MANJUNATHA G
and
H.C. CHITTAPPA

Bearing fault classification using statistical features and machine learning approach

Bearing degradation is the most common source of faults in machines. In this context, this work presents a monitoring scheme to diagnose bearing faults using machine learning approach. In this approach classification of healthy and faulty conditions of the bearing is carried out using artificial neural network (ANN). A set of statistical features are extracted from the acquired vibration signals. The decision tree technique is used to select significant features out of all statistical extracted features. The selected features were classified using different classifiers. Based on the various classifier results obtained, the ANN classifier achieve the maximum classification accuracy which is recommended for online monitoring and fault diagnosis of the bearing in various machines.

Key words: Bearing fault, diagnosis, ANN, classifiers, statistical features.

1.0 Introduction

Continuous monitoring of the machines is essential to reduce the breakdowns in order to increase the productivity as its role is inevitable. The defect in bearing may cause the failure in machinery and that causes a severe loss in industry. The failure in bearing reduces the efficiency and hence decreases the productivity in industrial operation. The acoustic emission and vibration are the two widely used measuring parameters which is used for the condition monitoring of machines. Vibration signals are widely used in condition monitoring of bearing. Fault detection is achieved by comparing the signals of bearing running under normal and faulty conditions. The faults considered in this study are inner race fault, outer race fault and rolling element fault. By the application of seismic transducers, the vibration levels are measured for different fault conditions.

Researchers have proposed many schemes for vibration signal analysis in the past. Adequate fault sensitive features have been extracted from the vibration signals and analysis has been done using intelligent decision making techniques. A review of various vibration feature extraction techniques in time domain, frequency domain, and joint time frequency domain for fault diagnosis in rotating machines is presented [1,2]. Statistical time domain features have been effectively used for extracted feature set along with various other features [3, 4, 5]. A comprehensive review of "Artificial algorithms" used in rotating machinery fault diagnosis has been done by Liu et al. [6]. They have discussed advantages and limitations of KNN, Naïve Bayes, SVM, ANN and Deep learning algorithms. Zhang et al. [7] have presented a systematic review of "Machine learning and deep learning algorithms" for bearing fault diagnostics. A review on "Meta Classification Algorithms" using WEKA has been presented by Bal and Sharma [8] in their research paper. This paper evaluates the importance of statistical features for the fault classification using different classifiers.

1.1 Experimental set-up

Fig.1 shows the experimental set up. It is comprised 1 HP motor, bearing 1, bearing 2 and loading platform. Motor shaft is supported by end bearings. Single point faults were seeded to the test bearing (bearing 2) using laser cutting method. The fault diameter of 0.1mm, 0.3mm and 0.5mm were seeded

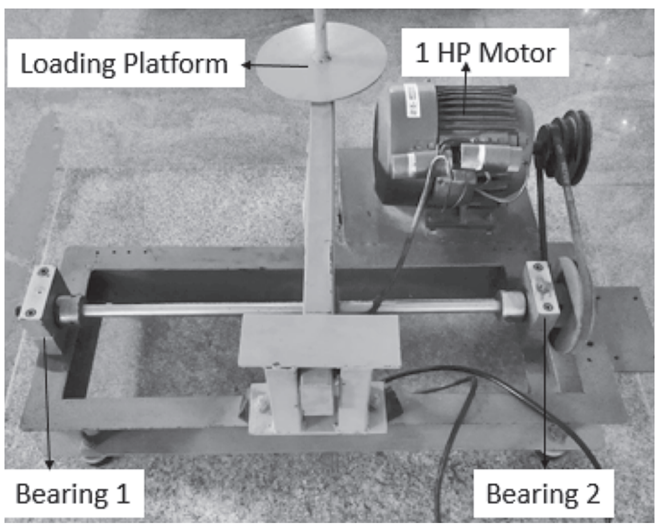

Fig.1: Experimental set up

Mr. Manjunatha G, Research Scholar, Department of Mechanical Engineering, UVCE, Bangalore University, Assistant Professor, School of Mechanical Engineering, REVA University and Dr. H.C. Chittappa, Assistant Professor, School of Mechanical Engineering, REVA University, Professor, Department of Mechanical Engineering, UVCE, Bangalore University, Karnataka, India. E-mail: manjunathag16@gmail.com / dr.hcchittappa@gmail.com

individually at the inner race, rolling element and outer race. Test rig was reinstalled with faulted bearings and accelerometers (not shown) attached to the housing were used to collect vibration data. This experiment runs at constant load of 200 N and speed of 1200 rpm which is same as real time application.

2.0 Statistical features extraction and selection process

Statistical feature extraction [9], [10], [11], [12], [13]–[15] is the least time consuming feature extraction method compared to other feature extraction techniques. As the name suggests, in this method, various mathematical and statistical features are calculated using well-known mathematical equations. In present study, a set of selected parameters were standard error, standard deviation, sample variance, kurtosis, skewness, range, minimum value, maximum value and sum. Table 1 shows proposed set of statistical features for the characterization of the available signals in the time domain analysis. Feature selection process involves selection of features that are highly efficient and able to classify representing conditions in case of condition monitoring. In this study, 10 fold cross validation is performed using decision tree. Decision tree represents knowledge with a tree structure comprised with a set of branches and nodes. Feature vectors were classified using decision tree. The decision tree algorithm was applied to the statistical features and resultant tree is shown in Fig.2.

3.0 Results and discussion

The condition monitoring of bearing is performed by different classifiers using machine learning approach. The results of the study are discussed.

TABLE 1: PROPOSED SET OF STATISTICAL FEATURES FOR THE TIME DOMAIN ANALYSIS, WHERE x(i) IS A SAMPLE FOR i = 1, 2,…AND N IS THE NUMBER OF POINTS FOR EACH ACQUIRED VIBRATION SIGNAL

Sl. No	Statistical feature	Mathematical equation
1	Standard error	$X_1 = \frac{X_2}{N}$
2	Standard deviation	$X_2 = \sqrt{\left[\frac{1}{N\sum_{i=1}^{N}(x_i - x_m)^2}\right]}$
3	Sample variance	$X_3 = X_2^2$
4	Kurtosis	$X_4 = \{\frac{n(n+1)}{(n-1)(n-2)(n-3)}\sum\left(x_i - \frac{x_m}{s}\right)^4\} - \frac{3(n-1)^2}{(n-2)(n-3)}$
5	Skewness	$X_5 = \frac{n}{n-1}\sum_{i=1}^{n}((x_i - x_m)/s)^3$
6	Maximum value	$X_5 = \max(x_i)$
7	Minimum value	$X_6 = \min(x_i)$
8	Range	$X_7 = \max(x_i) - \min(\mathrm{x_i})$
9	Sum	$X_8 = \sum_{i=1}^{n} x(i)$

3.1 STATISTICAL ANALYSIS FOR FAULT DIAGNOSIS

The statistical parameter values are computed from the acquired vibration signals by using mathematical equations. The parameters are computed for the healthy and different fault conditions of bearing. The computed parameters are tabulated in Table 2.

3.2 FAULT CLASSIFICATION

Artificial Neural Network (ANN) classifier is used in this experimentation to classify the faults in bearing. This steps in this machine learning algorithm are the feature extraction, then feature selection and then train the system for fault classification. The statistical parameter data are used as the input

TABLE 2: COMPARISON OF STATISTICAL PARAMETERS

	Static parameters	Healthy	IR 0.1	RE 0.1	OR 0.1	IR 0.3	RE 0.3	OR 0.3	IR 0.5	RE 0.5	OR 0.5
1.	Mean	0.01	0.02	0.02	0.01	0.02	0.01	0.01	0.01	0.01	0.01
2.	Standard error	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
3.	Median	0.01	0.02	0.02	0.01	0.02	0.01	0.01	0.00	0.01	0.01
4.	Mode	0.03	0.07	0.00	0.05	0.04	0.03	0.04	0.06	0.03	0.05
5.	Standard deviation	0.06	0.28	0.13	1.05	0.27	0.17	0.13	0.57	0.17	0.64
6.	Sample variance	0.00	0.07	0.01	1.11	0.07	0.03	0.01	0.33	0.02	0.41
7.	Kurtosis	-0.10	4.45	-0.03	3.88	15.9	15.9	0.27	0.49	0.11	19.4
8.	Skewness	-0.18	-0.05	-0.02	0.07	0.16	0.21	0.01	0.03	0.03	-0.14
9.	Range	0.43	3.07	1.16	10.5	4.57	3.88	1.16	4.97	1.45	12.3
10.	Minimum	-0.22	-1.50	-0.57	-5.21	-2.22	-1.78	-0.54	-2.53	-0.73	-5.97
11.	Maximum	0.21	1.57	0.58	5.30	2.35	2.10	0.62	2.44	0.72	6.32
12.	Sum	256	524	482	248	503	290	297	210	248	342

TABLE 3: CONFUSION MATRIX

I	II	III	IV	V	VI	VII	VIII	IX	X	Class
20	0	0	0	0	0	0	0	0	0	I-Healthy
0	20	0	0	0	0	0	0	0	0	II-IR 0.1
0	0	19	0	0	0	1	0	0	0	III-RE 0.1
0	0	0	20	0	0	0	0	0	0	IV-OR 0.1
0	0	0	0	20	0	0	0	0	0	V-IR 0.3
0	0	0	0	1	18	1	0	0	0	VI-RE 0.3
0	0	0	0	0	1	19	0	0	0	VII-OR 0.3
0	0	0	0	0	0	0	20	0	0	VIII-IR 0.5
0	1	0	0	0	0	0	0	19	0	IX-RE 0.5
0	0	0	0	0	0	0	0	0	20	X-OR 0.5

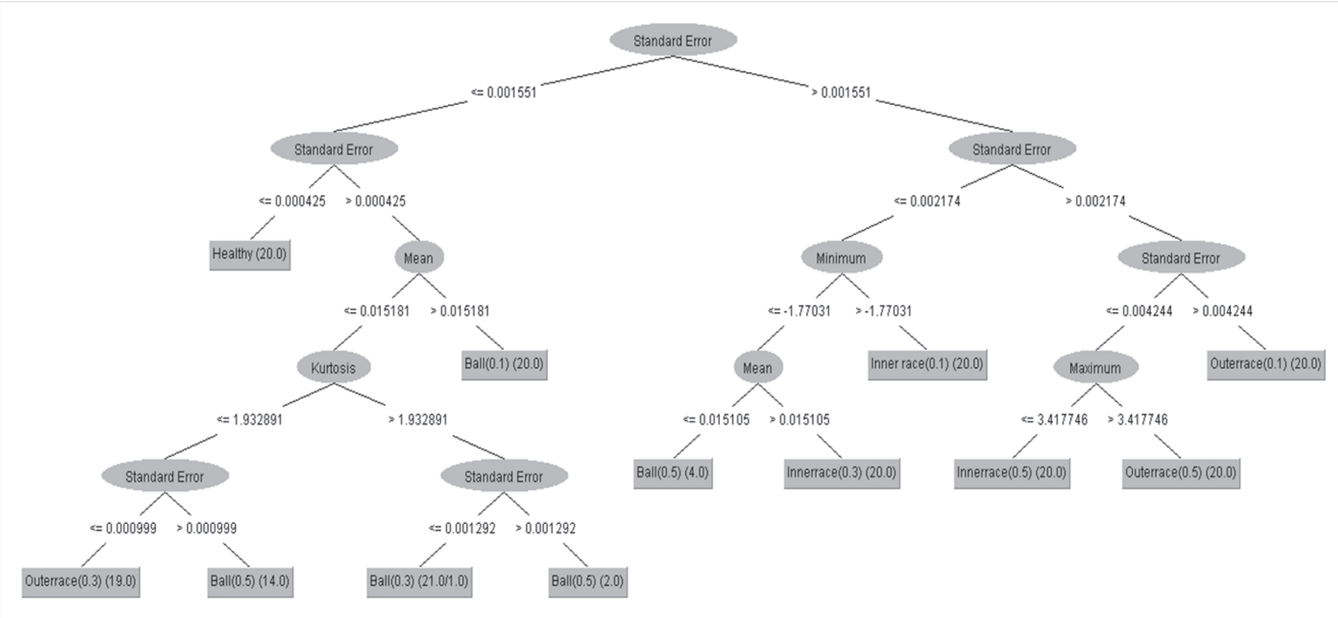

Fig.2: Decision tree for statistical features

data for the classifier to train the system. The statistical data extracted from the vibration signals of the bearings are tabulated in Table 1. To test the performance of the classifier the confusion matrix is used. The confusion matrix of ANN is shown in Table 2.

Table 2 shows confusion matrix for the results of ANN classification. The correctly classified instances are shown in diagonal elements. First cell in the row called as "class" represents the count of elements belonging to "class-I". It is found that all 20 elements belonging to "class-I" are correctly classified as "class-I" and there is no misclassification in class-I. Second row second cell in table above indicates the number of elements that belong to "class-II (Inner race (IR) seeded with 0.1 mm fault)" in which all the 20 instances belong to "class-II". Similarly, leading diagonal elements in all other rows give the number of correctly classified instances. Ninth element of ninth row shows the number of instances under "class-IX (Rolling Element (RE) seeded with 0.5 mm fault)". Out of 20 instances in "class-IX", 19 instances were classified as correct and remaining 01 instance is misclassified as "class-II". Out of 200 instances only 05 instances were found as misclassifications, giving an overall classification efficiency of 97.5%. As the classification efficiency is considerably high and error is very minimal, ANN classifier with statistical features can be used in numerous real-time applications.

4.0 Conclusion

Vibration signals acquired from the experimental set up are studied through machine learning approach. MATLAB was used to extract statistical features from signals acquired which in turn were used as input for further classification. J48 algorithm, a featured selection process was implemented to identify best contributing features. Further, ANN classifier was used for feature classification. It yields 97.5% accuracy for the extracted vibration signals. This approach can be applied in various real time applications.

References

1. H. Yang, J. Mathew, and L. Ma, (2003): "Vibration Feature Extraction Techniques for Fault Diagnosis of Rotating Machinery – A Literature Survey," Asia Pacific Vib. Conf. Aust., no. November, pp. 1–7.
2. S. N. Engin, K. Gulez, and M. N. M. Badi, (1999): "Advanced signal processing techniques for fault diagnostics – a review," *Mathematical and Computational Applications*, vol.4, no.1. pp.121-136.
3. M. Delgado et al., (2011): "Bearing Faults Detection by a Novel Condition Monitoring Scheme based on Statistical-Time Features and Neural Networks," *IEEE*, no. c.
4. B. R. Nayana and P. Geethanjali, (2018): "Identification of Bearing Faults Using Statistical Time Domain Features and Fused Time-Domain Descriptor Features,".
5. P. K. Kankar, S. C. Sharma, and S. P. Harsha, (2011): "Fault diagnosis of ball bearings using machine learning methods," Expert Syst. Appl., vol.38, no.3, pp. 1876–1886.
6. R. Liu, B. Yang, E. Zio, and X. Chen, (2018): "Artificial intelligence for fault diagnosis of rotating machinery : A review," *Mech. Syst. Signal Process*., vol.108, pp. 33–47, Aug.
7. S. S. Zhang, S. S. Zhang, B. Wang, and T. G. Habetler, (2019): "Machine Learning and Deep Learning Algorithms for Bearing Fault Diagnostics - A Review," Proc. 2019 IEEE 12th Int. Symp. Diagnostics Electr. Mach. Power Electron. Drives, SDEMPED 2019, pp. 257–263.
8. R. Bal and S. Sharma, (2016): "Review on Meta Classification Algorithms using WEKA," *Int. J. Comput. Trends Technol*., vol.35, no.1, pp. 38–47.
9. Fault diagnosis of automobile gearbox based on machine learning techniques. Procedia Engineering, 97(0): pp. 2092 – 2098, (2014).
10. Jalel Chebil, Meftah Hrairi, and Nazih Abushikhah. (2011): Signal analysis of vibration measurements for condition monitoring of bearings. *Australian Journal of Basic and Applied Sciences*, vol.5(1), pp.70-78.
11. V.E. Gai. (2014): Method of diagnostics of the state of rolling element bearing on the basis of the theory of active perception. In *Mechanical Engineering, Automation and Control Systems* (MEACS), International Conference on, pp.1-4.
12. P. K. Kankar, Satish C. Sharma, and S. P. Harsha. (2011): Fault diagnosis of ball bearings using machine learning methods. *Expert Syst. Appl.*, 38(3), pp. 1876–1886.
13. Abd Kadir Mahamad and Takashi Hiyama. (2011): Fault classiûcation based on artiûcial intelligence methods for induction motor. Internation Journal of Innovative Computing, *Information and Control,* 7(9), pp. 5477 – 5494.
14. Thomas W. Rauber, Francisco de Assis Boldt, and Fla´vio Miguel Vareja˜o. (2015): Heterogeneous feature models and feature selection applied to bearing fault diagnosis. *IEEE Transactions on Industrial Electronics,* 62(1), pp. 637–646.
15. Sadok Sassi, Bechir Badri, and Marc Thomas. (2007): A numerical model to predict damaged bearing vibrations. *Journal of Vibration and Control*, 13(11), pp. 1603– 1628.