude of the errors in a set of forecasts, without considering their direction. It measures accuracy for continuous variables. The MAE is a linear score which means that all the individual differences are weighted equally in the average. The root mean squared value (RMSV) is a quadratic scoring rule which measures the average magnitude of the error. The difference between forecast and corresponding observed values are each squared and then averaged over the sample. Finally, the square root of the average is taken. Since the errors are squared before they are averaged, the RMSE gives a relatively high weight to large errors. This means the RMSE is most useful when large errors are particularly undesirable.

The MAE and the RMSE can be used together to diagnose the variation in the errors in a set of forecasts. The RMSE will always be larger or equal to the MAE; the greater difference between them, the greater the variance in the individual errors in the sample. If the RMSE=MAE, then all the errors are of the same magnitude. Both the MAE and RMSE can range from 0 to ∞. They are negatively-oriented scores: Lower values are better. MAE values for all the three types of NDs are approximately near to zero (0.007). When compared to ND (0.0075) and DNBND (0.0075), the CBND algorithm is having less value (0.007). Hence it gives maximum accuracy compared to other two. Fig. 3 show the comparative results between different types of Nested dichotomy algorithms.

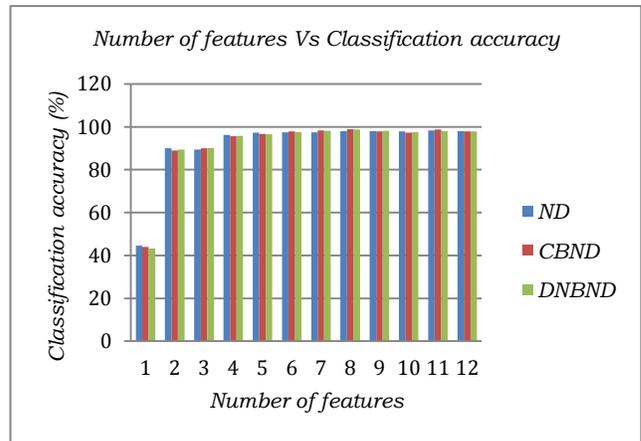


Figure 3. Number of features Vs Classification accuracy

The ND, CBND, DNBND algorithm results have been compared with other machine learning classifiers, namely, decision tree and support vector machine (SVM). Table 8 shows the overall classification accuracy of the different machine learning algorithms such as decision tree, support vector machine (Jegadeeshwaran & Sugumaran 2015), nested dichotomy algorithms. It indicates that the class balanced nested dichotomy (CBND) algorithm gives the maximum classification accuracy for the problem concerned.

S. No.	Name of the Classifier	Classification Accuracy (%)
1	Decision tree	97.45
2	Support Vector Machine	98.36
3	Nested dichotomy	98.00
4	Class balanced nested dichotomy	98.91
5	Data near balanced nested dichotomy	98.73

Table 8. Comparative study

6. CONCLUSION

Nested dichotomy algorithm has recently been shown to be a very promising meta learning scheme for multi-class problems. They yield class probability estimates in a natural way and produce accurate classification result. In this paper a meta learning scheme was presented to improve the classification accuracy in brake fault diagnosis. It deals with vibration based fault diagnosis of automobile hydraulic brake system. Nine classical fault conditions were simulated and were tested on a hydraulic brake setup. For each fault condition, the vibration signal was acquired using a piezo electric transducer. Set of statistical features were extracted from the vibration signal using feature extraction techniques. Feature selection was then carried out. The selected features were classified using ensemble algorithm techniques such as ND, CBND and DNBND. For all the above algorithms, random forest tree was taken as base algorithm. Referring Table 4, the vibration signals relevant to seven fault condition, namely, air in brake fluid, brake oil spill, disc brake pad wear – Inner, disc brake pad wear inner & outer, uneven disc pad wear (inner) and good condition have been correctly classified by the CBND model. In the remaining three classes, the misclassification accuracy is only 1.09 %. Moreover, all the misclassified data points were classified as other fault conditions and none of the fault condition data points have been misclassified as good condition. algorithms were compared. From the above results one can confidently say that statistical features with class balanced nested dichotomy (CBND) algorithm were found to be good candidate and it can be used for practical applications of fault diagnosis of hydraulic brake system. As

the results are encouraging one, future study can also be possible by conducting the experiment on real vehicle with real road conditions with the above feature – classifier combination.

NOMENCLATURE

DT	decision tree
ND	nested dichotomy
CBND	class balanced nested dichotomy
DNBND	Data near balanced nested dichotomy

REFERENCES

- Addin O. & Sapuan S.M. (2008). A Naïve-Bayes classifier for damage detection in engineering materials. *Materials and Design*, vol. 28, pp. 2379–2386.
- Burgess C. J. C. (1998). A tutorial on support vector machines for pattern recognition, *Data Mining and Knowledge Discovery*, vol. 2, pp. 955–974.
- Dong L. L, Frank E & Kramer S. (2005) Ensembles of Balanced Nested Dichotomies for Multi-Class Problems, *Knowledge Discovery in Databases: PKDD 2005*, 2005, pp. 84-95.
- Frank, E. & Kramer, S., Ensembles of nested dichotomies for multi-class problems. In: Proc. Int. Conf. on Machine Learning. ACM, 2004, pp.305 – 312.
- Huang C. H., (2009). Feature selection for text classification with Naïve Bayes, *Expert Systems with Applications*, vol. 36, pp. 5432–5435.
- Jack L. B. & Nandi A. K. (2000). Comparison of neural networks and support vector machines in condition monitoring application. In *Proceedings of COMADEM 2000*, Houston, TX, USA, pp. 721–730.
- Jegadeeshwaran R. & Sugumaran V., (2013). Method and Apparatus for Fault Diagnosis of Automobile Brake System Using Vibration Signals”, *Recent Patents on Signal Processing*, 3, 2013, 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, pp. 14-23.
- 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.
- Miller R. R., Marshall R. J., David Aexander Baiey & Nicholas Charles Griffin, (2004). Brake condition monitoring, US0011596 (A1).
- 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. pp. 1-7.

- Nowicki, R., Slowinski, R., & Stefanowski, J. (1992). Evaluation of vibro-acoustic diagnostic symptoms by means of the Rough Sets Theory. *in Computers in Industry*, vol. 20, pp. 141–152.
- Rajakarunakaran S., Venkumar P., Devaraj D., & Surya Prakasa Rao K., (2008). Artificial neural network approach for fault detection in rotary system, *Applied Soft Computing*, vol. 8, pp. 740–748.
- Reinecke, E., (1988). Apparatus for the measurement and / or of a braking torque, US4790606.
- Rodríguez, J. J., César García-Osorio & Jesús Maudes, Forests of nested dichotomies, *Pattern Recognition Letters*, 31, 2010, pp. 125–132.
- Sakthivel N.R., Indira V., Nair B.B., & Sugumaran V. (2011). 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)*, vol. 2, pp. 23–36.
- Sakthivel N.R., Sugumaran V., & Babudevasenapati S. (2010). Vibration based fault diagnosis of monoblock centrifugal pump using decision tree. *International Journal of Expert Systems with Application*, vol. 2, pp. 38–61.
- Sakthivel N.R., Sugumaran V., & Nair B. B. (2010). Comparison of Decision Tree-Fuzzy and Rough set-Fuzzy Methods for Fault Categorization of Mono-block Centrifugal Pump, *Mechanical Systems and Signal Processing*, vol. 24, pp. 1887–1906.
- 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.
- Saravanan N., & Cholairajan S. & Ramachandran K.I., (2009). Vibration based fault diagnosis of spur bevel gear box using fuzzy technique, *Expert Systems with Applications*, vol. 36, pp. 3119–3135.
- Shen Yin, Steven X. Ding, Adel Haghani, Haiyang Hao, & Ping Zhang. (2012). A comparison study of basic data-driven fault diagnosis and process monitoring methods on the benchmark - Tennessee Eastman process, *Journal of Process Control*, Volume 22, Issue 9, October 2012, Pages 1567-1581.
- Soman, K. P., & Ramachandran, K. I. (2005). *Insight into wavelets from theory to practice*. Prentice-Hall of India Private Limited, New Delhi, India.
- Sugumaran V., Muralidharan V., & Ramachandran K.I., (2007). Feature selection using Decision Tree and classification through Proximal Support Vector Machine for fault diagnostics of roller bearing. “Mechanical Systems and Signal Processing”, vol. 21, pp. 930–942.
- Sugumaran V., & Ramachandran K.I. (2007). Automatic rule learning using decision tree for fuzzy classifier in fault diagnosis of roller bearing. *Mechanical System and Signal Processing*, vol. 21, pp. 2237–2247.
- Sugumaran V., & Ramachandran K. I. (2011). Effect of number of features on classification of roller bearing faults using SVM and PSVM. *Expert Systems with Applications*. Vol. 38, pp. 4088–4096.
- Suykens, J. A. K., Van Gestel T., Vandewalle J., & De Moor B. (2003). A support vector machine formulation to PCA analysis and its Kernel version, *ESAT-SCD-SISTA Technical Report*.

BIOGRAPHIES



R. Jegadeeshwaran, was born in Erode city, Tamil Nadu, in 1981. The author completed his Under Graduate course in Mechanical Engineering at Institute of Road and Technology, Erode, Tamil Nadu, India, in the year 2002. He pursued his Masters degree in Mechatronics Engineering at Kongu Engineering College, Perundurai, Tamil Nadu, India, in 2009. At present he is pursuing his doctoral degree in VIT University, Chennai, Tamil Nadu.

He worked as an Assistant professor, at Kongu Engineering College, Perundurai for a period of two years. At present he is working as an assistant professor at VIT University, Chennai, Tamil Nadu, India. He has six international / National conference publications. He has published eight papers in peer reviewed international Journals. He is doing research in the field of fault diagnosis,



Dr. V. Sugumaran received the B.E. degree in Mechanical Engineering from the Amrita Institute of Technology & Science, 1998 and the M.Tech in Production Engineering, from The National Institute of Engineering, 2003. Ph.D. degree in Fault Diagnosis, from Amrita School of Engineering, Amrita University, Coimbatore, Tamil Nadu, India, in 2008.

From 2000 to 2009, he was an Associate Professor, with the Amrita School of Engineering, Coimbatore, India. From 2009 to 2011, he was working with SRM University, Chennai, Tamil Nadu. Since 2011, he has been an Associate Professor with the School of Mechanical and Building Sciences, VIT University, Chennai, Tamil Nadu, India. He is the author of one books, more than 88 international journal publications. He has also filed twelve patents. His research interests include Condition Monitoring & Fault Diagnosis, Machine learning / Data mining in manufacturing and Mechanical Engineering. He presented papers in 43 International / National Conferences. He is also acting as a Reviewer for many international journals and editor for four international journals.