Comparison of the supervised machine learning techniques using WPT for the fault diagnosis of cylindrical roller bearing

Umang Parmar[1]*, D.H. Pandya[2]

[1]**Department of Mechanical Engineering, LDRP Institute of Technology and Research, Gandhinagar, INDIA*
[2] *Department of Mechanical Engineering, LDRP Institute of Technology and Research, Gandhinagar, INDIA*
Corresponding Author: e-mail: umangparmar138@gmail.com, Tel +91-9687965686
ORCID iDs: http:/orcid.org/0000-0003-4339-5121 (Parmar); http:/orcid.org/0000-0002-1619-6170 (Pandya)

Abstract

In this paper, a comparative study of the Artificial Intelligence (AI) techniques for the condition monitoring of the cylindrical roller bearing is presented. For the feature selection, wavelet analysis is applied using the 'sym2' as the mother wavelet. Nine features are considered for the training and evaluation of the AI techniques and then effectiveness is compared. Bearing sample data consists of four different conditions as having defective inner race, defective outer race, having defects on roller and a healthy bearing. For the preparation of the sample bearing, laser machine is used for introduction of the micro size defects on the surfaces. Support Vector Machine (SVM), Artificial neural network (ANN), and logistic regression are used with feature ranking method for the data training purpose and their effectiveness of identifying the condition is the major purpose. Feature ranking method is the new way of filtering the right data in right sequence for the data training. In results, Logistic regression found more accurate in comparison with the ANN and SVM for the cylindrical bearing.

Keywords: Feature Extraction, SVM, ANN, Logistic regression, Wavelet Analysis.

1. Introduction

Cylindrical roller bearing has its own advantage of carrying high radial load in any rotating machineries. Hence it is widely used in the industries. Vibrational measurement is one of the methods of in the condition monitoring which aims to define the condition of the bearing when it is in the operation. Early detection of the bearing failure using condition monitoring decreases the chances of the shutdown and reduces the cost of maintenances and restricts production loss. Signals taken from both healthy conditions bearing and from the defective bearings are analyzed based on their vibrational parameter responses and training of the data helps in the generation of the model which auto detect the condition of bearing. So many researchers have tried different methods and techniques for making the auto fault detection in the field of bearing.

Vibration response for the cylindrical bearing using the mathematical model and computational analysis is observed different at the different rpm range. At low speed (1000 rpm) it shows chaotic nature which turns into the periodic with increase in the speed (around 2000 rpm). At higher rpm (more than 5000 rpm) it shows the sub and super harmonic effect (Yadav et al., 2017). System for the auto feature extraction for the bearing fault is proposed using the signal processing based on the concept of wavelet. Identifying the similarity between the signal taken from the bearing in operation and the predefined wavelet function is the key

aspect in this methodology. Based on the similarity index the best suitable mother wavelet is identified and used for the fault detection (Rafiee et al., 2010).

In the process of the finding the suitable mother wavelet for the given bearing, Minimum Shannon Entropy value is suggested and based on that wavelet selected for the feature extraction. After that statistical data is trained using three AI techniques in which two methods are of supervised machine learning (Support Vector Machine- SVM, learning vector quantization and one is an unsupervised machine learning technique named self-organizing map. Among all three, SVM found as the best method for the ball bearing (Kankar et al., 2011). ANN technique with the Multilayer perceptron and additionally feature selection based on the correlation is applied along with support vector method (SVM) technique with sequential minimal optimization for the ball bearing fault diagnosis and it proved that performance of the ANN (93.5%) is better than the SVM in the classification of ball bearing defects (Ankit Darji et al.,).

A unique method of Discrete wavelet transform (DWT) in the fault detection is also useful because of the impulse response of the signals. It was noted that the response of the impulse is periodically over the time span and that period of repetition gives a lot of information regarding the presence of defects. DWT as a tool worked in the ball bearing and gave clear bifurcation between outer race and inner race defects (Prabhakar et al., 2002). A new approach of continuous wavelet transform is also given for pattern recognition in the bearing fault detection. Using the fact of every defect and healthy bearing are having unique vibration signatures and classifying them can help in the fault detection is the new approach (Saxena et al., 2016). In the comparative study of the Fast Fourier Transform, Continues Wavelet Transform and Wavelet Packet Transform (WPT) using the maximum energy criteria, WPT stands out as best technique for the ball bearing. (Pandya et al., 2015).

In vibrational analysis, few pre-processing signal analysis methods like wavelet transform of the vibrational signal and finding out the optimize features for the feature extraction is also studied and found very effective before applying the ANN technique. One proposed procedure state that only limited features from all available features can give better accuracy. When reduction of input data occurs, training time decreases in the ANN and in few iterations the require output can be achieved (Samanta and Al-Balushi, 2003). New concept of pseudo-intrinsic mode functions IMF introduced for the understanding of the effects IMF at various resonant frequency bands. Hilbert–Huang transform technique at the resonant frequency band with the concept of spectral kurtosis is new tool in the analysis. This method has an advantage over the conventional method of selecting envelope for analysis. Signals having non-linear and non-stationary behavior is easily analyzed using the IMF2 and IMF3 compared to IMF1 (Darji et al., 2020).

A flexible analytical wavelet transforms (FAWT) and discrete wavelet transforms as a decomposition technique compared for the classification of the ball bearing conditions the three different techniques: Random Forest (RF), K-Nearest Neighbors (KNN) and SVM. It was found that FAWT is superior with SVM when feature selection was done using Permutation entropy (Sharma et al., 2021). In the self-aligned ball bearing, when KNN and ANN was compared based on wavelet transform, result stated that the ANN shows better accuracy in the classification then the KNN (Gunerkar et al., 2019). In one more comparative analysis of the ANN and SVM using the CWT, Morlet wavelet is used as mother wavelet based on the energy and entropy data for ball bearing and result shows that SVM gives better performance in classification because of its generalization capability (Agrawal and Jayaswal, 2020).

This research work is aimed towards selection of the right wavelet based on the effective features which are extracted after the decomposition of the signal at level 7 using the WPT. Here maximum energy to entropy ratio and maximum Shannon entropy are considered for deciding the best suitable mother wavelet. Comparison of three techniques names Multinomial logistic regression (MLR), ANN and SVM are presented and effectiveness of selected features on model training is calculated,

2. Machine learning techniques

Faulty bearing in the operation leads to the big shut down and damage because of it can be very large. Every time manual or visual inspection can satisfy the need and risk of breakdown increases with the time. So, need of some special method or technique arise which give flexibility in terms of some prediction of the faults. One of these kinds of method is Supervised Machine learning were using some of the input data, an output can be predicted based on the pattern of input. Here data training plays a very important role as model is generated based on the training data only. In the bearing fault detection, major types of failures are well defined so that in the output knowing conditions supervised machine learning techniques is very helpful and quick as well.

In the fault detection of bearings, classification of defects is formulated as neural network in the ANN. After that different networks are trained using the available sample data. Here, number of hidden layers and neurons highly affect the accuracy so defining the right number is very important. Accuracy of the network is defined as the ratio of number of the samples correctly classified to the total number of samples (Heo and Lee, 2018). In ball bearing, 93% correct classification is achieved in the two hidden layers and twelve input nodes using back propagation algorithm in the ANN where energy and kurtosis were the main features extracted. (Pandya et al., 2012).

Support vector machine initially made for the two-class bifurcation and then it was modified for the more than 2 classes. This method is based on the kernel function in which any inputs like n_1, n_2, n_3.. are classified based on the distance between them. This distance calculation is from the hyperplane and it helps in the transforming any input data in the high-dimensional map. The main advantage of the method is the less effort in the computation and less timing in the training. Kurnel trick is the key feature behind

this method. So many kernels are developed for making the different shapes hyperplanes form the variety of the input data. It is very clear that the limitation of linear hyperplane is overcomes by these different shapes hyperplanes (Hsu et al., 2002).

Multinomial logistic regression is a technique which categorized dependent variables into any required output levels. This is like the logistic regression but here multiple classes in outcomes can be achieved. Any output class in MLR considered as the state and probability of falling the input in that state is calculated for categorization of the data. In the ball bearing, heathy and faulty condition are identified by this method and results shows that 99% accuracy in the healthy condition and 88% in the fault identification can be achieved using the prediction performance index (Kerdsiri, 2007).

3. Signal Processing

3.1 Data Collection: In this research, cylindrical roller bearing is used which is subjected to the rotation of minimum 200 rpm to maximum 3000 rpm using the setup shown in the figure 1. Solid concreate foundation is prepared and over that mild steel plate is bolted. To avoid the external noise and environmental effects in terms of vibration, the whole setup is isolated. Over the shaft, at the one end heathy cylindrical roller bearing is kept and at the other end, four different sets: one healthy and three defective bearings are placed for data collection. Shaft is connected to the DC motor and speed is controlled by the servo controller. This setup has cylindrical bearing NJ 305 and specification of it is shown in the table. Three defective bearing are consisting of defects on inner race, outer race and roller which is made using the laser cutting operation.

For the vibration analysis, signals are taken using CoCo-80. It is a portable recorder having feature of dynamic signal analysis and it collects and store vibration data. Signals taken from the recorded is then analyzed using signal processing technique called wavelet transform.

Figure 1. Experimental set up

Table 1. Details of Cylindrical bearing

Bearing	Cylindrical Roller bearing NJ 305
No of roller	10
Bore Diameter	25 mm
Out diameter	62 mm
Width	17 mm
Weight	0.25 kg
Number of Rows	1
Static and Dynamic load rating	36.5 and 46.5 kN

3.2 Data Preparation: Every signal has a data in terms of the time domain and frequency domain. When time domain data is taken out, it gives only amplitude of the signal at that time and when frequency domain data is considered it gives the overall idea about how amplitude is changing. Bearing fault's signals having data in terms of the big change in amplitude at intervals so time-frequency domain techniques only can satisfy the need. Wavelet packet transform is of the time-frequency domain technique which is widely used in the bearing fault diagnosis. As wavelet itself is a signal with zero mean and satisfying the predefined mathematical function, so selecting a proper wavelet for given bearing is a really challenging task.

Mother wavelet selection method consists of some criterions and here in this research two of them are followed. One is the Maximum Energy to Shannon Entropy and second one is the Maximum relative wavelet energy. Firstly, all the input data are decomposed at the different level according to their natural frequency calculation and then after considering level 7 of decomposition using spectral kurtosis concept, calculation is made for both criterions. Results have shown that the out of 7 selected wavelets (db4, db10, db44, Rbio5.5, Sym2, coif5, dmey), Sym 2 is the best mother wavelet for cylindrical bearings (Parmar and Pandya, 2020).

Once mother wavelet selected then further data processing can be done after the denoising of the signals with the mother wavelet. So here sym2 is used for the denoising of all available sample data and in the next step data processing is done on it using different artificial intelligence techniques.

3.3 Data Processing: This is the one important phase in AI, which starts with the feature extraction then it goes for the optimum feature selection and at last normalized data goes for the training. Whole procedure in detail is explained step by step below:

1. As Sample data, 15 readings for 15 different speeds and for 4 different conditions are taken, these are further denoised at level 7 by sym2 wavelet.
2. From the denoised signals, ten features are extracted as statistical data which are x and y axis kurtosis and energy value at level 7 decomposition sub band, energy of signal, mean value, standard deviation, energy of signal, skewness, kurtosis of signal, crest factor, max value and mean squared value.
3. Level 7 decomposition sub-band count is $2^7 = 128$, so total $128 \times 4 = 512$ (x axis and y axis kurtosis and energy) and 8 other features so total 520 features are normalized from 0 to 1.
4. Normalized data then subjected to the ANN, SVM and MLR model training.
5. Multiple time data is trained by taking different combination of the features and according to that optimized featured are identified.
6. For the identified features, Result is compared for all models.

Here mentioned all features are statistical parameters of the signal and values of their coming from the predefined equations. Kurtosis (K) is a statistical parameter of the degree at which the ends of any signal distribution depart from the normal distribution. It is denoted by K and measured using the Equation (1) for N number of samples of signal x_j.

$$K = \frac{\frac{1}{N} \sum x_j^4}{\left(\frac{1}{N} \sum x_j^2\right)} \tag{1}$$

Very similar term to this is skewness which is defined as the Skewness measures as the lack of symmetry in comparison with the normal distribution and the value of the skewness is calculated using Equation (2).

$$Skewness = \frac{\frac{1}{N} \sum_{j=1}^{N} (x_l - \bar{x})^3}{(\frac{1}{N} \sum_{j=1}^{N} (x_l - \bar{x})^2)^3} \tag{2}$$

For the same signal x_j, Signal Energy is ultimately measure of signal strength and can be find using the below equation, For the same signal x_j, Signal Energy is ultimately measure of signal strength and can be find using the below Equation (3).

$$E = \sum_{j=-\infty}^{\infty} \left| x_j^2 \right| \tag{3}$$

The mean value, M is the value around which the signal is centered over the period as per the Equation (4) and in place of the value of x_j if square of the x_j is considered than it becomes the mean square value (MSV) as shown in the Equation (5). Apart from these, Maximum value of the any signal is considered as the largest and smallest values that a signal takes on some period N_1 and N_2.

$$M = \frac{1}{N_2 - N_1 + 1} \sum_{j=N_1}^{N_2} \left| x_j \right| \tag{4}$$

$$MSV = \frac{1}{N_2 - N_1 + 1} \sum_{j=N_1}^{N_2} \left| x_J^2 \right| \tag{5}$$

The standard deviation (S.D) is a common measurement of the development of a distribution. The formula is given below as Equation (6).

$$S.D = \sqrt{\frac{\sum_{j=1}^{N} (x_l - \bar{x})^2}{N-1}} \tag{6}$$

One more feature here considered, which is ratio of the Peak value and Root Mean Square, called as the Crest Factor (C.F) and Equation (7) is used to find the value of it.

$$C.F = \frac{Max|x_j|}{\sqrt{\frac{1}{N}\sum_{j=1}^{N}|x_J^2|}}$$

(7)

Using all the above equations, the data set is prepared for the application of the ranking feature method and the results of all three techniques are discussed in the this research work.

4. Results and Discussion

In this research work, features extraction for the cylindrical rolling bearing is done with the sample size of 900 signals and training is carried out for the ANN, SVM and MLR model. Here input data having the 520 features and 900 signals for 15 different speeds (At an interval of 200 RPM between 200 to 3000 RPM). Total 520 features are taken in the consideration and feature wise effect of them is calculated. In the table 2, order no is given as per the accuracy of each feature and after that according to the given order number all models are again trained and accuracy of it is given in the fig 2.

Table 2. Ordering the features as per classification accuracy

Feature Name	No of data	ANN Accuracy	Order No	SVM Accuracy	Order No	MLR Accuracy	Order No
Kurtosis X-Axis (Level 7)	128	41.7	7	78.2	5	75.2	3
Kurtosis Y-Axis (Level 7)	128	62.4	3	85.3	2	70.4	4
Energy X-Axis (Level 7)	128	71.8	2	84.2	3	79.3	2
Energy Y-Axis (Level 7)	128	82.3	1	89.4	1	84.3	1
Mean value	1	32.6	12	51.4	12	30.7	12
Std deviation	1	47.4	4	73.4	6	65.1	5
Energy	1	45.7	5	81.7	4	61.2	6
Skewness	1	39.6	9	71.3	7	45.3	9
Kurtosis of signal	1	34.4	10	59.4	10	38.3	10
Crest factor	1	32.7	11	55.3	11	35.4	11
Max value	1	40.9	8	68.3	8	49.2	8
Mean-squared value	1	42.9	6	63	9	56.8	7
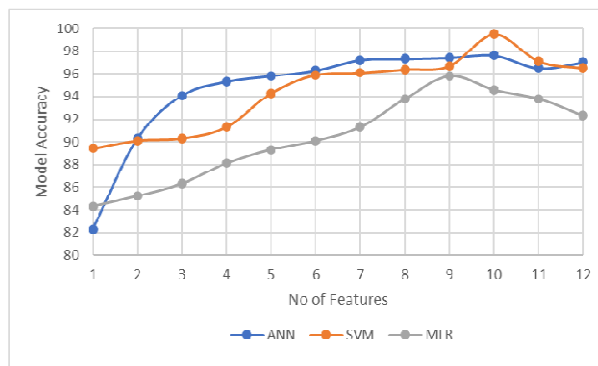
Figure 2. Model accuracy vs no of features

The model accuracy for ANN and SVM both are highest for top 10 ordered features and for MLR it is top 9 features. At the end all three model are again trained for their best ordered featured and class wise results is shown in the table 3. Here all conditions having 225 sample data and four conditions of bearing so total 900 sample data are trained.

Table 3. Classification of the Sample data using AI Technique.

AI Technique	ANN	SVM	MLR
Bearing condition (sample size)	Rightly classified No of Sample Data		
Healthy (225)	222	225	219
ORD (225)	221	224	217
IRD (225)	215	223	212
RD (225)	220	224	216
Total (900)	878	896	864
Accuracy (%)	97.6	99.6	96.0

5. Conclusion

In the present research work, Sym2 is taken as mother wavelet for the cylindrical bearing's auto fault diagnosis. Feature selection is based on the ordered number and overall classification accuracy of each features. Out of the 12 selected features, ANN and SVM is finally trained according to the top 10 effective features where MLR is trained according to top 9 ordered features. Final accuracy of the top ten ranked features in ANN is 97.6 %, for SVM it is 99.6 % and MLR shows 96% accuracy. Comparison of the previous research (Umang Parmar. 2021) for the same set of data without ranking feature method shows that ANN is 96% accurate in the training but not so good in the validation and testing while SVM shows the 95.6% efficiency which proves that ranking feature method helps in getting the good results. From the results, it is concluded that SVM showing better performance in the fault diagnosis compared to the ANN and MLR.

Nomenclature

AI	Artificial Intelligence
SVM	Support Vector Machine
ANN	Artificial Neural Network
RPM	Revolution Per Minute (Unit of the speed)
DWT	Discrete Wavelet Transform
WPT	Wavelet Packet Transform
IMF	Intrinsic Mode Functions
FAWT	Flexible Analytical Wavelet Transforms
RF	Random Forest (classifier)
KNN	K-Nearest Neighbors (classifier)
MLR	Multinomial logistic regression
X_j	Mathematical Symbol for the signal
K	Kurtosis
N	Variable (defining sample size)
E	Energy
M	Mean Value
MSV	Mean Square Value
SD	Standard Deviation
CF	Crest Factor

Acknowledgement

References

Agrawal P., Jayaswal P., 2020, Diagnosis and classifications of bearing faults using artificial neural network and support vector machine. *Journal of Institution of Engineers of India*, Series C., Vol. 101, pp. 61–72. https://doi.org/10.1007/s40032-019-00519-9

Darji A., Darji P.H., Pandya D.H., 2019, Fault Diagnosis of Ball Bearing with WPT and Supervised Machine Learning Techniques. In: Tanveer M., Pachori R. (eds) *Machine Intelligence and Signal Analysis*. Advances in Intelligent Systems and Computing, Vol 748. Springer, Singapore. https://doi.org/10.1007/978-981-13-0923-6_25

Darji A. A., Darji P. H., and Pandya D. H., 2020, Envelope Spectrum Analysis with Modified EMD for Fault Diagnosis of Rolling Element Bearing. In: Gupta V., Varde P., Kankar P., Joshi N. (eds) *Reliability and Risk Assessment in Engineering*. Lecture Notes in Mechanical Engineering. Springer, Singapore. https://doi.org/10.1007/978-981-15-3746-2_8

Gunerkar R.S., Jalan A.K. & Belgamwar S.U., 2019, Fault diagnosis of rolling element bearing based on artificial neural network. *Journal of Mechanical Science and Technology*, Vol. 33, pp. 505–511. https://doi.org/10.1007/s12206-019-0103-x

Heo S., Lee J.H., 2018, Fault detection and classification using artificial neural networks. *IFAC-PapersOnLine*, Vol. 51, No. 18, pp. 470–475. https://doi.org/10.1016/j.ifacol.2018.09.380

Hsu C.W., Lin C. J., 2002, A comparison of methods for multi-class support vector machines. *IEEE Transaction on Neural Networks* Vol. 13, No. 2, pp. 415–425. DOI: 10.1109/72.991427

Kankar P.K., Sharma S.C., Harsha S.P., 2011, Rolling element bearing fault diagnosis using wavelet transform, *Neurocomputing*, Vol. 74, pp. 1638–1645. https://doi.org/10.1016/j.neucom.2011.01.021

Kerdsiri T., Gullayanon R., 2017, Early fault detection based on ball bearing vibration analysis using multinomial logistic regression. *Proceedings of the 2017 International Conference on Intelligent Systems, Metaheuristics & Swarm Intelligence - ISMSI '17*, pp. 152–156. doi:10.1145/3059336.3059367

Kumar H.S., Pai P.S., Sriram N.S., Vijay G.S., 2013, ANN based evaluation of performance of wavelet transform for condition monitoring of rolling element bearing. *Procedia Engineering*, Vol. 64, pp. 805–814. https://doi.org/10.1016/j.proeng.2013.09.156

Parmar U., Pandya D.H., 2020, Selection of mother wavelet for fault detection of cylindrical bearing, *Vibration and Acoustics Research Journal*, Vol. 2, No. 1, pp. 13-26. DOI: 10.13140/RG.2.2.11094.57928

Parmar U., Pandya D.H., 2021, Experimental investigation of cylindrical bearing fault diagnosis with SVM, *Materials Today: Proceedings*, Vol. 44, No. 1, pp. 1286-1290. https://doi.org/10.1016/j.matpr.2020.11.327

Pandya D.H., Upadhyay S.H. and Harsha S.P., 2015, Fault diagnosis of high-speed rolling element bearings using wavelet packet transform, *International Journal of Signal and Imaging Systems Engineering*, Vol. 8, No. 6, pp. 390–401. https://doi.org/10.1504/IJSISE.2015.072922

Pandya D.H., Upadhyay S.H., Harsha S.P., 2012, ANN based fault diagnosis of rolling element bearing using time-frequency domain feature, *International Journal of Engineering Science and Technology*, Vol. 4, No. 6, pp. 2878-2886.

Prabhakar S., Mohanty A.R., Sekhar A.S., 2002, Application of discrete wavelet transform for detection of ball bearing race faults, *Tribology International*, Vol. 35, pp. 793–800. https://doi.org/10.1016/S0301-679X(02)00063-4

Rafiee J., Rafiee M.A., Tse P.W., 2010, Application of mother wavelet functions for automatic gear and bearing fault diagnosis, *Expert Systems with Applications*, Vol. 37, pp. 4568–4579. https://doi.org/10.1016/j.eswa.2009.12.051

Samanta B. and Al-Balushi K.R., 2003, Artificial neural network based fault diagnostics of rolling element bearings using time-domain features, *Mechanical Systems and Signal Processing*, Vol. 17, No. 2, pp. 317–328. https://doi.org/10.1006/mssp.2001.1462

Saxena M., Bannet O.O., Gupta M., Rajoria R.P., 2016, Bearing fault monitoring using CWT based vibration signature, *12th International Conference on Vibration Problems*, ICOVP, *Procedia Engineering*, Vol. 144, pp. 234–241. https://doi.org/10.1016/j.proeng.2016.05.029

Sharma S., Tiwari S.K., Singh S., 2021, Integrated approach based on flexible analytical wavelet transform and permutation entropy for fault detection in rotary machines, *Measurement*, Vol. 169, 108389, ISSN 0263-2241. https://doi.org/10.1016/j.measurement.2020.108389

Yadav H. K. and Pandya D. H., 2017, Non-linear dynamic analysis of cylindrical roller bearing, *11th International Symposium on Plasticity and Impact Mechanics*, IMPLAST, *Procedia Engineering*, Vol. 173, pp. 1878–1885. https://doi.org/10.1016/j.proeng.2016.12.242

Biographical notes

Umang Parmar received his BE degree in Mechanical Engineering from Indus Institute of Technology, Ahmedabad. He did his masters in production engineering from the LDRP Institute of Technology and Research, Gandhinagar and doing his PhD in Fault diagnosis of cylindrical bearing with localized defects using soft computing technique.

D. H. Pandya did his doctorate from IIT, Roorkee. His areas of interest are vibration, condition monitoring, signal processing, expert system, and machine design. He had more than 65 publications in referred journals and international conferences.