

Identifying Condition Indicators for Artificially Intelligent Fault Classification in Rolling Element Bearings



Mohd Atif Jamil and Sidra Khanam

Abstract Bearing condition monitoring is significant in industries due to increased machine reliability and decreased production loss due to machinery breakdown. With the advancement of Artificial Intelligence (AI), Machine Learning (ML) techniques are reasonably useful to build condition monitoring systems for real-world applications. ML algorithms help distinguish faulty bearings from healthy ones and classify the related fault types using the extracted time-domain and frequency-domain features. This study recognizes distinctive features or condition indicators that effectively separate different fault groups and are worthy of training an ML model. Box plot and scatter plot of fault features are used to identify these condition indicators. Vibration datasets representing various faults are taken from the open-source Case Western Reserve University (CWRU) bearing database. A number of time-domain features are extracted from the ensemble data of bearing fault classes, consisting of healthy bearing, inner race fault, ball fault, and outer race fault. Our investigation indicates that more than one condition indicator is better for separating the fault categories. Six different ML models are trained using the condition indicators and the best-performing model is found through the classification accuracy, training time, and prediction speed of the classifier.

Keywords Condition monitoring · Bearing faults · Vibration features · Condition indicators · Machine learning · Fault classification

Nomenclature

CWRU Case Western Reserve University
CF Crest Factor

M. A. Jamil (✉) · S. Khanam
Department of Mechanical Engineering, Aligarh Muslim University, Aligarh, India
e-mail: majamil@myamu.ac.in

S. Khanam
e-mail: sidrakhanam@zhcet.ac.in

IF	Impulse Factor
K_u	Kurtosis
REB	Rolling Element Bearing
S_k	Skewness
\bar{X}	Mean
σ	Standard Deviation

1 Introduction

Modern machinery has a high starting cost, and its efficient operation depends on minimal operating and maintenance expenses. Rotating machinery must have rolling element bearings (REBs). As Rolling Element Bearings (REBs) are essential component of rotating machinery, their health monitoring is essential to reduce unintentional machinery shutdowns, minimize downtime for maintenance, and to enhance reliability and safety. Condition monitoring and fault diagnostics of REBs have emerged as key characteristics to meet these requirements. Condition monitoring is a procedure of knowing machinery health by capturing the operational information and examining it to put a figure/label on the state of equipment. It helps to detect and diagnose potential problems early in their development and fixing them by suitable recovery activities before they become hazardous enough to cause machine failure and other severe consequences. Subsequently, there is a requirement for a large amount of data for analysis. But the relationship between the bearing health and the condition monitoring data produced is not always well understood [1]. So, it is a dare to extract meaningful information from the data for condition monitoring.

For machine monitoring and diagnosis, a number of reliable techniques are well-established. These techniques include visual examination, stator current analysis, temperature monitoring, and vibration-based monitoring. The most often used parameter for identifying damage and monitoring machine condition is the vibration signature [2, 3]. Condition monitoring of bearings is also carried out using acoustic emission [4], non-contact infrared thermography [5, 6], and laser Doppler vibrometer [7]. Recent advances in sensing technology and the internet of things have introduced an intelligent framework to monitor bearing healthiness [8]. The bearing fault diagnosis is more problematic than detection because different faults can have analogous characteristics and different faults can happen at the same time. The two stages of the fault diagnosis procedure are vibration feature extraction and the classification of defects. Typically, features are retrieved using the Fast Fourier Transform (FFT), Hilbert Transform (HT), Short-Time Fourier Transform (STFT), Wavelet Transform (WT), both in continuous and discrete form, and the envelope analysis [9]. In order to identify the most important features, effective dimensionality reduction and feature selection methods have been used, namely, linear discriminant analysis, Principal Components Analysis (PCA), Sequential Floating Forward Selection (SFFS) [10], and Genetic Algorithm (GA), etc. Using ML techniques, fuzzy logic, and Deep

Learning (DL) approaches, fault classification deals with detecting bearing fault category. For condition monitoring of REBs, many AI techniques have been used, such as Artificial Neural Network (ANN) [11], Support Vector Machine [12, 13], K-Nearest Neighbors (KNN) [14, 15], and Hidden Markov Model (HMM) [16]. The Deep Neural Network (DNN) has also been documented in recent literature to determine the Remaining Useful Life (RUL) of bearings [17].

The identification of condition indicators, or features in the bearing data whose behavior change predictably as the bearing deteriorates due to presence of defects, is a crucial stage in the creation of condition monitoring algorithms. It may be recognized whether a bearing is healthy or defective using condition indicators. For fault classification and RUL estimation, they can be extracted from preprocessed data. The objective of feature extraction is to identify the most condensed and informative set of features (distinct patterns) to improve the effectiveness of the ML classifier, thereby achieving accurate classification.

To diagnose faulty and healthy rolling element bearing states, statistical characteristics of time signals can be used as condition indicators. The mean value of a particular signal or the standard deviation of the signal, for instance, may shift by a significant amount as the health of the bearing deteriorates. The worsening in the bearing healthiness may also be visible in some higher-order moments, such as kurtosis and skewness of the signal. The threshold values that differentiate healthy operation from defective one can be defined with such features, or changes in the bearing state can also be revealed by sudden or abrupt changes in the corresponding values. In the present analysis, the time-domain features namely, mean, median, kurtosis, skewness, crest factor, and impulse factor, are extracted to check their individual and combined effectiveness in classifying different fault categories using box plot and scatter plot approach, respectively.

2 Experimental Data and Machine Learning Implementation

Figure 1 illustrates the setup used to collect vibration data for healthy and defective ball bearings available at the bearing data centre of Case Western Reserve University (CWRU). The arrangement at its left has a 2 hp induction, a torque transducer in the center, and a dynamometer attached on the right. Single point faults of 0.007", 0.014", 0.021", and 0.028" in diameter were formed artificially on the inner race, rolling balls, and outer race of test bearings by electro-discharge machining (EDM).

At sampling frequencies of 12,000 and 48,000 Hz, vibration data was obtained for motor speeds of 1797 to 1720 rpm using two accelerometers that are positioned on the fan and drive ends of the housing of motor. The CWRU bearing data center documented the data and made it accessible to the public [18]. It can be used as a benchmark dataset to assess how well ML algorithms perform.

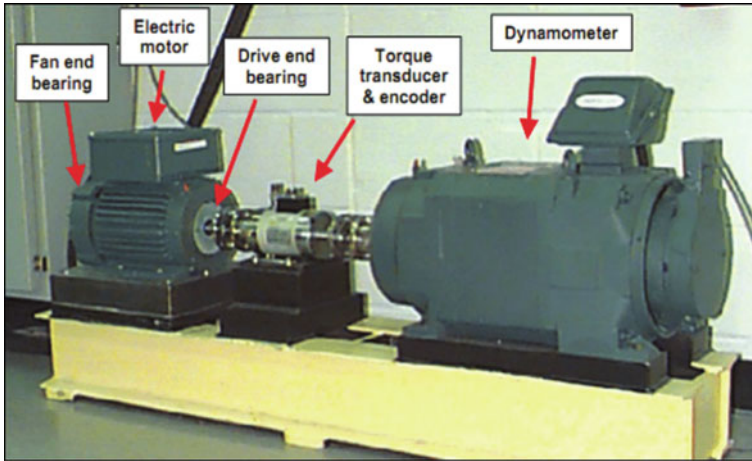


Fig. 1 Setup for CWRU bearing data collection [18]

2.1 Data Subset

The subset of CWRU data considered for the present analysis consists of 10 samples of 6000 data points/observations, each from four fault categories; healthy bearing, inner race fault, ball fault, and outer race fault located at 6 o'clock angular position. The fault codes 1, 2, 3, and 4 are assigned to these fault categories. Samples are collected at 12,000 samples per second at an average motor speed of 1772 rpm, and the fault size is 0.007".

2.2 Fault Features

The features used have a significant impact on the effectiveness of ML-based bearing defect detection techniques. The time-domain features described in Table 1 have been extracted from the raw vibration data and are assessed to be used to train the ML models for bearing fault classification.

x_i corresponds to some time-series data for $i = 1, 2, \dots, n$; n is the total no. of observations in a sample, $\bar{x} = \frac{1}{n} \sum_{i=1}^n |x_i|$ represents the absolute mean, σ denotes the standard deviation, and $x_{max} = \max|x_i|$.

Table 1 Description of time-domain features extracted from bearing vibration data

S. No	Feature	Feature description
1	Mean	$\bar{X} = \frac{1}{n} \sum_{i=1}^n x_i$
2	Median	$Med = \frac{(n+1)^{th} \text{ term}}{2}; n \text{ is even}$ $Med = \frac{n^{th} \text{ term} + (\frac{n}{2} + 1)^{th} \text{ term}}{2}; n \text{ is odd}$
3	Kurtosis	$K_u = \frac{\sum_{i=1}^n (x_i - \bar{x})^4}{(n-1)\sigma^4}$
4	Skewness	$S_k = \frac{\sum_{i=1}^n (x_i - \bar{x})^3}{(n-1)\sigma^3}$
5	Crest factor	$CF = \frac{x_{max}}{x_{rms}}$
6	Impulse factor	$IF = \frac{x_{max}}{\bar{x}}$

2.3 Cross-Validation Approach

The k-fold cross-validation procedure is implemented to assess machine learning models by splitting the original sample into k equal-sized subsamples. A single subsample is retained as the validation data, and the remaining k – 1 subsamples are used as training data. The cross-validation process is then repeated k times, with each of the k subsamples used exactly once as the validation data. The k outcomes from the folds are then combined to produce a single estimation. A 8-fold cross-validation is used on 40 samples, 10 each from the four fault categories considered in the present work.

3 Results and Discussion

Ensembled raw vibration data and power spectrum of different fault categories are presented. Boxplots are used to understand how each of the time-domain features; mean, median, kurtosis, skewness, crest factor, and impulse factor perform individually in differentiating between types of faults. Scatter plots of the combination of features are obtained to investigate how well a particular combination separates different kinds of faults and can be used as condition indicators to train machine learning models. Finally, the best-performing ML classifier that has been trained with the extracted features is shown by its confusion matrix.

3.1 Raw Data

Figure 2 shows the ensemble vibration signal of about 50 ms and the power spectrum containing measurements of healthy bearings corresponding to fault code 1 and faulty

bearings with three fault types: inner race fault, ball fault, and outer race fault located at the 6 o'clock position, each with a size of 0.007". These fault classes are assigned with fault codes 2, 3, and 4.

The time-domain ensemble vibration plot includes data of all the healthy and faulty conditions. The trend in the raw vibration signals of different fault categories gives an idea of their relative amplitude. As shown in the power spectrum plot, the resonance due to impact caused by interaction of the respective defect with the mating element excites the resonant frequencies of the bearing. The resonant frequency band

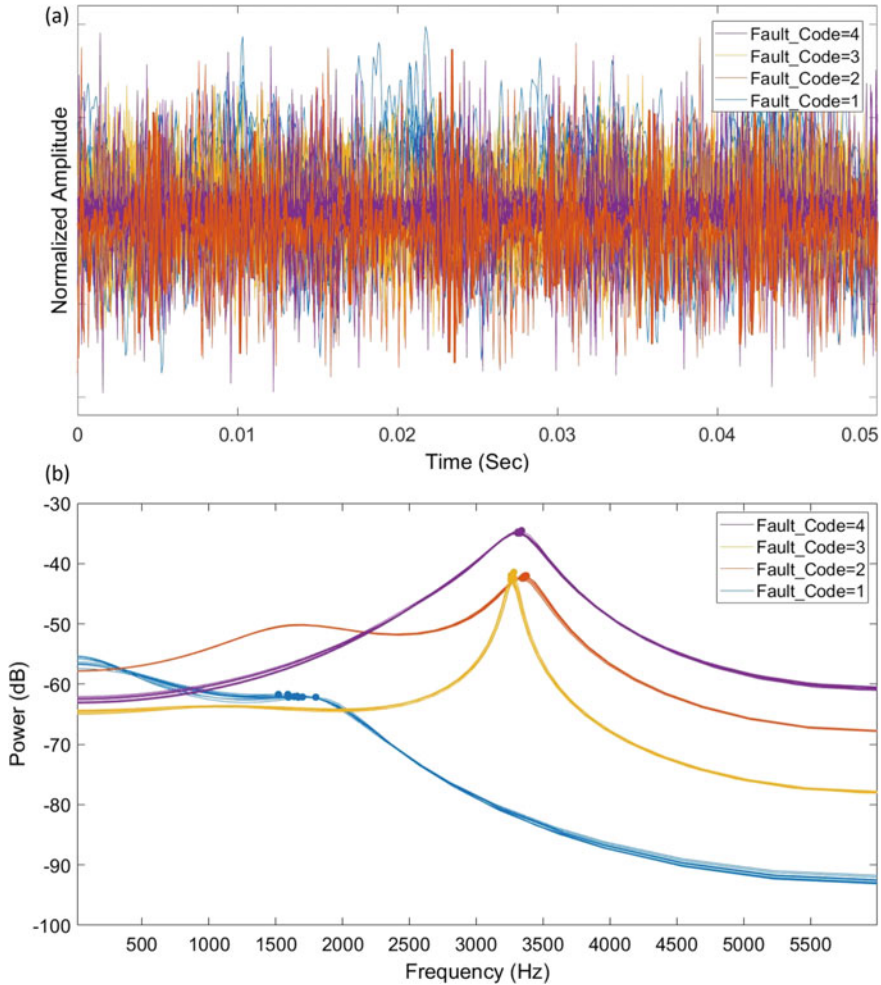


Fig. 2 **a** Ensemble vibration signal and **b** Power spectrum for different fault categories designated as 1, 2, 3, and 4, shown by blue, red, yellow, and purple colors, respectively

is almost the same for the inner race fault, ball fault, and outer race fault, in contrast to the healthy bearing.

3.2 *Box Plot*

A box plot, commonly referred to as a box and whisker plot, is a form of a chart that is used in explanatory data analysis in descriptive statistics. Box plots use the quartiles (or percentiles) of the data and averages to visually depict the distribution of numerical data. Box plots are used to display numeric data distributions, particularly when comparing values between various groups. In the present work, the box plots of individual features are compared with respect to fault classes of the bearing to check whether a feature can be considered as a condition indicator to train a ML classifier.

The box plots of time-domain features, mean, median, kurtosis, skewness, crest factor, and impulse factor, are shown in Fig. 3. In the case of (a) Mean and (d) Skewness, it can be observed that box plots for fault class 3 and 4 overlap, which implies that the distribution of mean and skewness values for these two fault classes are in the same respective ranges and therefore they may not distinguish between fault classes 3 and 4. Likewise, the (b) Median and (e) Crest Factor for the fault classes 2 and 4 are not very effective in classifying them because they lie in the same range owing to their corresponding overlaying box plots. For (c) Kurtosis, it can be seen that the box plots for fault classes 1 and 3 somewhat overlap, suggesting that the distribution of kurtosis values for these two fault classes is in the same range. As a result, they may not be able to differentiate between these classes.

As the boxes don't overlap for any of the four fault classes in the plot of (f) Impulse Factor only, there is a difference between the associated data groups, so Impulse Factor can be used to distinguish the fault categories under consideration. The other five features may not be considered promising for classifying the faults when used individually.

3.3 *Scatter Plot*

To indicate how much one variable affects another, scatter plots are used to exhibit data points on a horizontal and a vertical axis. A marker is used to represent each row in the data table, and the position of the marker depends on the values of the columns that are set up on the X and Y axes.

The scatter plots of the time-domain features in combination with each other are presented in Fig. 4. The scatter plot of (a) Skewness versus Mean shows a significant overlapping of data samples of fault classes 3 and 4, which indicates that these two classes are not clearly distinguished by Skewness and Mean. Similar is the case of (b) Kurtosis versus Impulse Factor where the fault classes 1 and 3 are not well

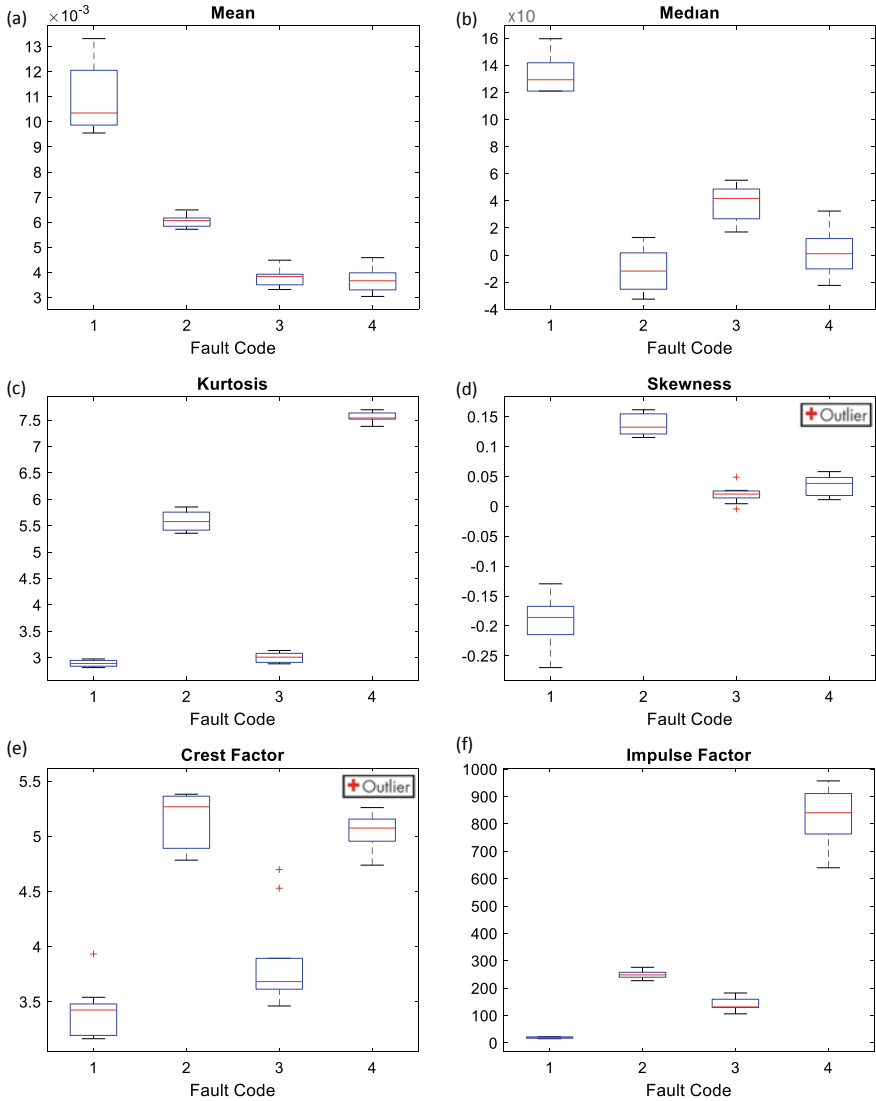


Fig. 3 Box plot of time-domain features for fault classification—**a** Mean **b** Median **c** Kurtosis **d** Skewness **e** Crest factor and **f** Impulse factor

differentiated. The plot of (c) Crest Factor versus Mean shows a somewhat better separation of the fault classes but with a minor closeness of the samples of fault classes 3 and 4. On the other hand, the scatter plots of (d) Impulse Factor versus Mean, (e) Kurtosis versus Skewness, and (f) Kurtosis versus Mean show good parting of all the fault classes with maximum separation in the case of (f) Kurtosis versus Mean.

Thus, these are good condition indicators and may be used to train the ML models for fault classification.

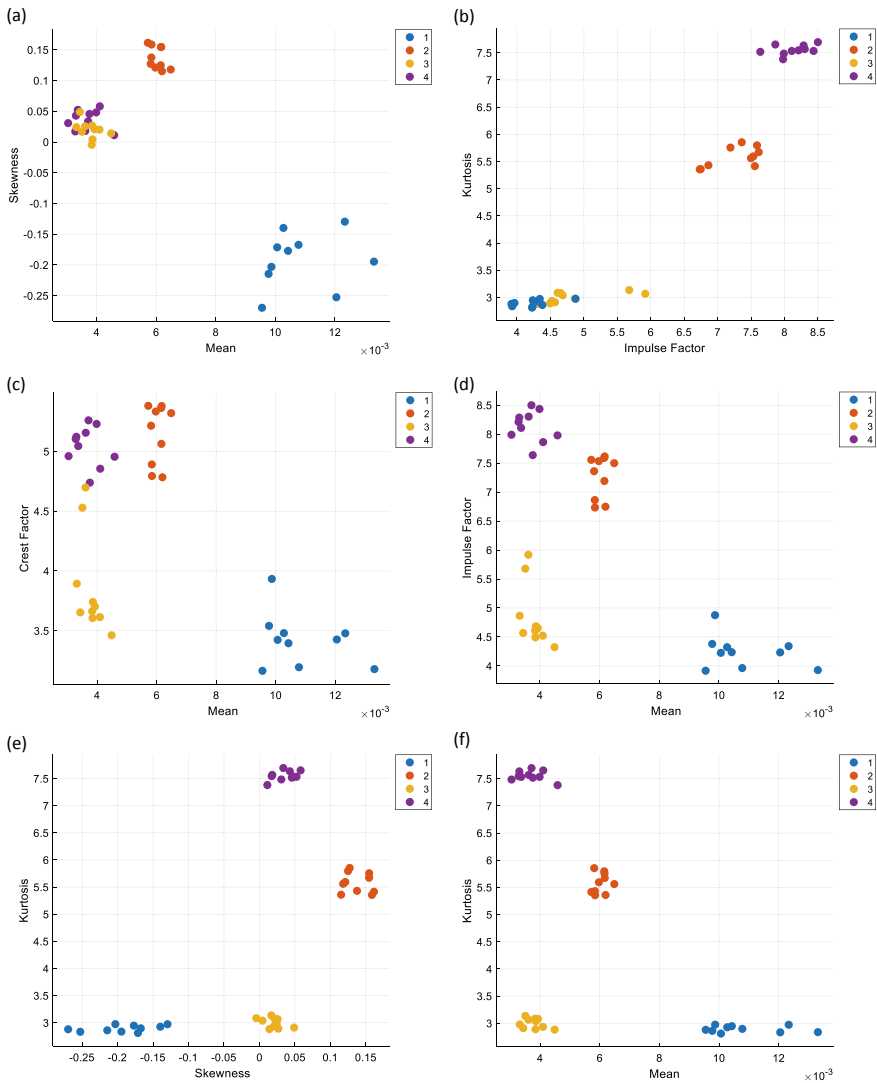


Fig. 4 Scatter plots for different pairs of time-domain features—**a** Skewness versus Mean **b** Kurtosis versus Impulse Factor **c** Crest Factor versus Mean **d** Impulse Factor versus Mean **e** Kurtosis versus Skewness **f** Kurtosis versus Mean

3.4 Confusion Matrix

The confusion matrix of a ML classifier is a summary of classification problem prediction outcomes. Count values are used to describe the number of accurate and inaccurate predictions for each class. It gives an insight into the errors being made by the classifier. The confusion matrices for (a) K-Nearest Neighbor, (b) Decision Tree, (c) Ensemble Classifier, (d) Discriminant Analysis, (e) Support Vector Machine, and (f) Gaussian Naïve Bayes models are shown in Fig. 5. The associated results of fault classification accuracy, time to train the classifier, and the corresponding prediction speed of these classification models are reported in Table 2.

It is worth noting that the classification accuracy obtained by SVM and Gaussian Naïve Bayes is the same, but the latter outperforms in terms of training time and prediction speed. The same is applicable when comparing the Ensemble Classifier and Discriminant Analysis. Overall, the Gaussian Naïve Bayes model performs best in the current analysis.

4 Conclusions

Investigating different fault types using a box plot shows that a single time-domain feature may not be sufficient to classify the faulty behavior, especially in multi-class fault classification problems. One cannot distinguish between all the fault types as the box plots of some of the time-domain features overlap, due to which these features are not enough to set fault types apart. The scatter plot of a combination of features reveals that two condition indicators are better than one for separating different faults. Different combinations of features may be tried to see which ones are better at classifying the defects.

In the present analysis, all the time-domain features except the impulse factor are insufficient to classify the fault categories when used independently. However some of their combination in different pairs may successfully distinguish between the fault classes, which can be easily understood from the scatter plots. Further, the combination of mean and kurtosis results in the best prediction of different fault categories. Therefore, these time-domain features and the other extracted features showing a bit lesser prediction are good candidates to train ML models. The condition indicators thus identified are used to train the ML models. Finally, the best-performing model may be selected by checking its accuracy using a confusion matrix. In case the accuracy of two or more classifiers match, the training time and prediction speed parameters may be used to recognize the best classifier. In this case, out of 6 ML classifiers, namely, k-Nearest Neighbors (k-NN), Support Vector Machines (SVM), Naïve Bayes classifier, Discriminant Analysis, Ensemble Classifier, and Decision Trees. The Gaussian Naïve Bayes is discovered to be the most effective.

There is no pre-determined number regarding how many features are enough to train a machine learning model. So, as a future direction, the feature extraction step

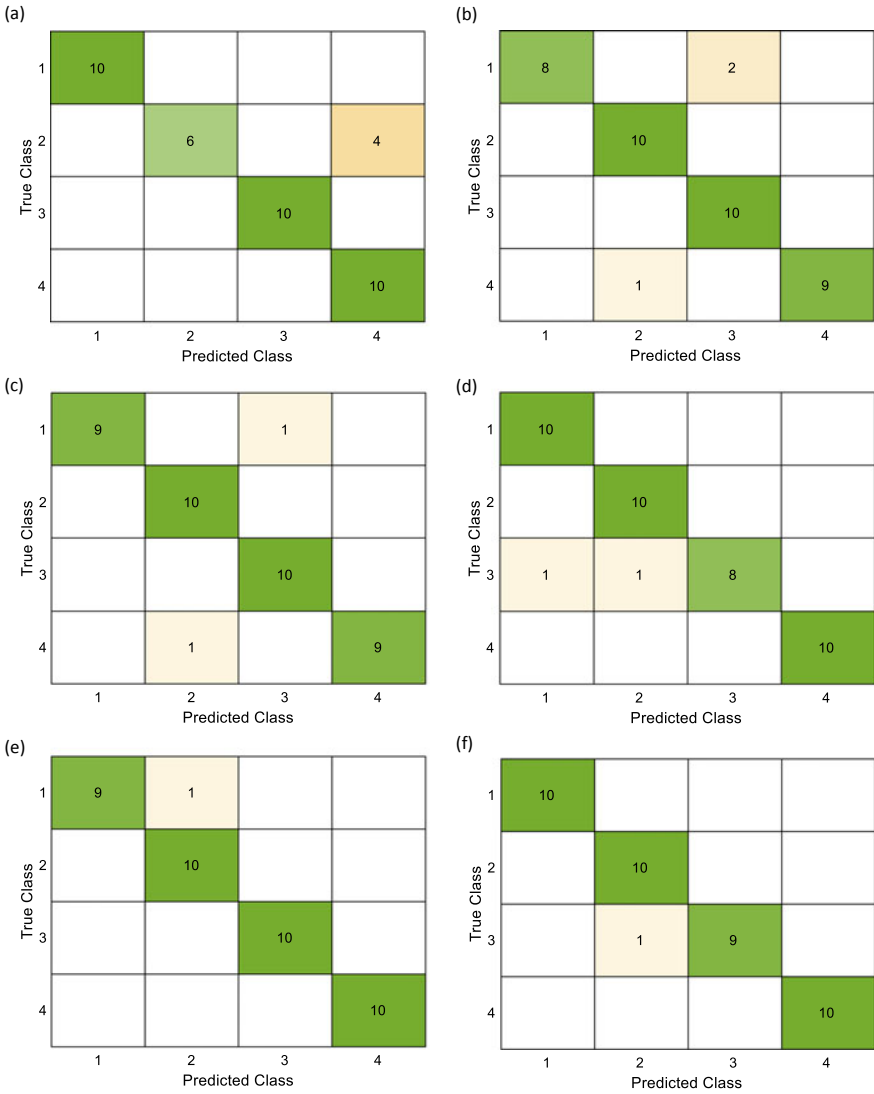


Fig. 5 Confusion matrices showing fault classification results—**a** K-nearest neighbor **b** Decision tree **c** Ensemble classifier **d** Discriminant analysis **e** Support vector machine **f** Gaussian Naïve Bayes

may be reconsidered, and machine learning models be trained with different sets of features to check for the possible improvement in the accuracy of ML models. It is also important to remember that ML models can benefit from a high-dimensional set of distinguishing features and can effectively differentiate the fault types.

Table 2 Performance results of different ML models

ML classifier	Accuracy (%)	Training time (S)	Prediction speed (Obs/S)
K-nearest neighbor (KNN)	90	0.69843	640
Decision tree	92.5	9.6881	1200
Ensemble classifier	95	1.8984	350
Discriminant analysis	95	0.72548	1000
Support vector machine (SVM)	97.5	0.79327	730
Gaussian Naïve Bayes	97.5	0.64558	980

References

- McArthur SDJ, Strachan SM, Jahn G. The design of a multi-agent transformer condition monitoring system. *IEEE Trans Power Syst.* 2004;19(4):1845–52. <https://doi.org/10.1109/TPWRS.2004.835667>.
- Goyal D, Pabla BS. The vibration monitoring methods and signal processing techniques for structural health monitoring: a review. *Arch Comput Methods Eng.* 2016;23(4):585–94. <https://doi.org/10.1007/s11831-015-9145-0>.
- Sehgal S, Kumar H. Damage detection using derringer's function based weighted model updating method. *Conf Proc Soc Exp Mech Ser.* 2014;5:241–53. https://doi.org/10.1007/978-3-319-04570-2_27.
- Glowacz A. Fault diagnosis of single-phase induction motor based on acoustic signals. *Mech Syst Signal Process.* 2019;117:65–80. <https://doi.org/10.1016/j.ymsp.2018.07.044>.
- Younus AMD, Yang BS. Intelligent fault diagnosis of rotating machinery using infrared thermal image. *Expert Syst Appl.* 2012;39(2):2082–91. <https://doi.org/10.1016/j.eswa.2011.08.004>.
- Lim GM, Bae DM, Kim JH. Fault diagnosis of rotating machine by thermography method on support vector machine. *J Mech Sci Technol.* 2014;28(8):2947–52. <https://doi.org/10.1007/s1206-014-0701-6>.
- Vass J, Šmíd R, Randall RB, Sovka P, Cristalli C, Torcianti B. Avoidance of Speckle noise in laser vibrometry by the use of kurtosis ratio: application to mechanical fault diagnostics. *Mech Syst Signal Process.* 2008;22(3):647–71. <https://doi.org/10.1016/j.ymsp.2007.08.008>.
- Tokognon Jr. CA, Gao B, Tian GY, Yan Y. Structural health monitoring framework based on internet of things: a survey. *IEEE Internet Things J.* 2017;4(3):619–635. <https://doi.org/10.1109/SYSCO.2016.7831328>.
- Shi DF, Wang WJ, Qu LS. Defect detection for bearings using envelope spectra of wavelet transform. *J Vib Acoust Trans ASME.* 2004;126(4):567–73. <https://doi.org/10.1115/1.1804995>.
- Pudil P, Novovičová J, Kittler J. Floating search methods in feature selection. *Pattern Recognit Lett.* 1994;15(11):1119–25. [https://doi.org/10.1016/0167-8655\(94\)90127-9](https://doi.org/10.1016/0167-8655(94)90127-9).
- Haj Mohamad T, Samadani M, Nataraj C. Rolling element bearing diagnostics using extended phase space topology. *J Vib Acoust Trans ASME.* 2018;140(6):1–9. <https://doi.org/10.1115/1.4040041>.
- Widodo A, Yang BS. Support vector machine in machine condition monitoring and fault diagnosis. *Mech Syst Signal Process.* 2007;21(6):2560–74. <https://doi.org/10.1016/j.ymsp.2006.12.007>.
- Jamil MA, Khan MAA, Khanam S. Feature-based performance of SVM and KNN classifiers for diagnosis of rolling element bearing faults. *Vibroeng Procedia.* 2021;36–42. <https://doi.org/10.21595/vp.2021.22307>.

14. Pandya DH, Upadhyay SH, Harsha SP. Fault diagnosis of rolling element bearing with intrinsic mode function of acoustic emission data using APF-KNN. *Expert Syst Appl.* 2013;40(10):4137–45. <https://doi.org/10.1016/j.eswa.2013.01.033>.
15. Jamil MA, Khanam S. Fault classification of rolling element bearing in machine learning domain. *Int J Acoust Vib.* 2022;27(2):77–90. <https://doi.org/10.20855/ijav.2022.27.21829>.
16. Zhou H, Chen J, Dong G, Wang H, Yuan H. Bearing fault recognition method based on neighbourhood component analysis and coupled hidden markov model. *Mech Syst Signal Process.* 2016;66–67:568–81. <https://doi.org/10.1016/j.ymsp.2015.04.037>.
17. Xia M, Li T, Shu T, Wan J, De Silva CW, Wang Z. A two-stage approach for the remaining useful life prediction of bearings using deep neural networks. *IEEE Trans Ind Inform.* 2019;15(6):3703–11. <https://doi.org/10.1109/TII.2018.2868687>.
18. Case Western Reserve University (CWRU) Bearing Data Center [Online]. <https://csegroups.case.edu/bearingdatacenter/pages/welcome-case-western-reserve-university-bearing-data-center-website>.