

PAPER • OPEN ACCESS

Fault diagnosis of helical gearbox using acoustic signal and wavelets

To cite this article: SK Pranesh *et al* 2017 *IOP Conf. Ser.: Mater. Sci. Eng.* **197** 012079

View the [article online](#) for updates and enhancements.

Related content

- [Fault Diagnosis of Wind Turbine Gearbox Based on Wavelet Transform](#)
Hao Li and Changliang Liu
- [The Application of Radial Basis Function \(RBF\) Neural Network for Mechanical Fault Diagnosis of Gearbox](#)
Pengbo Wang
- [Condition Monitoring and Fault Diagnosis for Wind Turbine Gearbox Based on Waterfall](#)
Hongwei Xin, Xinjian Feng, Ye Xin et al.



IOP | ebooks™

Bringing you innovative digital publishing with leading voices to create your essential collection of books in STEM research.

Start exploring the collection - download the first chapter of every title for free.

Fault diagnosis of helical gearbox using acoustic signal and wavelets

Pranesh SK¹, Siju Abraham², V. Sugumaran³, M.Amarnath⁴

¹School of Electronics Engineering VIT University Chennai Campus
Chennai, Indiaskpran@gmail.com

^{2,3}School of Mechanical and building Sciences VIT University Chennai Campus
Chennai, India v_sugu@yahoo.com

⁴Assistant Professor, Dept. of Mechanical Engineering IITDM, Jabalpur
amaranth.cmy@gmail.com

Abstract: The efficient transmission of power in machines is needed and gears are an appropriate choice. Faults in gears result in loss of energy and money. The monitoring and fault diagnosis are done by analysis of the acoustic and vibrational signals which are generally considered to be unwanted by products. This study proposes the usage of machine learning algorithm for condition monitoring of a helical gearbox by using the sound signals produced by the gearbox. Artificial faults were created and subsequently signals were captured by a microphone. An extensive study using different wavelet transformations for feature extraction from the acoustic signals was done, followed by wavelet selection and feature selection using J48 decision tree and feature classification was performed using K star algorithm. Classification accuracy of 100% was obtained in the study

Keywords: K star algorithm, Daubechies wavelet, Acoustic signals, Helical gearbox, J48.

1. Introduction:

Machines are the source of production of all the vital components that is required. Gears are responsible for the supply and transmission of the power for these machines making them a very essential part. Gears have cut teeth or cogs which mesh with each other to transmit torque. The gear tooth takes all the load showing its significance. Each and every tooth of the gear contributes for the transmission of the load. It is essential to ensure the proper maintenance of the gear tooth through real time condition monitoring system as the failure of even a single tooth can result in tremendous damage which can result in decreasing the efficiency for transmission or disruption. Gear failures can be caused by several factors such as incorrect design, installation, acid corrosion, poor lubrication etc. The effects of a faulty gearbox can be catastrophic as it can damage the machinery since the broken



tooth may disrupt the rotatory motion of the gear. It is therefore essential to reduce the risk of gear failures. The situation can be monitored by examining the sound or vibration signals coming out of the box as the gears are enclosed inside the gearbox. The aim of this study is to reduce the risk of gear failures by real time monitoring of acoustic signals.

Acoustic signals were chosen for the experiment as the accelerometers are more expensive than microphones which would be used for recording the acoustic signals. Artificial fault conditions in gears were created for obtaining acoustic signals of different fault conditions. This was done so that the detection of faults could be done at early stages to avoid mishaps and thus resulting in better efficiency and transmission.

The processes involved were feature extraction, wavelet selection, feature selection and feature classification. Statistical features [1], histogram features [2] or wavelet features are the usually used features. FFT [5] (Fast Fourier Transform) was a very useful tool as the computation time for the Fourier Transform was reduced significantly. It was not suitable for the study as it was inconvenient for non-stationary signals. Though STFT [6] [18] (Short Time Fourier Transform) proved to be better than FFT for non-stationary signals, wavelet transform were preferred as it provided useful information for fault identification and also provides multiresolution as performed by M. Lang et al . Discrete wavelet transform [3] [4] was preferred over continuous wavelet transform as the processing of data took place faster in discrete wavelet transform despite the provision of accurate data. With further analysis Daubechies 5 was found to produce the best feature set which was used for selection and classification.

The most popular techniques for feature selection are artificial networks [7] [8], fuzzy [9], decision tree [10] [11] [19], principal component analysis and genetic algorithm. J48 decision tree [10] [11] [19] was used for wavelet selection and feature selection as it is very compact, simple and a fast computation speed.

Commonly used classifiers for feature selection are Support vector machine (SVM) [12] [20], Proximal support vector machine (PSVM), artificial neural network (ANN) [13], Naïve Bayes (NB) and Bayes Net, Fuzzy, Decision tree (DT) [14] [15] etc. The SVM based models suffers from higher training time and computational complexity when the number of patterns increases. Despite good results provided by artificial neural network classifier, the training of artificial neural network classifier is complex and time consuming. For a real time system, the classifier used should consume least time with high classification accuracy. J48 decision tree was thus used.

Artificial faults were created in the gears by chipping off 20%, 40%, 60% and 80% of the gear tooth and also by removing one entire gear tooth (100%) and another half of the next gear tooth (150%) all at full load capacity of the machine. Using J48 decision tree a set of wavelet features were extracted using 'db5' wavelet feature selection and classification. Simulations were performed and results were obtained.

2. Experimental studies

The experimental set up which is shown in Fig. 1 consists of a 5 HP two stage helical gearbox [17] [21] which is driven by 5.5 HP 3 phase induction motor at 1200 rpm. The DC generator is driven by the mechanical output from the gearbox. A load needs to be provided to the DC motor for the functioning. A resistor bank is connected to the DC motor such that the power can be dissipated in it. Resistor bank is preferred over dynamometers as additional torsional vibrations may occur due to torque fluctuations resulting in formations of inaccurate acoustic signals. Additional vibrations to the test rig are thus eliminated in the generator and resistor bank arrangement. The gearbox, generator and motor are mounted on stiff I-beams which are anchored to a big concrete block. To measure the vertical vibration signals which are generated on the bearing housing of the 16 teeth pinion, B&K 4332 accelerometer is used. Meshing gear frequencies are calculated as 320 Hz and its multiples. Data

sets were collected when the helical gear was working at normal conditions and also when artificial faults were created by chipping off 20%, 40%, 60%, 80%, 100% and also by removing one entire gear tooth and another half of the next gear tooth (150%). The signals were curtailed to 3 kHz when a low pass filter was used and was sampled at 8kHz. B&K TYPE 2626 charge amplifier was used to condition the microphone output.

A pinion is connected to a DC motor to generate 2kW power, which dissipates power in the resistor bank. Subsequently the actual load on the gearbox is only 2.6 HP which is 52% of its rated 5 HP. Utilization of the load varies from 50% to 100% in the industrial environment. Tyre couplings are installed between the electrical machines and gear box so that the backlash can be restricted to the gears. The experimental setup is shown in Fig. 1

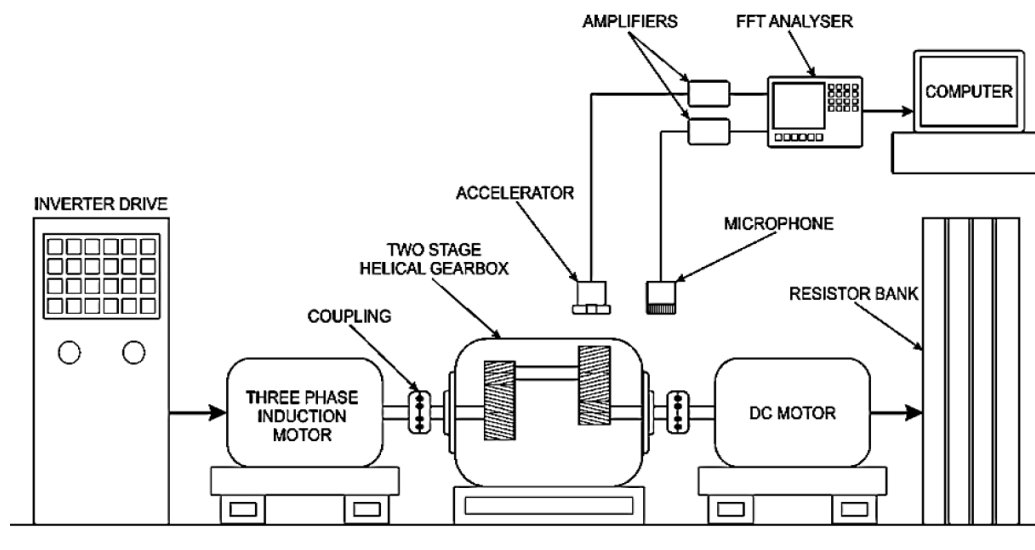


Figure:1 Experimental Setup

Table 1 Specifications of helical gearbox.

	First stage	Second stage
Number of teeth	43/13	73/16
Pitch circle diameter (mm)	198/65	202/48
Pressure angle	20	20
Helix angle	20	15
Modules	4.5/5	2.75/3
Speed of shafts	80 rpm (input)	1200 rpm (output)
Mesh frequency	59 Hz	320 Hz
Step-up ratio	15	
Power transmitted	5 HP	

3. Feature extraction

The acoustic signal recorded from the gearbox by the microphone was used to perform the fault diagnosis. The conversion of acoustic time-domain signal into time-frequency-domain data is done

using Discrete Wavelet Transform (DWT) through wavelet decomposition. The wavelet decomposition forms a trend. The obtained trend is again decomposed into next level trend and details. The same process is repeated for several levels of trends to give multiple levels of details. For this study, a signal length of 2048 (2^{11}) is chosen and subsequently the signals can be decomposed into 11 levels.

At each level, the detail co-efficients were used to compute energy content using the following formula.

$$V_i = \sum_{i=1}^n X_i^2$$

Where x_i = details coefficients

n = number of details coefficients

Then, the features were defined as the energy content at each level. The feature vector is defined as

$$V = (V_1, V_2, V_3, \dots, V_m)$$

When m – (number such that length of signal) = 2^m

$V_1, V_2, V_3 \dots$ are energy content at given level

The following discrete wavelet transformations were used in this study.

- (1) Coiflet : coif – 1, 2, 3, 4, 5
- (2) Reverse biorthogonal: rbio – 1.1, 1.3, 1.5, 2.2, 2.4, 2.6, 2.8, 3.1, 3.3, 3.5, 3.7, 3.9, 4.4, 5.5, 6.8
- (3) Daubechies wavelet : db – 1, 2, 3, 4, 5, 6, 7, 8, 9, 10
- (4) Discrete Meyer : dmey
- (5) Sym wavelet : sym – 2, 3, 4, 5, 6, 7, 8
- (6) Biorthogonal wavelet: bior – 1.1, 1.3, 1.5, 2.2, 2.4, 2.6, 2.8, 3.1, 3.3, 3.5, 3.7, 3.9, 4.4, 5.5, 6.8
- (7) Haar

4. Wavelet selection

Processing of the time domain signals was done using 54 different discrete wavelet transforms from the seven wavelet families. The features which were extracted from the wavelet transform are passed to the J48 algorithm to find the maximum classification accuracy. The features extracted using Daubechies 5 gave the maximum classification accuracy among all the DWTs mentioned above. Daubechies 5 was thus selected for subsequent operations. Daubechies wavelet which is represented as 'db n' is a family of orthogonal wavelets which are characterised by highest number of vanishing points(n) for a given support width of 2^{n-1} . Of all these possible solutions for the point and orthogonality conditions, the solution whose scaling filter produces the maximum phase is selected.

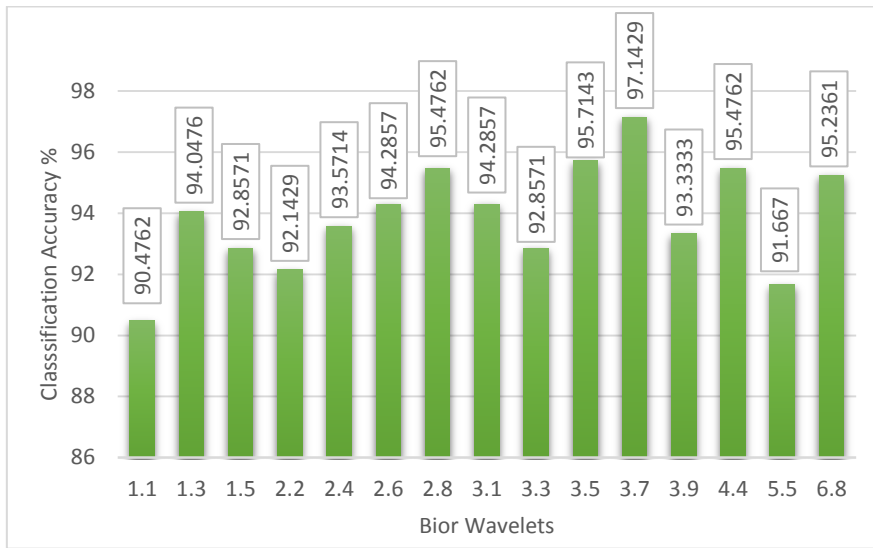


Figure: 2 Classification accuracy of bior wavelets

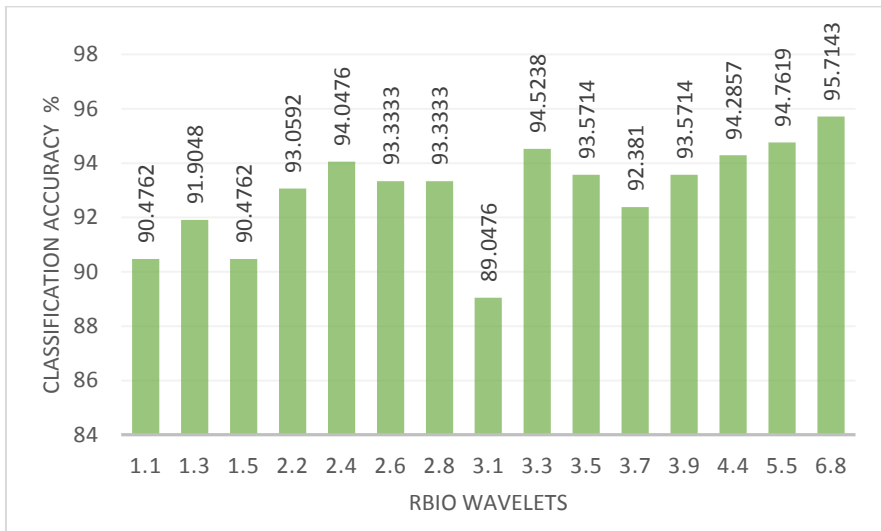


Figure:3 Classification accuracy of rbio wavelets

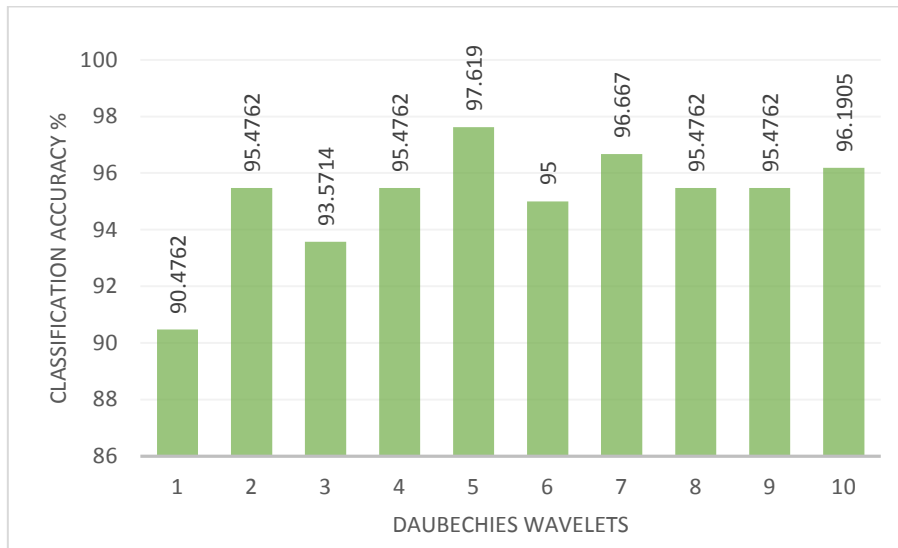


Figure:4 Classification accuracy of Daubechies wavelets

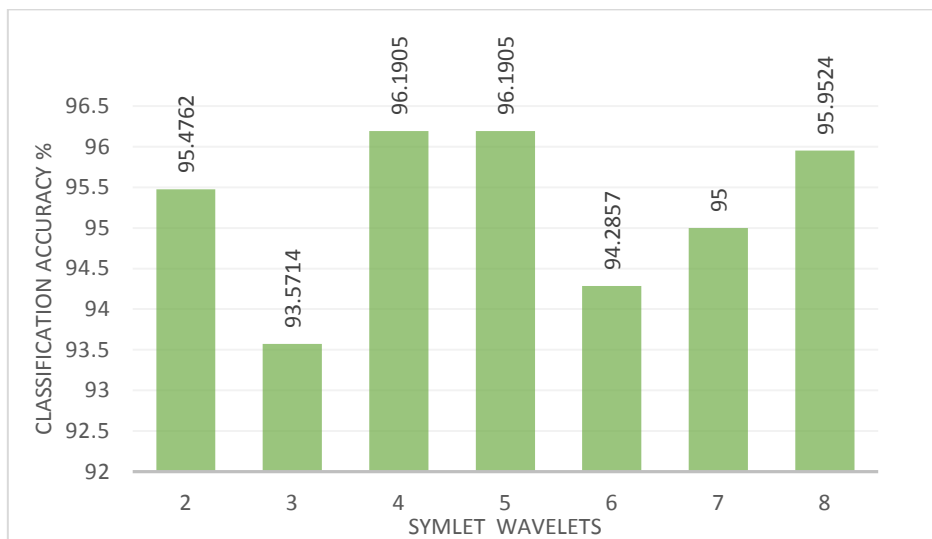


Figure:5 Classification accuracy of Symlet wavelets



Figure:6 Classification accuracy of coiflet wavelets

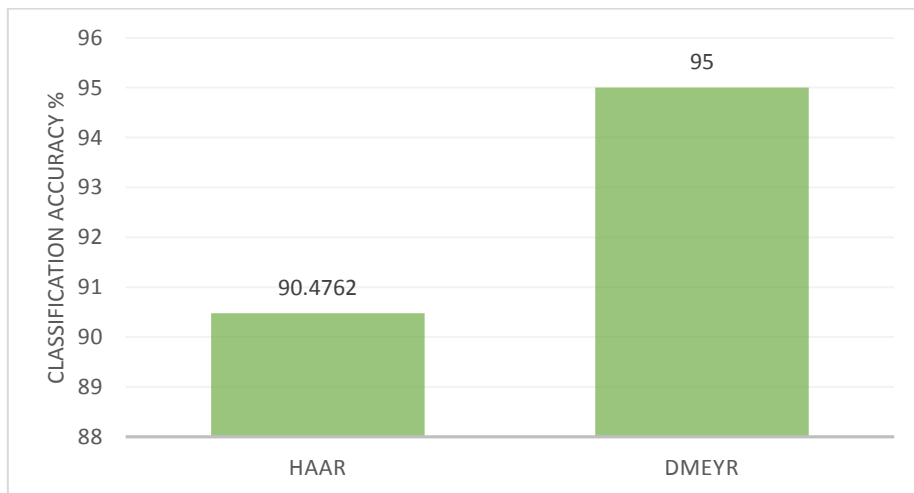


Figure:7 Classification accuracy of Haar and dmeyr wavelets

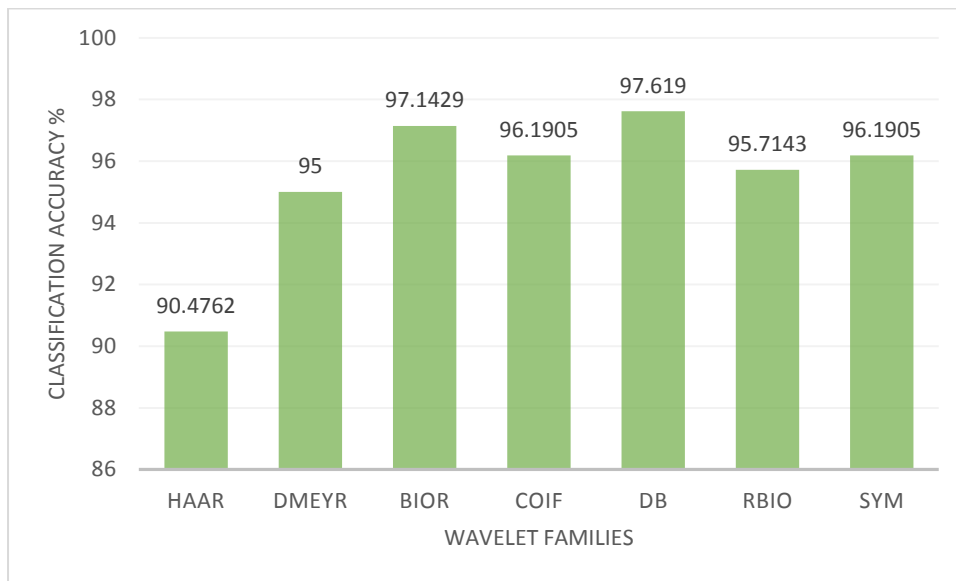


Figure:8 Classification accuracy of various wavelet families

5. Feature selection

J48 classifier is a simple decision tree used for classification. It creates a binary tree. The decision tree approach is most useful in classification problem. With this technique, a tree is constructed to model the classification process. Once the tree is built, it is applied to each tuple in the database and results in classification for that tuple.

While building a tree, J48 ignores the missing values i.e., the value for that item can be predicted based on what is known about the attribute values for the other records. The basic idea is to divide the data into range based on the attribute values for that item that are found in the training sample. J48 allows classification via either decision trees or rules generated from them.

J48 uses entropy based information gain as the selection criteria. As per information theory, entropy is a measure of the uncertainty in a random variable. The expected reduction in entropy due to the partitioning of the examples according to the given feature gives the information gain. It is a measure of the capability of a given attribute to separate its training examples according to the target function.

Information gain (S, A) of a feature A relative to a collection of examples S , is defined as $Gain(S, A) = Entropy(S) - \sum_{v \in Values(A)} \frac{|S_v|}{|S|} Entropy(S_v)$ (1)

Where, $S_v = (\{s \in S | A(s) = v\})$

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

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

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

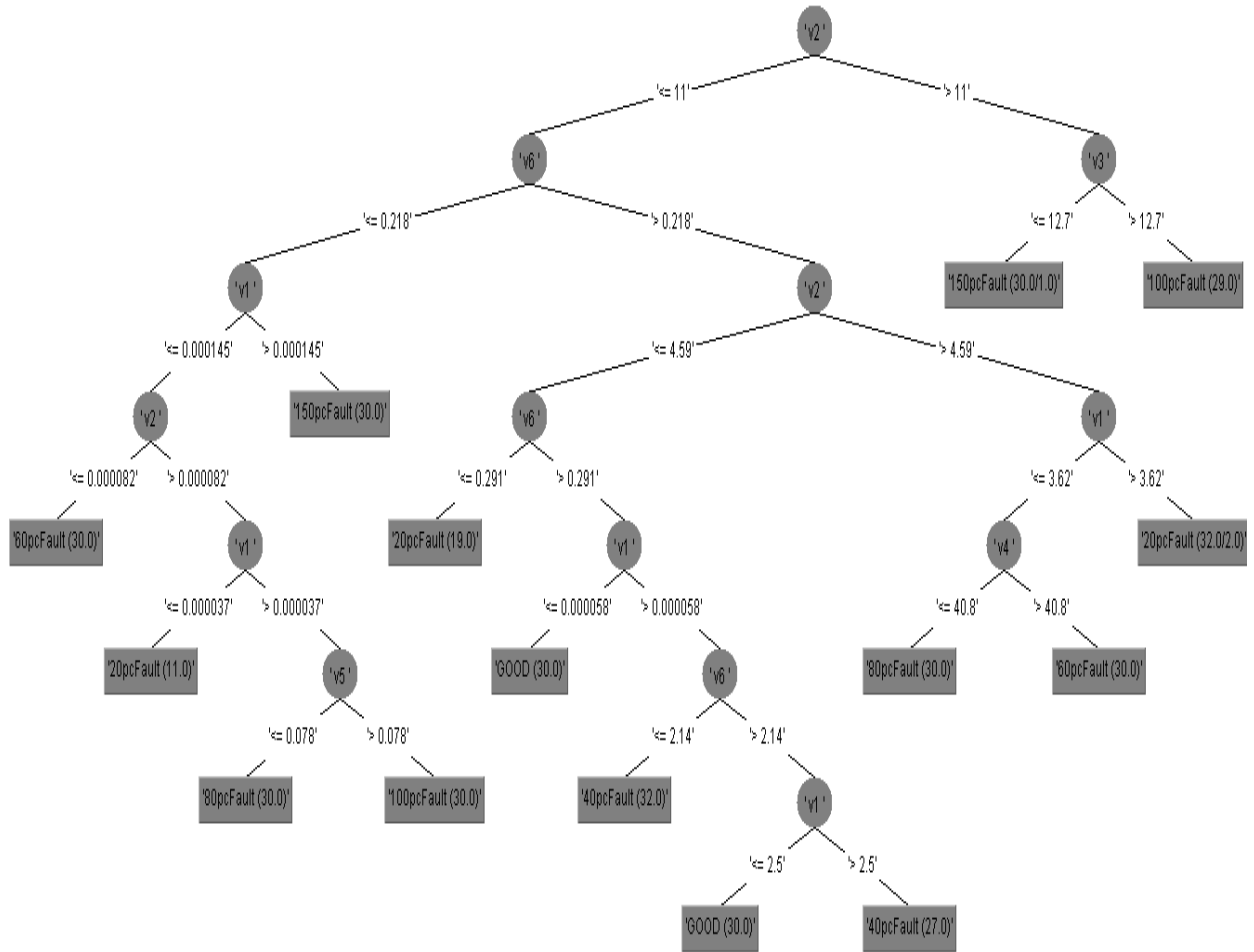


Figure:9 Feature selection

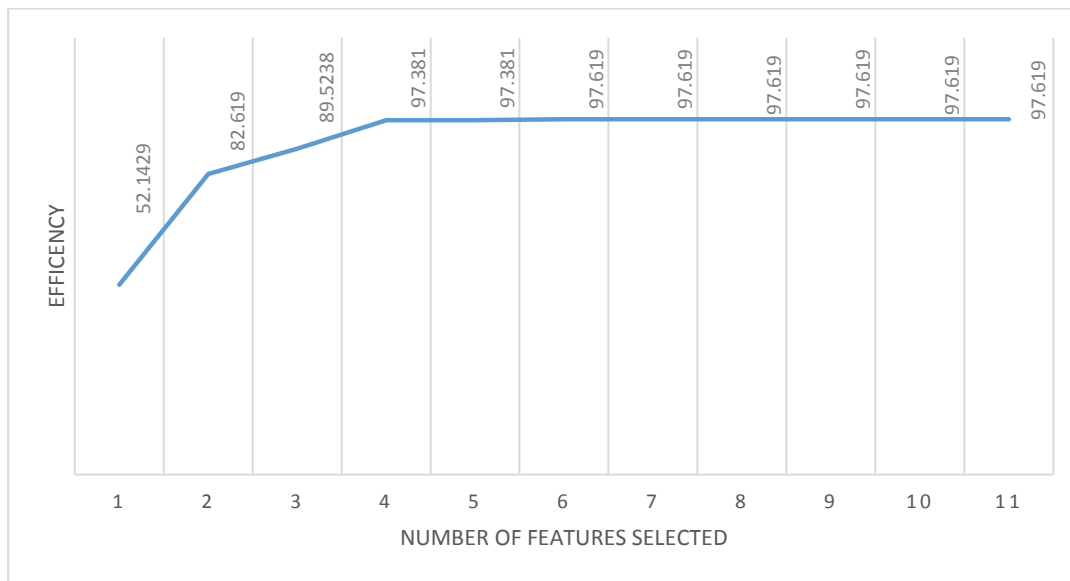


Figure: 10 Effect of features on classification accuracy

The features selected are *V1*, *V2*, *V3*, *V4*, *V5* and *V6* as this combination gave the maximum classification accuracy.

6. Feature classification

The K^* algorithm [16] is a method of cluster analysis which mainly aims at the partition of “ n ” observation into “ k ” clusters in which each observation belongs to the cluster with the nearest mean. K^* algorithm can be described as an instance based learner which uses entropy as a distance measure. The benefits are that it provides a consistent approach to handling of real valued attributes, symbolic attributes and missing values. K^* is a simple, instance based classifier, similar to K Nearest Neighbour (K - NN). New data instances, x , are assigned to the class that occurs most frequently amongst the k -nearest data points $y(j)$, where $j = 1, 2, \dots, k$. Entropic distance is then used to retrieve the most similar instances from the data set. By means of entropic distance as a metric has a number of benefits including handling of real valued attributes and missing values

$$K^*(y_i, x) = -\ln P^*(y_i, x)$$

Where P^* is the probability of all transformational paths from instance x to y . It can be useful to understand this as the probability that x will arrive at y via a random walk in IC feature space. It will be performed optimization over the percent blending ratio parameter which is analogous to KNN "sphere of influence", prior to assessment with other Machine Learning methods.

7. Results and discussions

Analysis of the acoustic signals from the gearbox under good and faulty conditions (20%, 40%, 60%, 80%, 100% and 150%) was done followed by the feature extraction, selection and classification using discrete wavelet transform and K star algorithm. Using Daubechies 5 wavelet transform ($v1 - v11$) 11 features were extracted from the acoustic signals. The features which contribute for the feature classification were selected using J48 decision tree. The maximum classification accuracy was obtained when the features *V1*, *V2*, *V3*, *V4*, *V5* and *V6* were selected for training and testing. The remaining features were ignored as they reduced the classifier's accuracy.

7.1 Effect of number of features

(i) The decision tree gave a preview of the relative importance of the features extracted using ‘db5’ wavelets. The feature at the top gives the maximum contribution, *i.e.* V2, the contribution reduces down the tree.

(ii) The information gained as the entropy reduced gave the measure of the discriminating capability of the feature in a given data set and thus useful features were selected.

(iii) The classification accuracy with J48 algorithm is lower with less features and increases as the number of features increases, reaching a maximum value at 6 features and thereafter it remains constant at 97.619%.

7.2 Feature classification using K star algorithm

(i) Features V1, V2, V3, V4, V5 and V6 which contributed for the classification were selected for training and testing from the J48 decision tree

(ii) When the number of instances from the data set of 60 increased, the classification accuracy was 100% initially and was constant for some range of values of the global blend and further diminished as the value of the global blend decreased. The reduction of accuracy may be due to increase in complexity of the problem or unnecessary confusion when the number of features increases.

(iii) The summary of stratified cross validation is illustrated in Table 2. The confusion matrix gives an insight into the details of misclassification. In confusion matrix, the first row corresponds to the total number of data points for the ‘GOOD’ condition of the gearbox and the first column in first row corresponds to the correct classification as ‘GOOD’.

(iv) The second row in the confusion matrix represents the total number of 60 data points for 20% fault condition and the first column represents misclassification of those data points as ‘GOOD’. The cell corresponding to the second row and second column in the confusion matrix represents the number of data points in 20% fault condition that have been correctly classified as 20% fault condition.

(v) As shown in the confusion matrix, it is observed that all the data points have been classified correctly. Thus giving 100% accuracy.

	Good	20%fault	40%fault	60%fault	80%fault	100%fault	150%fault
Good	60	0	0	0	0	0	0
20%fault	0	60	0	0	0	0	0
40%fault	0	0	60	0	0	0	0
60%fault	0	0	0	60	0	0	0
80%fault	0	0	0	0	60	0	0
100%fault	0	0	0	0	0	60	0
150%fault	0	0	0	0	0	0	60

Table 2 Confusion matrix using K star algorithm

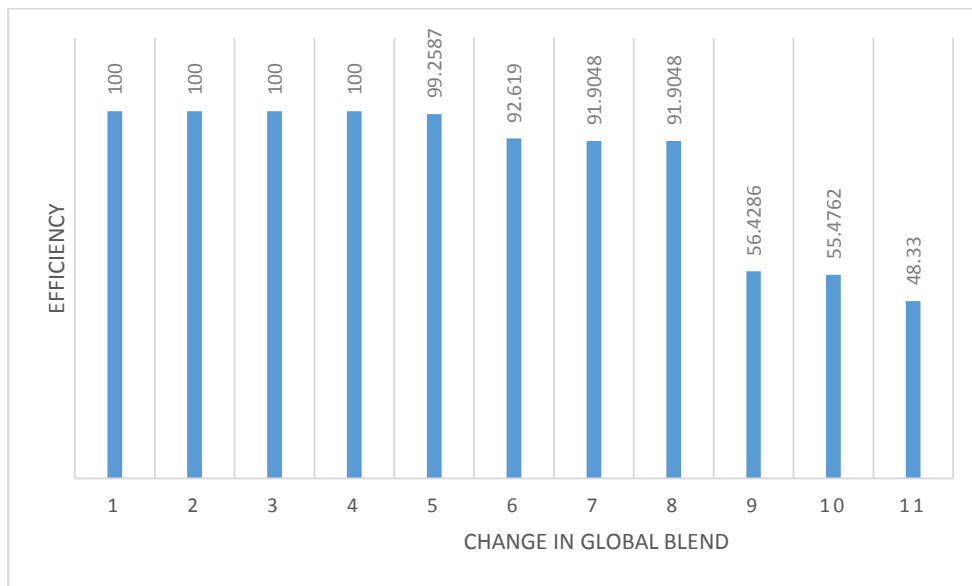


Figure: 11 Efficiency with variation of global blend

Total Number of Instances 420

Correctly Classified Instances 420 100%

Incorrectly Classified Instances 0 0%

Kappa statistics = 1

Mean Absolute error= 0.0017

Root mean squared error= 0.0077

Relative absolute error= 0.6744%

Root relative square error= 2.2044%

The True Positive (TP) rate is a measure of instances which are considered as true instances in the same class and should be 1 in ideal cases. If a class is declared 'GOOD' all the instances in the same class are classified correctly and the TP rate is 1. Classes with artificial faults of 20%, 40%, 100% and 150% show TP rate as 1 implies that all the instances were classified correctly. The False Positive (FP) rate is a measure of non-satisfying conditions which were interpreted as satisfying conditions for perfect classification. From Table 2, it can be implied that FP is zero. Precision is the fraction of the retrieved instances that are relevant while recall is the fraction of relevant instances that are retrieved and both should be ideally one. In the present study they both confirm to an ideal condition.

8. Conclusion

The classification of features of the faults of the gearbox was done by extracting the discrete wavelet features from the acoustic signals. K star algorithm which was used for classification gave 100% accuracy. The proposed system of using J48 decision tree for selection along with the classification using K star algorithm is extremely effective and can be used for the real time condition monitoring of gearbox with minimal expenditure on the microphone and various components. This system can be implemented properly as it is cost effective and also reduces the risk of failure of the gearbox. This also enables prediction of the faults in the gearbox. Further developments can be made to increase the speed of the process of computation as K star algorithm is a time consuming process.

9. References

- [1] Lambrou T, Kudumakis P, Speller R, Sandler M and Linney A 1998 Classification of Audio Signals Using Statistical Features on Time and Wavelet Transform Domains *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 1998)* **Vol. 6** pp 3621-24
- [2] Sakthivel N, Indira V, Nair B and 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.
- [3] Lang M, Guo H, Odegard J, Burrus C and Well R 1996 Noise reduction using an undecimated discrete wavelet transform *IEEE Signal Processing Letters* **Vol. 3** Issue 1 pp 10-12
- [4] Shensa M 1992 The discrete wavelet transform: wedding the a trous and Mallat algorithms *IEEE Transactions on Signal Processing* **Vol. 40** Issue 10 pp 2464-82
- [5] Bergland G 1969 A guided tour of The Fast Fourier Transform *IEEE Spectrum* **Vol. 6** Issue 7 pp 41-52
- [6] Kwok H and Jones D 2000 Improved instantaneous frequency estimation using an adaptive short-time Fourier transform *IEEE Transactions on Signal Processing* **Vol. 48** Issue 10 pp 2964-72
- [7] Lipo W and Daniel L 1993 Artificial Neural Networks: Oscillations, Chaos and Sequence Processing *IEEE Computer Society Press*
- [8] Kazuo T and Michio S 1992 Stability analysis and design of fuzzy control systems *Fuzzy Sets and Systems Elsevier* **Vol. 45** Issue 2 pp 135-56
- [9] Saravanan N and Ramachandran K 2009 Fault diagnosis of spur bevel gear box using discrete wavelet features and Decision Tree classification *Expert Systems with Applications: An International Journal* **Vol. 36** Issue 5 pp 9564-73
- [10] Sugumaran V, Ramachandran K and Muralidharan V 2007 Feature selection using Decision Tree and classification through Proximal Support Vector Machine for fault diagnostics of roller bearing *Mechanical Systems and Signal Processing Elsevier* **Vol. 21** Issue 2 pp 597-1196
- [11] Glenn M and Mangasarian O 2005 Multicategory Proximal Support Vector Machine Classifiers *Machine Learning Springer* **Vol. 59** Issue 1 pp 77-97
- [12] Park D, El-Sharkawi M, Marks R and Altas L 1991 Electric load forecasting using an artificial neural network *IEEE Transactions on Power Systems* **Vol. 6** Issue 2 pp 442-49
- [13] Cristina O, Louis W 2003 A complete fuzzy decision tree technique *Fuzzy Sets and Systems Elsevier* **Vol. 138** Issue 2 pp 221-54
- [14] Friedl M and Brodley C 1997 Decision tree classification of land cover from remotely sensed data *Remote Sensing of Environment Elsevier* **Vol. 61** Issue 3 pp 399-409
- [15] Deeman Y and Mohammed A 2013 Intrusion Detection System Based on K-Star Classifier and Feature Set Reduction *IOSR Journal of Computer Engineering* **Vol. 15** Issue 5 pp 107-12
- [16] Lin J and Zuo M 2003 Gearbox Fault Diagnosis using Adaptive Wavelet Filter *Mechanical Systems and Signal Processing Elsevier* **Vol. 17** Issue 6 pp 1219-35

- [17] Griffin D and Jae L 1984 Signal modification from Short Time Fourier Transform *IEEE Transactions on Acoustics, Speech and Signal Processing* **IEEE** Vol. **32** Issue 2 pp 236-43
- [18] Tina R and Sherekar S 2013 Performance Analysis on Naïve Bayes and J48 Classification algorithm for data classification *International Journal of Computer Science and Applications* Vol. **6** Issue 2
- [19] Margaret H and Sridhar S 2006 Data mining, Introductory and Advanced Topics *Pearson education* 1st Edition
- [20] Amarnath M and Praveen K 2014 Local fault detection in helical gears via vibration and acoustic signals using EMD based statistical parameter analysis *Measurement* Vol. **58** pp 154–64