Advances in Parallel Computing Algorithms, Tools and Paradigms
D.J. Hemanth et al. (Eds.)

doi:10.3233/APC220016

An Investigation of Ensemble Learning Algorithms for Fault Diagnosis of Roller Bearing

Gangavva C[a,1], J Alamelu Mangai[a] and Mohit Bansal[b]
[a]*Department of CSE, Presidency University, Bangalore, India*
[b]*Kaizen Secure Voiz Private Limited, Chennai, India*

Abstract. Roller Bearing (RB) is one of the critical mechanical components in rotating machineries. Failure of a bearing may cause the fatal breakdown of an entire machine and inestimable financial losses due to its continuous rotation. Hence, it is significant to diagnose the fault accurately at an early stage so that it helps in predictive maintenance of the machine from malfunctioning. In the recent developments, Machine Learning (ML) has shown a drastic change in the way we predict, analyze and interpret the results. In this paper, a diagnostic technique is being proposed to identify the bearing faults that employs ensemble learning algorithms such as Bagging, Extra Tree and Gradient Boosting classifiers. The proposed method includes 1) Pre-processing of vibration data 2) Extracting statistical features such as Mean, Standard Deviation, Kurtosis, Crest Factor and Mel-Frequency Cepstral Co-efficient (MFCC) features and 3) Training the Ensemble Learning algorithms for classifying the various faults based on extracted features. For experimentation, vibration data is collected from the Case Western Reserve University (CWRU) Laboratory to diagnose 12 different fault types associated with Inner Race (IR), Outer Race (OR), Ball fault and normal bearing of varying diameters. Results shows that Ensemble learning algorithms performs better based on MFCC features as compared to statistical features.

Keywords. Fault Diagnosis, Roller Bearing, Machine Learning, Ensemble Learning, Statistical and MFCC features

1. Introduction

Rotating equipment is extensively used in many fields including aerospace, automobile, medical, agriculture and so on. Due to the continuous rotation of roller bearing under varying working conditions, failure may occur frequently and hence it may result in an unexpected severe loss in safety and economy [1]. Hence, it is important to diagnose bearing faults accurately to prevent repeated failures and additional expense as part of condition monitoring.

Over the years, many intelligent fault diagnosis methods have been developed based on vibration signals. Fault diagnosis of rotating machineries based on Machine Learning (ML) has received more importance because of its high performance in classifying various faults accurately. The measured vibration signals are pre-processed

[1] Gangavva C, Presidency University, Bangalore, India; E-mail: gangavc31@gmail.com

and a set of features are extracted, then these features are used as input to ML model to classify the faults [2]. An ensemble of Ensemble Learning (EL) algorithms has demonstrated an accuracy of 97.91% for fault classification based on statistical features [3]. In [4], Random Forest (RF) ensemble learning has been employed and it has shown the performance accuracy of 99%in classifying bearing faults by pre-processing vibration data using Fast Fourier Transform (FFT) to reveal intrinsic features.

In this paper, various Ensemble Learning algorithms are investigated by extracting MFCC and Statistical features from vibration signals to diagnose the faults of roller bearing. The paper is structured as follows: Various Feature extraction and ML techniques used in fault diagnosis are discussed in Section-2. Investigations of Ensemble Learning algorithms for fault diagnosis are discussed in section-3. An experimental set-up and conclusion is given in section-4 and 5 respectively.

2. Related Work

Condition monitoring of roller bearing based on vibration signal is the most popular technique. As per the literature survey, extracting accurate features from raw vibration signals is a challenging task due to its non-linear characteristic. Therefore, various feature extraction methods are employed in frequency, time and time-frequency domains [5] in order to extract the relevant features. In [6], ML models such as Gaussian Mixture Models and Hidden Markov Models were discussed to classify the faults by extracting linear, Non-linear, MFCC and kurtosis features using Multi-Scale Fractal Dimension (MFD).

A combination of Ensemble Empirical Mode Decomposition (EEMD) and RF was applied for fault diagnosis of rolling bearing by comparing EEMD with wavelet method [7]. A comparative analysis of random forest (RF), artificial neural networks (ANN) and support vector machine (SVM) was presented in [8] for intelligent diagnosis of rotating machinery. An application of Improved Ensemble Learning Algorithm i.e. AdaBoost algorithm and exponential Smoothing Predictive Segmentation technique was used for Fault Diagnosis of Rolling Bearing in Rail Train by collecting the bearing’s vibration data provided by the Rail Traffic Control of State Key Laboratory, Safety of Beijing Jiaotong University and CWRU dataset for the experiment [9]. Kangkan Bhakta et al employed Gradient Boosting (GB) Ensemble learning Algorithm and identified various types of bearing faults with an accuracy of 99.58% by pre-processing the signals through analysis of real cepstrum [10]. A hybrid ensemble learning method was developed to diagnose the fault of Rotating Machinery by combining the ML algorithms like Support Vector Machine, K-Nearest Neighbors, and decision tree classifiers with Bootstrap samples [11]. Secure privacy and control data possession to secure the health records with duplication removing technique as given in [16].

3. Proposed methodology

The proposed multiclass fault diagnostic method to identify the bearing faults under various running conditions is shown in Figure 1. The method diagnoses the different types of faults such as IR, OR, Ball fault and Normal bearing based on the feature extracted from vibration data. Initially, vibration data is collected, MFCC and statistical features are extracted from pre-processed vibration signals and then the feature dataset is given as input to various Ensemble Learning algorithms such as Bagging, Extra Tree and Gradient Boosting Classifiers to diagnose the different fault types with improved diagnostic capability of the classifiers. Details of each step are discussed in the following subsections.

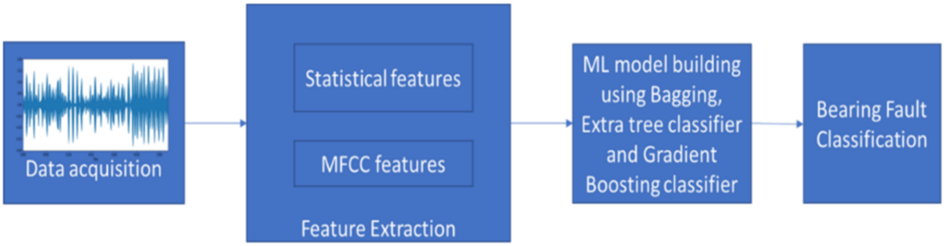


Figure 1. Block Diagram of the Proposed Fault Diagnosis Method

3.1. Data Acquisition

Vibration data of RB is collected from the CWRU [12] bearing repository which has both normal and faulty bearing data.

The vibration data is measured for three operating conditions:

- Baseline – no fault, sampling rate of 12 kHz with rotations of 1797 per minute.
- Outer race fault- Sampling rate of 12 kHz with rotations of 1797 per minute.
- Inner race fault- Sampling rate of 12 kHz with rotations of 1797 per minute.

*3.2. **Feature Extraction***

Features represent the various characteristics information present in the vibration signal. In the proposed method, four statistical features and twenty MFCC features are extracted from the pre-processed vibration signals and this feature set is given as an input to the ensemble machine learning techniques for classifying the faults of bearing.

(1) Statistical features – It includes Mean, Standard Deviation, Kurtosis, and Crest Factor. Each of these features and its corresponding formula is given in Table 1.

Table 1. Statistical features

S.No.	**Statistical features**	**Formula**
1	Mean	$\frac{\text{Sum of terms}}{\text{Count of terms}}$
2	Standard Deviation	$\sigma = \sqrt{\frac{\sum(xi-\mu)^2}{N}}$
3	Kurtosis	K=$\frac{\mu}{\sigma^4}$
4	Crest Factor	C= 20 log10 $\left(\frac{x_{peak}}{x_{rms}}\right)$

(2) Mel Frequency Cepstral Coefficients (MFCC) features - MFCC is observed as a best feature extraction (FE) technique for speech signal. Initially, the vibration signal is windowed and framed. Then, for each frame the frequency spectrum is captured by applying short-time Fourier transform (STFT). Next, the spectrum of short frame is obtained and then segmented by a scale of Mel frequency to highlight the features of a signal. Finally, the frequency spectrum is split up by discrete cosine transform (DCT) in order to reduce the feature of spectra to a lower dimensional feature, which is termed as Mel-Frequency Cepstral Coefficient (MFCC) feature [13].

3.3. Ensemble Learning Algorithms used in the proposed model for Fault Diagnosis

Ensemble learning techniques will combine the results from individual models to enhance the classification performance of the entire model. Among many EL algorithms, following three will perform better in fault diagnosis.

1) Bagging – Bagging (Bootstrap AGGregatING) involves the process of training several independent classifiers and then combining the outcomes of each individual classifier to make the final decision. The decision is made based on majority voting in classification problems and average is considered in regression. Random forest is one of the well-known techniques of bagging. Pseudocode for Bagging is described below.

<u>Pseudocode for Bagging</u>

```
Input: S-Training Samples, L - base learning classifier, I-Iterations
Bagging_func(S, I, L)
  for i ← 1 to I
  Si ← a bootstrap sample of S
 hi ← apply L to Si
 return h1 , h2 , . . . , hi
```

2) Extra Tree Classifier-In classification problems, extra tree classifier creates additional trees in sub samples of datasets and applies majority voting to improve the accuracy [14]. This classifier applies random thresholds for each features of sub-sample to obtain the best of the thresholds as a splitting rule. Pseudocode for extra tree algorithm is described below.

Pseudocode for Extra Tree classifier

Split_a_node(S)
Input: Learning subset 'S' corresponding to the node which we want to split
Output: a split [a < a_c] or nothing

- If Stop_split(S) is TRUE then return nothing.
- Else select 'k' attributes {a1,..., ak } among all candidate attributes.
- Choose 'k' splits {s1,...,sk }, where si = Pick_a_random_split(S, ai), ∀i = 1,..., k.
- Return a split s∗ such that Score(s∗, S) = maxi=1,...,k Score(si, S).

Pick_a_random_split(S,a)
Inputs: a subset 'S' and an attribute 'a'
Output: a split
– Let a^S_{max} and a^S_{min} denote the maximal and minimal value of a in S.
– Choose a random cut-point a_c uniformly in [a^S_{min}, a^S_{max}]
– Return the split [a < a_c].
Stop_split(S)
Input: a subset 'S'
Output: a boolean value
– If |S| < n_{min}, then return TRUE.
– If all attributes are constant in S, then return TRUE.
– If the output is constant in S, then return TRUE.
– Otherwise, return FALSE.

3) Gradient Boosting (GB) Classifiers - Gradient Boosting algorithm trains several models in a very successive and sequential manner. GB minimizes the loss function of a model by adding weak classifier using a gradient descent procedure. Pseudocode for GB is described below.

Pseudocode for Gradient Boosting algorithm

Input : Training set {(xi ,yi)} where i=1…n; A differentiable loss function L(y,f(x)),'M' iterations.
1. Initialize base learner with a constant value:
   F0(x) = argmin $\sum_{i=1}^{n} L(y_i, \gamma)$
2. for m = 1 to M
   a) Compute pseudo –residuals:
   $$\text{rim} = -\left[\frac{\partial L(y_i, f(x_i))}{\partial f(x_i)}\right]_{f=f_{m-1}} , for\ i = 1, \ldots \ldots \ldots n;$$
   b) Fit a base learner hm(x) to the targets rim using data set {(xi ,γi)}for i=1,…..,n.
   c) Compute multiplier rm by solving the following optimization problem.
   $$\gamma_m = argmin \sum_{i=1}^{n} L(y_i, f_{m-1}(x_i) + \gamma h_m)x))$$
   d) Update the model :
   $$f_m(x) = f_{m-1}(x) + \gamma_m h_m)x)$$
3. Output $f_m(x)$

4. Experimental Set-up

An experimental set-up of bearing data collection from CWRU dataset is shown in Figure 2. It consists of a 2 hp motor, a torque transducer, a dynamometer, control electronics, fan end and drive end bearings which are supported by the motor shaft. Vibration data was recorded by conducting experiments using a 2 horsepower Reliance

Electric motor by placing accelerometer close to and distant from the motor bearings. These bearings were induced with faults using electro-discharge machining (EDM) at the IR, rolling component (i.e. ball) and OR with diameters ranging from 0.007 to 0.040 inches. Vibration data was noted with motor speeds of 1797 to 1720 RPM with a load of 0 to 3 horsepower by reinstalling faulted bearings into the test motor.

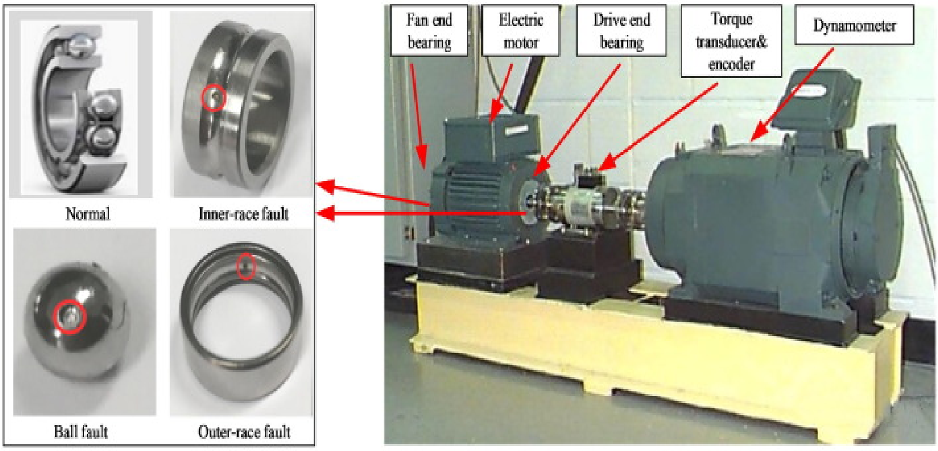


Figure 2. Experimental set-up of CWRU bearing data collection

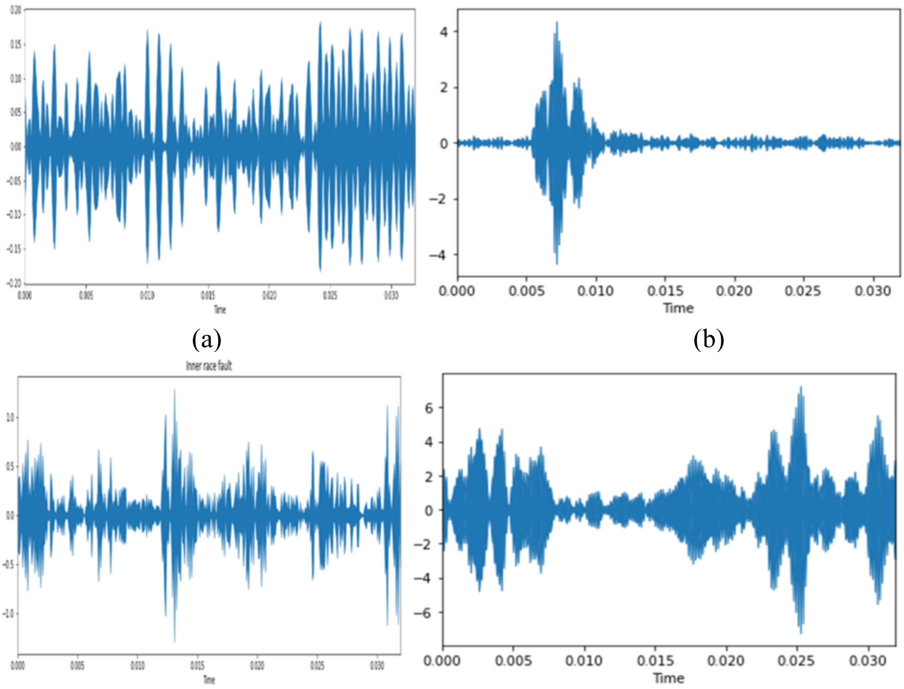
5. Results and Discussion

In this study, the vibration data is collected from CWRU dataset for three different bearing conditions. The collected data for each bearing condition and the class distribution is shown in table 2.

Table 2.Class distribution and description of the CWRU dataset of 12DriveEnd

Class	Type of Bearing Fault	No.of Samples	Data description
0	0.007-Ball	319	Ball fault level =0.007
1	0.007-InnerRace	315	Inner Race fault level=0.007
2	0.007-OuterRace12	318	Outer Race fault level at 12’o clock = 0.007
3	0.014-Ball	317	Ball fault level =0.014
4	0.014-InnerRace	317	Inner Race fault level=0.014
5	0.014-OuterRace6	317	Outer Race fault level at 6’o clock = 0.014
6	0.021-Ball	317	Ball fault level =0.021
7	0.021-InnerRace	318	Inner Race fault level=0.021
8	0.021-OuterRace12	317	Outer Race fault level at 12’o clock = 0.021
9	0.028-Ball	314	Ball fault level =0.028
10	0.028-InnerRace	314	Inner Race fault level=0.028
11	Normal	634	Normal Bearing

The time-domain characteristics of the raw vibration signals for different faults areshown in Figure 3(a)-(d).

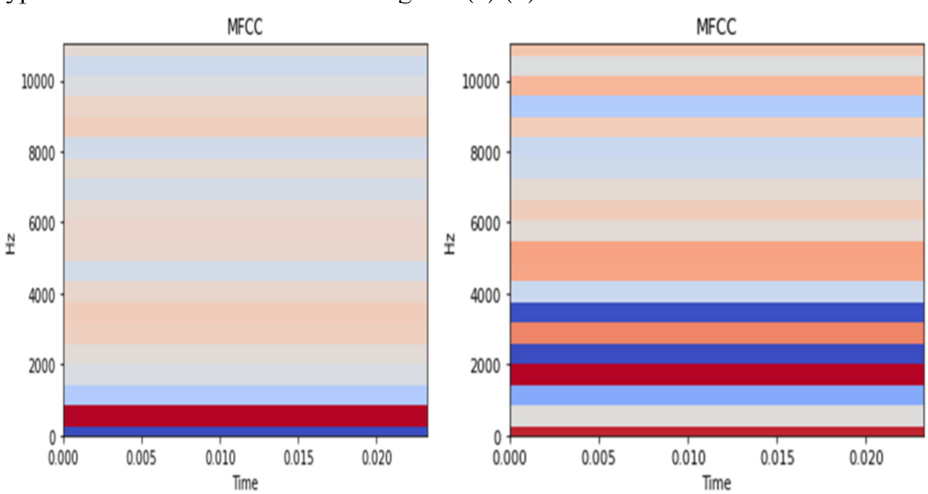


(a) (b)

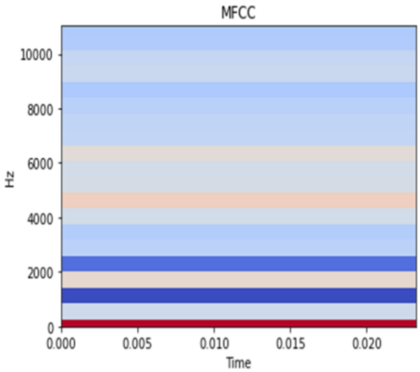
(c) (d)

Figure 3. Time-domain characteristics for(a)Normal Bearing(b) Outer Race Fault(c) Inner Race fault (d) Ball Fault

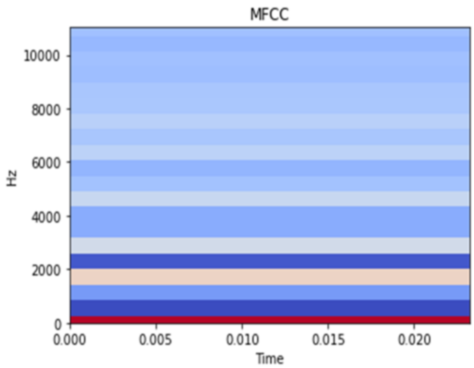
After pre-processing the vibration data, statistical features and twenty MFCC features were extracted from the pre-processed vibration signal in time-frequency domain. Then, the extracted Statistical and MFCC features were given as inputs to the ensemble learning techniques for multi-class classification. The MFCC feature for each type of the fault is shown below in Figure 4(a)-(d).



(a) (b)



(c)



(d)

Figure 4. MFCC features for (a) Normal Bearing(b) Outer Race fault (c) Inner Race fault (d) Ball fault

The practical application of the proposed method is verified by diagnosing 12 different types of faults using Python's ML package called Sk learn. The performance metrics for fault diagnosis of roller bearing based on both statistical and MFCC features are shown in Table 3.

Table 3. Performance Metrics

S.No.	Algorithm	Accuracy	
		Statistical features	MFCC features
1	Bagging Classifier	94.8%	98.03%
2	Extra Tree Classifier	95.5%	99.75%
3	Gradient Boosting Classifier	95.16%	98.98%

Based on results, it is observed that Ensemble Learning Algorithms perform better based on MFCC features as compared to statistical features.

6. Conclusion

In this paper, a comparative study of various ensemble learning algorithms for fault diagnosis of rotating machinery is proposed with the experimental results of the same. MFCC based Ensemble Learning classifiers demonstrate very promising results in terms of accuracy compared to statistical features. In future, the performance of these Ensemble Models for Fault Diagnosis will be analyzed based on Image extracted from the vibration signals of roller bearings for various fault conditions.

7. References

[1] Jiang, Qinyu, Faliang Chang, and Bowen Sheng. "Bearing fault classification based on convolutional neural network in noise environment." IEEE Access 7 (2019).

[2] Abid, Anam, Muhammad Tahir Khan, and Javaid Iqbal. "A review on fault detection and diagnosis techniques: basics and beyond." Artificial Intelligence Review (2020): 1-26.

[3] Patil, Priyanka S., et al. "Ensembles of Ensemble Machine Learning approach for Fault detection of Bearing." Solid State Technology 63.6 (2020): 14442-14455.

[4] Sikder, Niloy, et al. "Fault diagnosis of motor bearing using ensemble learning algorithm with FFT-based preprocessing." 2019 International Conference on Robotics, Electrical and Signal Processing Techniques (ICREST). IEEE, 2019.

[5] Yang, Hongyu, Joseph Mathew, and Lin Ma. "Vibration feature extraction techniques for fault diagnosis of rotating machinery: a literature survey." Asia-Pacific Vibration Conference. No. 42460. 2003.

[6] Nelwamondo, Fulufhelo V., Tshilidzi Marwala, and Unathi Mahola. "Early classifications of bearing faults using hidden Markov models, Gaussian mixture models, mel-frequency cepstral coefficients and fractals." International Journal of Innovative Computing, Information and Control 2.6 (2006): 1281-1299.

[7] Qin, Xiwen, et al. "The fault diagnosis of rolling bearing based on ensemble empirical mode decomposition and random forest." Shock and Vibration 2017 (2017).

[8] Qin, Xiwen, et al. "The fault diagnosis of rolling bearing based on ensemble empirical mode decomposition and random forest." Shock and Vibration 2017 (2017).

[9] Han, Lu, et al. "Fault diagnosis of rolling bearings in rail train based on exponential smoothing predictive segmentation and improved ensemble learning algorithm." Applied Sciences 9.15 (2019): 3143.

[10] Bhakta, Kangkan, et al. "Fault diagnosis of induction motor bearing using cepstrum-based preprocessing and ensemble learning algorithm." 2019 International Conference on Electrical, Computer and Communication Engineering (ECCE).

[11] Tian, Jing, et al. "An ensemble learning-based fault diagnosis method for rotating machinery." 2017 Prognostics and System Health Management Conference.

[12] Available at https://engineering.case.edu/bearingdatacenter

[13] Yang, Hongyu, Joseph Mathew, and Lin Ma. "Vibration feature extraction techniques for fault diagnosis of rotating machinery: a literature survey." Asia-Pacific Vibration Conference. No. 42460. 2003.

[14] Geurts, Pierre, Damien Ernst, and Louis Wehenkel. "Extremely randomized trees." Machine learning 63.1 (2006): 3-42. Available at https://scikit-learn.org/stable/