Section A-Research paper
ISSN 2063-5346

# Identification of Faults for Centrifugal Pump MODWPT and SVMA

**Mit Shah[1], Pina Bhatt[2], Keval Bhavsar[2,*], Umang Parmar[2,†]**

*[1]Department of Mechanical Engineering, Centre for Research, Silver Oak University, Ahmedabad, Gujarat – 382481, India*
*[2]Department of Mechanical Engineering, Aditya Silver Oak Institute of Technology, Silver Oak University, Ahmedabad, Gujarat – 382481, India*

## Abstract

Centrifugal pumps are very essential components for fluid transfer in the industry and in self-priming centrifugal pumps, the bearing is the key element. The performance of the pumps highly depends on the bearing conditions and high-performance leads to smooth operation and lesser downtime. But with time when defects introduce in the bearing, the detection of the type of defect is the major challenge. Conventional ways of detecting defects lead to high maintenance costs and time. For the detection of defects in the bearing, vibration signal analysis is widely used in the last decade and this research also focuses on the Signal processing techniques for such as Discrete Wavelet Transform (DWT) and Modulated Wavelet Packet Transform (MODWPT). Along with this, soft computing techniques, including Ensemble Bagging Tree Algorithm (ETA), Decision Tree Algorithm (DTA) and Support Vector Machine Algorithm (SVMA) are compared to make the classification more intelligent and to make this approach adaptive for fault detection. Due to the advantages of all these techniques in nonlinear problems, the results are impressive in comparison. A data set is first denoised using signal processing in which four different wavelets are selected and using the best-performing wavelet, thirteen different Statistical Time-Domain Features are extracted for training the models. Results show that the best performance is achieved using MODWPT and SVMA.

**Keywords:** Machine Learning; Fault Identification; Rotating Machine; Centrifugal Pump; Rolling Element Bearing; Support Vector Machine; SVMA; MODWPT; DWT.

[*] Corresponding Author Email Id: kevalbhavsar42@gmail.com
[†] Corresponding Author Email Id: umangparmar138@gmail.com

*Section A-Research paper*
*ISSN 2063-5346*

## 1. Introduction

Rotating machines are indispensable in Industry 4.0 as it provides effective power transmission and also is it helps in complete automation. Through Rotating machines, a highly productive production line is possible to establish. The only concern with these kinds of machines is their cost of maintenance. Due to the digital transformation of industries in the past decade, Now under Industry 4.0, Data is available in massive amounts and through that data, the condition-based monitoring (CBM) of any machine is the new challenge for industries. Major issues like the high downtime of machines and calculating the remaining useful life also become possible to answer with the CBM.

Recent development in the field of Machine Learning (ML) is now making predictive maintenance possible with very high-reliability value [1]. In Industry 4.0, the new path of predictive maintenance is starting from the data collected from all collected devices and then through remote monitoring based on past health condition data, possible machine fault alerts can be generated and automatic maintenance scheduling can also be channelized [2].

Detection of the fault can be done either by model-based techniques or by signal processing. The signal processing is further categorized into three basic types namely- time domain [3], frequency domain [4] and time-frequency domain techniques [5]. Vibration signals in the case of bearing are the source of information for fault detection. The common cause of high vibration from the bearing is either the less lubrication between their parts or any kind of defective internal part itself. During the vibration, the important frequencies which need attention are the inner race frequencies, out race frequencies and ball frequencies. Based on the frequency spectrum analysis, the identification of the bearing fault is possible [6]. There are three basic types of signal – namely stationary, cyclo-stationary and cyclo nonstationary [7]. To understand the response of these frequencies, signal processing is the very essential technique to be implemented and before implementing that, the nature of the signal needs to know. Another way of detecting the defect is the integration of the sensors with the machine element itself and this smart bearing can replace the conventional bearing [8].

The behaviour of the bearing fault is noticed as non-stationary which is always difficult to process. To deal with such problems, wavelet analysis is mostly used by the researcher in the past decade [9], [10]. Some researchers also have explored the Hilbert Transform (HT) in place of wavelet analysis [11].

A novel technique based on the neural network and fuzzy logic is also explored by many researchers. The bearing frequencies are always there in the recorded signals but over that, some noise signals and other part's frequencies might merge with it. In such cases, high pass filters with the right envelope are used [12]. In the recent past, the concept of nonlinear features is used over linear features in soft computing. The best example is to use the permutation entropy to provide better results in comparison to the conventional one [13].

Under machine learning, So many algorithms are explored in the field of fault detection and a few of them like, Artificial Neural Network (ANN) [14], K-nearest neighbour (KNN) [15], Deep Learning [16], Support Vector Machine (SVM) [17], [18] Multinomial Logistic Regression [19] are some which are performing well for the bearing fault detection. Among all the machine learning algorithms, majorly researchers are coming up with the SVM as the very effective algorithm for bearing fault detection [20].

This study (Fig. 1) uses the dataset captured by Chen Lu et al. [21] on self-priming centrifugal pumps from the identification of faults. The captured vibration signals are denoised using two Signal Processing Techniques (SPTs) namely Discrete Wavelet Transform (DWT) and Maximal Overlap Discrete Wavelet Packet Transform (MODWPT) and the mother wavelet is selected using the prominent criterion known as MESE. After the calculation of coefficients from both SPTs, statistical features are extracted. These features are ranked and selected based on the Relief F algorithm to ease the computational power and reduce the size/dimension of the feature vectors. At last there ranked and selected feature vectors are used to train the Machine Learning Algorithms (MLAs) for fault identification.

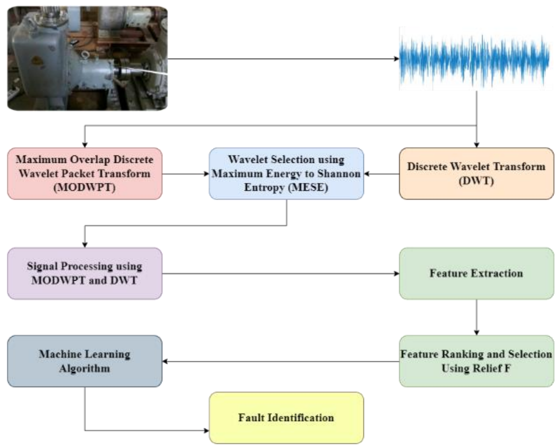



Figure 1: Methodology

## 2. Experimental Setup

This study uses the centrifugal pump dataset (Fig. 2 and Table 1) developed by Chen Lu et al. [21] consisting of vibration signals captured using an accelerometer for Regular Conditions (RC) as well as four faulty conditions of which three are in the bearing namely, Outer Race (ORF), Rolling Element (REF), and Inner Race (IRF) and the last one in the Impeller (IF).

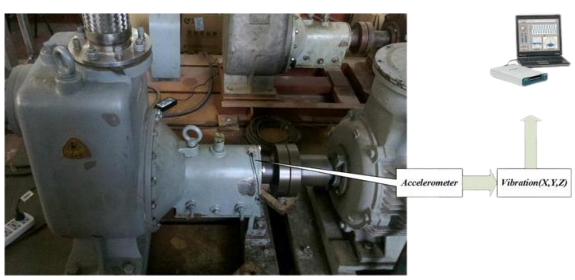



Figure 2: centrifugal pump setup for data acquisition

Table 1: Parameters used for data acquisition

| Parameter | Value |
|---|---|
| Sampling Rate | 10239 Hz |
| Rotational Speed | 2900 RPM |
| Acquisition Time | 2 seconds |

## 3. SPTs

The signals are pre-processed to increase the total number of instances for each condition to 100. Two prominent SPTs techniques namely, MODWPT and DWT are used to denoise the acquired vibration signals. DWT is an SPT developed by Daubechies [22], and Mallat [23]–[25] in the late 1980s to extract the Detail Coefficient (DC) (using a high pass filter) and Approximate Coefficient (AC) (using a low pass filter) consisting of the information of signal for that frequency band. The frequency band is calculated based on the decomposition level identified using the FFT analysis and various Fault Frequencies like Roller Spin Frequency. The output of the DWT (Fig. 3) shows that DCs are extracted at each level of decomposition while the ACs are extracted at the last level of decomposition. The MODWPT [26] (Fig. 4) uses the same method for the generation of coefficients but both the coefficients are extracted at all the levels of decomposition and hence more information is extracted from the signal.

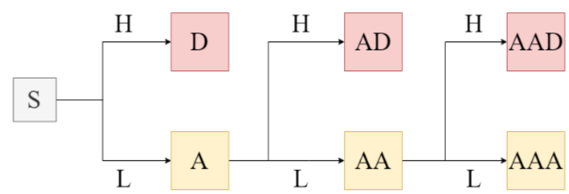



Figure 3: DWT

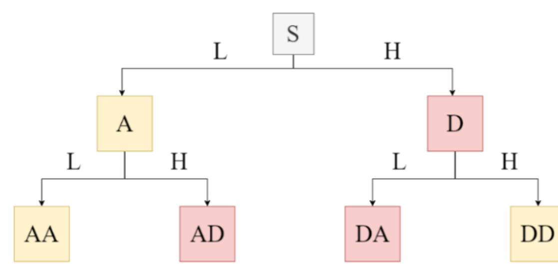



Figure 4: MODWPT

## 4. Feature Extraction, Ranking and Selection

The statistical features mentioned in Table 2 are extracted from the calculated coefficients of MODWPT and DWT to analyse the amplitude change and the distribution of energy and develop the feature vectors [27]. These features are then used as input to the feature ranking and selection model Relief F developed in the year 1992 by Kira and Rendell [28], [29]. It gives scores or more commonly known as weights based on the quality or the relevancy of the feature to the class [30]. For calculating the weights of the feature, at first random features are used to predict the classes using the nearest neighbour and the features that distinctively predict the classes are given more weights. The top 46 and 10 ranked features are selected for feature vectors developed using MODWPT and DWT respectively that are used as input to MLAs.

Table 2: Statistical Time-Domain Features

| **Name** | | |
|---|---|---|
| Impulse Factor | Shape Factor | Crest Factor |
| Minimum Amplitude | Maximum Amplitude | Skewness |
| Kurtosis | Root Mean Square | Peak 2 Peak Value |
| Sum | Variance | Standard Deviation |
| Mean | | |

## 5. Machine Learning Algorithms (MLA)

In this study, popular MLAs namely Ensemble Bagging Tree Algorithm (ETA), Decision Tree Algorithm (DTA) and Support Vector Machine Algorithm (SVMA) are used for the prediction of various faults based on the features extracted. At last, the MLAs are evaluated using four performance evaluators that are Precision (*Pr*), Recall (*Re*), Accuracy (*Ac*) and F1 Score ($F_1S$) (Table 3) [31].

Table 3: Performance Evaluators

| **Parameter** | **Formula** |
|---|---|
| Precision (*Pr*) | $\frac{TrPo}{TrPo + FaPo}$ |
| Recall (*Re*) | $\frac{TrPo}{TrPo + FaNe}$ |
| Accuracy (*Ac*) | $\frac{TrPo + TrNe}{TrPo + TrNe + FaPo + FaNe}$ |
| F1 Score ($F_1S$) | $\frac{2 \times Pr \times Re}{Pr + Re}$ |

Where, *Fa* means False, *Tr* means True, *Ne* means Negative and *Po* means Positive.

### ETA

In the year 1996, Breiman proposed an algorithm where multiple trees are combined to develop an ensemble for increasing the stability of the algorithm [32], [33]. The algorithm at first divides the feature vector into multiple samples in a randomised manner. This randomly sampled feature vector is used to train various trees, followed by the generation of a single ensemble by aggregating the predictions of all the trees (Fig. 5). This algorithm is developed to reduce the error as various trees used in this algorithm cannot have the same classification errors [34], [35].

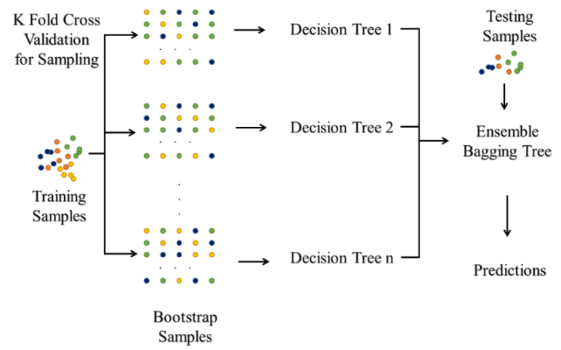



Figure 5: ETA

### DTA

In the year 1986, Quinlan proposed an algorithm which looks like the inverted tree (Fig. 6) [34]. The resemblance to the natural tree is given as the root of the tree from where the initial split is executed is called the root node. The branches where another split happens is called splitting node. The last part of any tree is the leaf, at this location, the predictions are made and this node is known as Leaf Node.

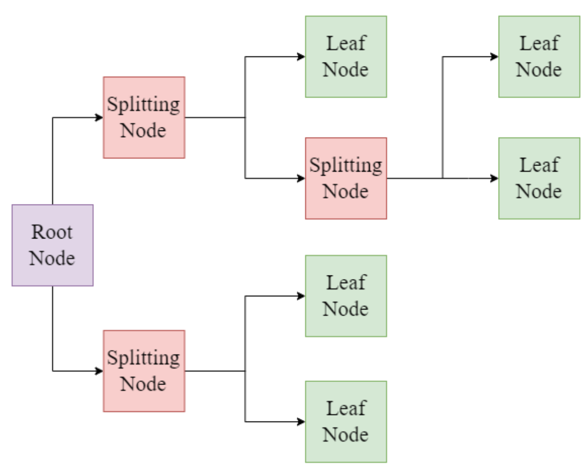



Figure 6: DTA

### SVMA

In the year 1995, Cortes and Vapnik developed an algorithm which plots the whole dataset in an n-dimensional area (n is the total number of features),

*Section A-Research paper*
*ISSN 2063-5346*

followed by the generation of boundaries to separate each class. At last, a hyperplane is constructed in such a manner that the distance between the boundary of the classes can be maximised (Fig. 7) [34], [36], [37].

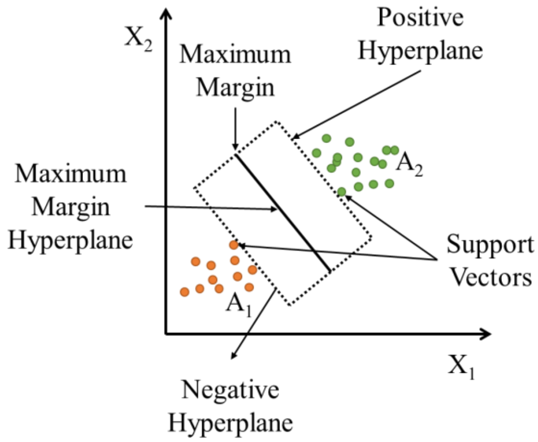
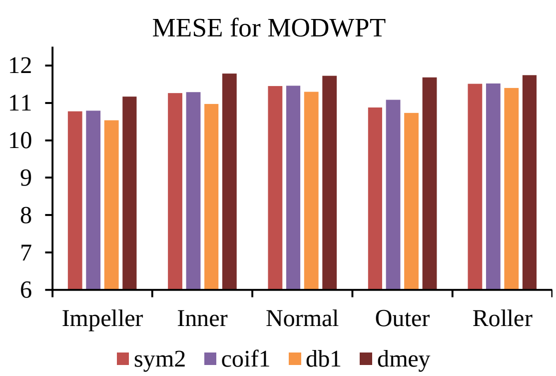



Figure 7: SVMA

## 6. Result and Discussion

In this study, two WTs are used MODWPT are used to remove the noise from the acquired vibration signals. Selection of the mother wavelet is one of the important steps for conduction of any WT and the criteria used in this study is the ratio of Energy to Shannon Entropy developed by Yan and Gao [38]. Four wavelets namely, Discrete Meyer (dmey), Symlet 2 (sym2), Daubechies 1 (db1) and Coiflet (coif1) are evaluated using the MESE equation [39]. Figure 8 shows that the highest value of MESE for MODWPT and DWT is achieved with dmey and db1 respectively.



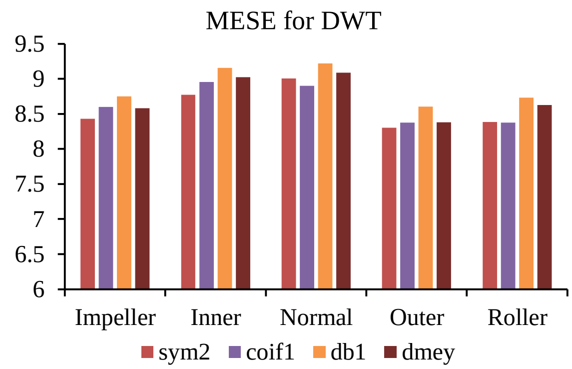




Figure 8: MESE graphs for MODWPT and DWT

The vibration signal is then processed using the above-mentioned mother wavelets for both the signal processing techniques i.e. MODWPT and DWT. The calculated coefficients from both techniques are used to extract the features mentioned in Table 2. The size of the developed feature vectors consisting of the features is huge and hence to reduce the size/dimension of the feature vector Relief F is used. The total number of features selected after ranking is 46 and 10 ranked features are selected for feature vectors developed using MODWPT and DWT respectively.

Tables 4 and 5 show the result of the performance evaluators used for evaluating the classification performance of the MLAs. It can be seen from Table 4 that the SVMA performs better compared to other MLAs with a 1 or 100% as the value of all the performance evaluators and based on Table 5 it can be stated that ETA performs better when compared with other MLAs with a Recall and $F_1S$ of 0.878, Precision of 0.877, and Ac of 87.8 %. Hence, the best MLAs of each signal processing technique are used to plot the confusion matrix.

Table 4: Performance of MLAs with MODWPT

| MLA | *Pr* | *Re* | *F₁S* | *Ac* |
|---|---|---|---|---|
| **SVMA** | **1** | **1** | **1** | **100 %** |
| DTA | 0.982 | 0.982 | 0.982 | 98.2 % |
| ETA | 0.986 | 0.986 | 0.986 | 98.6 % |

Table 5: Performance of MLAs with DWT

| MLA | *Pr* | *Re* | *F₁S* | *Ac* |
|---|---|---|---|---|
| SVMA | 0.849 | 0.848 | 0.845 | 84.8 % |
| DTA | 0.821 | 0.822 | 0.821 | 82.2 % |
| **ETA** | **0.877** | **0.878** | **0.878** | **87.8 %** |

*Section A-Research paper*
*ISSN 2063-5346*

The confusion matrix shown in Fig. 9 (a) show that SVMA with MDWPT shows that it clearly predicts all the faults as well as regular working condition correctly. Based on Fig. 9 (b) it can be stated that it only correctly predicts the IRF while mis predicts twenty-three instances of IF to four, twelve and seven instances of RC, ORF and REF respectively. It also mispredicts nine instances of RC to three instances of IF, one instance of ORF and five instances of REF. Further, twelve and fifteen instances of ORF and REF respectively are mispredicted to ten and seven instances of IF, two and seven instances of RC, and one instance of ORF.

| SVMA MODWPT | IF | IRF | RC | ORF | REF |
|---|---|---|---|---|---|
| IF | 100 | 0 | 0 | 0 | 0 |
| IRF | 0 | 100 | 0 | 0 | 0 |
| RC | 0 | 0 | 100 | 0 | 0 |
| ORF | 0 | 0 | 0 | 100 | 0 |
| REF | 0 | 0 | 0 | 0 | 100 |

(a)

| ETA DWT | IF | IRF | RC | ORF | REF |
|---|---|---|---|---|---|
| IF | 77 | 0 | 4 | 12 | 7 |
| IRF | 0 | 100 | 0 | 0 | 0 |
| RC | 3 | 0 | 91 | 1 | 5 |
| ORF | 10 | 0 | 2 | 86 | 2 |
| REF | 7 | 0 | 7 | 1 | 85 |

(b)

Figure 9: MESE graphs for MODWPT and DWT

## 7. Conclusion

In this paper, fault detection of bearing in self-priming centrifugal is presented. The data set considered for the current study is of a self-priming centrifugal pump which has five different conditions of the bearing. The presented methodology aims towards comparing the two different wavelet analyses namely DWT and MODWPT. For the selection of the right wavelet among the four selected wavelets, the maximum energy to Shannon entropy criteria is applied and results show that in the case of DWT, db1 is performing better while in the case of the MODWPT, Dmey is performing better in comparison. Further for both the analysis, the best mother wavelet is chosen for extracting the thirteen different features for training the model under the ETA, DTA and SWMA. The feature selection and ranking are done based on the Relief F algorithm. The outcome of this comparative study concludes that for DWT, ETA is giving the best result of 87.8% accuracy while the combination of MODWPT and SVMA has achieved 100% overall accuracy in classifying all five different conditions of the bearing for the selected data set.

## 8. Reference

[1] T. P. Carvalho, F. A. A. M. N. Soares, R. Vita, R. da P. Francisco, J. P. Basto, and S. G. S. Alcalá, ‘A systematic literature review of machine learning methods applied to predictive maintenance’, *Comput. Ind. Eng.*, vol. 137, p. 106024, Nov. 2019, doi: 10.1016/j.cie.2019.106024.

[2] Z. M. Çınar, A. Abdussalam Nuhu, Q. Zeeshan, O. Korhan, M. Asmael, and B. Safaei, ‘Machine Learning in Predictive Maintenance towards Sustainable Smart Manufacturing in Industry 4.0’, *Sustainability*, vol. 12, no. 19, p. 8211, Oct. 2020, doi: 10.3390/su12198211.

[3] J. Prakash Kumar, P. S. Chauhan, and P. Prakash Pandit, ‘Time domain vibration analysis techniques for condition monitoring of rolling element bearing: A review’, *Mater. Today Proc.*, vol. 62, pp. 6336–6340, 2022, doi: 10.1016/j.matpr.2022.02.550.

[4] Y.-T. Su and S.-J. Lin, ‘On initial fault detection of a tapered roller bearing: Frequency domain analysis’, *J. Sound Vib.*, vol. 155, no. 1, pp. 75–84, May 1992, doi: 10.1016/0022-460X(92)90646-F.

[5] Z. Germán-Salló and G. Strnad, ‘Signal processing methods in fault detection in

manufacturing systems', *Procedia Manuf.*, vol. 22, pp. 613–620, 2018, doi: 10.1016/j.promfg.2018.03.089.

[6] Ensor, L. C., and C. C. Feng., 'A Study of the use of Vibration and Stress Wave Sensing for the Detection of Bearing Failure.' ENGINEERING REPORT NO. 75-18 C, Sep. 30, 1975.

[7] D. Abboud, S. Baudin, J. Antoni, D. Rémond, M. Eltabach, and O. Sauvage, 'The spectral analysis of cyclo-non-stationary signals', *Mech. Syst. Signal Process.*, vol. 75, pp. 280–300, Jun. 2016, doi: 10.1016/j.ymssp.2015.09.034.

[8] B. T. Holm-Hansen and R. X. Gao, 'Integrated microsensor module for a smart bearing with on-line fault detection capabilities', in *IEEE Instrumentation and Measurement Technology Conference Sensing, Processing, Networking. IMTC Proceedings*, Ottawa, Ont., Canada: IEEE, 1997, pp. 1160–1163. doi: 10.1109/IMTC.1997.612382.

[9] V. Vakharia, V. Gupta, and P. Kankar, 'A multiscale permutation entropy based approach to select wavelet for fault diagnosis of ball bearings', *J. Vib. Control*, vol. 21, no. 16, pp. 3123–3131, Dec. 2015, doi: 10.1177/1077546314520830.

[10] P. K. Kankar, S. C. Sharma, and S. P. Harsha, 'Rolling element bearing fault diagnosis using wavelet transform', *Neurocomputing*, vol. 74, no. 10, pp. 1638–1645, May 2011, doi: 10.1016/j.neucom.2011.01.021.

[11] R. Senthil Kumar *et al.*, 'A combined HT and ANN based early broken bar fault diagnosis approach for IFOC fed induction motor drive', *Alex. Eng. J.*, vol. 66, pp. 15–30, Mar. 2023, doi: 10.1016/j.aej.2022.12.010.

[12] P. Bhonde and A. Panchwadkar, 'Automatic detection of bearing faults', *Mater. Today Proc.*, vol. 77, pp. 586–591, 2023, doi: 10.1016/j.matpr.2022.10.078.

[13] A. S. Minhas and S. Singh, 'A new bearing fault diagnosis approach combining sensitive statistical features with improved multiscale permutation entropy method', *Knowl.-Based Syst.*, vol. 218, p. 106883, Apr. 2021, doi: 10.1016/j.knosys.2021.106883.

[14] U. Parmar and Divyang H. Pandya, 'Optimized feature selection for condition-based monitoring of cylindrical bearing using wavelet transform and ANN.', in *Neural Networks, Machine Learning, and Image Processing*, CRC Press, 2022, pp. 55–65.

[15] H. S. Kumar and G. Upadhyaya, 'Fault diagnosis of rolling element bearing using continuous wavelet transform and K- nearest neighbour', *Mater. Today Proc.*, p. S2214785323016863, Apr. 2023, doi: 10.1016/j.matpr.2023.03.618.

[16] S. Shen *et al.*, 'A physics-informed deep learning approach for bearing fault detection', *Eng. Appl. Artif. Intell.*, vol. 103, p. 104295, Aug. 2021, doi: 10.1016/j.engappai.2021.104295.

[17] R. K. Mishra, A. Choudhary, A. R. Mohanty, and S. Fatima, 'Multi-domain Bearing Fault Diagnosis using Support Vector Machine', in *2021 IEEE 4th International Conference on Computing, Power and Communication Technologies (GUCON)*, Kuala Lumpur, Malaysia: IEEE, Sep. 2021, pp. 1–6. doi: 10.1109/GUCON50781.2021.9573613.

[18] U. Parmar and D. H. Pandya, 'Experimental investigation of cylindrical bearing fault diagnosis with SVM', *Mater. Today Proc.*, vol. 44, pp. 1286–1290, 2021, doi: 10.1016/j.matpr.2020.11.327.

[19] D. H. Pandya, S. H. Upadhyay, and S. P. Harsha, 'Fault diagnosis of rolling element bearing by using multinomial logistic regression and wavelet packet transform', *Soft Comput.*, vol. 18, no. 2, pp. 255–266, Feb. 2014, doi: 10.1007/s00500-013-1055-1.

[20] K. Adamsab, 'Machine learning algorithms for rotating machinery bearing fault diagnostics', *Mater. Today Proc.*, vol. 44, pp. 4931–4933, 2021, doi: 10.1016/j.matpr.2020.12.050.

[21] C. Lu, Y. Wang, M. Ragulskis, and Y. Cheng, 'Fault Diagnosis for Rotating Machinery: A Method based on Image Processing', *PLOS ONE*, vol. 11, no. 10, p. e0164111, Oct. 2016, doi: 10.1371/journal.pone.0164111.

[22] I. Daubechies, 'Orthonormal bases of compactly supported wavelets', *Commun. Pure Appl. Math.*, vol. 41, no. 7, pp. 909–996, Oct. 1988, doi: 10.1002/cpa.3160410705.

[23] S. G. Mallat, 'Multiresolution Approximations and Wavelet Orthonormal Bases of L 2 (R)', *Trans. Am. Math. Soc.*, vol. 315, no. 1, p. 69, Sep. 1989, doi: 10.2307/2001373.

[24] S. G. Mallat, 'A theory for multiresolution signal decomposition: the wavelet representation', *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 11, no. 7, pp. 674–693, Jul. 1989, doi: 10.1109/34.192463.

[25] S. G. Mallat, 'Multifrequency channel decompositions of images and wavelet

models', *IEEE Trans. Acoust. Speech Signal Process.*, vol. 37, no. 12, pp. 2091–2110, Dec. 1989, doi: 10.1109/29.45554.

[26] X. Yu, X. Ren, H. Wan, S. Wu, and E. Ding, 'Rolling Bearing Fault Feature Extraction and Diagnosis Method Based on MODWPT and DBN', in *2019 11th International Conference on Wireless Communications and Signal Processing (WCSP)*, Xi'an, China: IEEE, Oct. 2019, pp. 1–7. doi: 10.1109/WCSP.2019.8927859.

[27] K. Bhavsar, V. Vakharia, R. Chaudhari, J. Vora, D. Y. Pimenov, and K. Giasin, 'A Comparative Study to Predict Bearing Degradation Using Discrete Wavelet Transform (DWT), Tabular Generative Adversarial Networks (TGAN) and Machine Learning Models', *Machines*, vol. 10, no. 3, p. 176, Feb. 2022, doi: 10.3390/machines10030176.

[28] K. Kira and L. A. Rendell, 'A Practical Approach to Feature Selection', in *Machine Learning Proceedings 1992*, Elsevier, 1992, pp. 249–256. doi: 10.1016/B978-1-55860-247-2.50037-1.

[29] K. Kira, L. A. Rendell, and others, 'The feature selection problem: Traditional methods and a new algorithm', in *Aaai*, 1992, pp. 129–134.

[30] V. Vakharia, V. K. Gupta, and P. K. Kankar, 'Efficient fault diagnosis of ball bearing using ReliefF and Random Forest classifier', *J. Braz. Soc. Mech. Sci. Eng.*, vol. 39, no. 8, pp. 2969–2982, Aug. 2017, doi: 10.1007/s40430-017-0717-9.

[31] P. Akhenia, K. Bhavsar, J. Panchal, and V. Vakharia, 'Fault severity classification of ball bearing using SinGAN and deep convolutional neural network', *Proc. Inst. Mech. Eng. Part C J. Mech. Eng. Sci.*, vol. 236, no. 7, pp. 3864–3877, Apr. 2022, doi: 10.1177/09544062211043132.

[32] L. Breiman, 'Bagging predictors', *Mach. Learn.*, vol. 24, no. 2, pp. 123–140, Aug. 1996, doi: 10.1007/BF00058655.

[33] M. Patel and M. Gupta, 'Caravan Insurance Customer Profile Modeling with R', in *Data Mining Applications with R*, Elsevier, 2014, pp. 181–227. doi: 10.1016/B978-0-12-411511-8.00007-4.

[34] V. Kotu and B. Deshpande, 'Classification', in *Data Science*, Elsevier, 2019, pp. 65–163. doi: 10.1016/B978-0-12-814761-0.00004-6.

[35] E. Fathi and B. Maleki Shoja, 'Deep Neural Networks for Natural Language Processing', in *Handbook of Statistics*, Elsevier, 2018, pp. 229–316. doi: 10.1016/bs.host.2018.07.006.

[36] C. Cortes and V. Vapnik, 'Support-vector networks', *Mach. Learn.*, vol. 20, no. 3, pp. 273–297, Sep. 1995, doi: 10.1007/BF00994018.

[37] K. Roy, S. Kar, and R. N. Das, 'Selected Statistical Methods in QSAR', in *Understanding the Basics of QSAR for Applications in Pharmaceutical Sciences and Risk Assessment*, Elsevier, 2015, pp. 191–229. doi: 10.1016/B978-0-12-801505-6.00006-5.

[38] R. Yan and R. X. Gao, 'BASE WAVELET SELECTION FOR BEARING VIBRATION SIGNAL ANALYSIS', *Int. J. Wavelets Multiresolution Inf. Process.*, vol. 07, no. 04, pp. 411–426, Jul. 2009, doi: 10.1142/S0219691309002994.

[39] Umang Parmar and Divyang H Pandya, 'Selection of Mother Wavelet for Fault detection of cylindrical bearing', *Vib. Acoust. Res. J.*, vol. 2, no. 1, pp. 13–26, 2020, doi: 10.13140/RG.2.2.11094.57928.