



Bearing Fault Diagnosis Based on Artificial Intelligence Methods: Machine Learning and Deep Learning

Ahmed Ghorbel^{1,2} · Sarra Eddai¹ · Bouthayna Limam¹ · Nabih Feki¹ · Mohamed Haddar¹

Received: 21 April 2024 / Accepted: 6 August 2024
© King Fahd University of Petroleum & Minerals 2024

Abstract

This paper presents a comprehensive study on the application of Artificial Intelligence (AI) methods, specifically machine learning and deep learning, for the diagnosis of bearing faults. The study explores both data preprocessing-dependent methods (Support Vector Machine, Nearest Neighbor, and Decision Tree) and a preprocessing-independent method (1D Convolutional Neural Network). The experiment setup utilizes the Case Western Reserve University dataset for signal acquisition. A detailed strategy for data processing is developed, encompassing initialization, data loading, signal filtration, decomposition, feature extraction in both time- and frequency-domains, and feature selection. Indeed, the study involves working with four datasets, selected based on the distribution curves of the indicators as a function of the number of observations. The results demonstrate remarkable performance of the AI methods in bearing fault diagnosis. The 1D-CNN model, in particular, shows high robustness and accuracy, even in the presence of load variations. The findings of this study shed light on the significant potential of AI methods in improving the accuracy and efficiency of bearing fault diagnosis.

Keywords Intelligent fault diagnosis · Machine learning · Deep learning · Indicators · CWRU dataset

1 Introduction

Rotating machinery, including bearings, forms the backbone of many industrial processes [1]. Their efficient operation is crucial for productivity and profitability. However, these machines are prone to wear and tear, and their failure can lead to significant downtime and financial loss [1, 2]. Traditionally, the diagnosis of faults in rotating machines has relied on manual inspection and conventional vibration analysis techniques. Though effective, these methods can be time-consuming, require expert knowledge, and may not detect faults at an early stage [2].

In recent years, Artificial Intelligence (AI) has emerged as a powerful tool for detecting faults in rotating machinery [3]. AI-based methods, particularly Machine Learning (ML) and Deep Learning (DL), are capable of analyzing large

volumes of data, learning complex patterns, and accurately predicting errors. These technologies have the potential to automate the diagnostic process, improve early fault detection, and reduce the need for manual inspection [1, 3, 4]. AI has been widely applied to the identification of bearing faults in rotating machinery. To this end, a variety of AI methods have been employed, each with its unique advantages. Artificial Neural Networks (ANNs) are capable of learning and recognizing patterns from data, making them suitable for diagnosing different types of bearing faults [5, 6]. Support Vector Machines (SVMs) are particularly effective in handling high-dimensional data and have been used to achieve high classification accuracy [7–9]. Genetic Algorithms (GAs) have been used for feature selection in bearing fault diagnosis, optimizing the selection of features that are fed into the diagnostic model, thereby improving the model's performance [10–12]. The research work in Ref. [13] combines the genetic algorithm with SVM for the diagnosis of bearing defects. Fuzzy Logic systems handle uncertainties in the data and have been used to enhance the bearings' condition monitoring [14, 15]. Random Forests have been used for bearing fault diagnosis due to their ability to handle large datasets and provide importance scores for features [16–18].

✉ Ahmed Ghorbel
gh.ahmed.enis@gmail.com

¹ Laboratory of Mechanics, Modelling and Production, National School of Engineering of Sfax, University of Sfax, Sfax, Tunisia

² Higher Institute of Applied Sciences and Technology of Kairouan, University of Kairouan, Kairouan, Tunisia



The K-Nearest Neighbors (K-NN) algorithm has been effectively applied in the diagnosis of bearing faults [19, 20]. This method classifies a new test sample based on the majority of its K-Nearest training samples. In the context of bearing fault diagnosis, K-NN can be particularly useful due to its simplicity and effectiveness [21, 22]. Amaury's study [23] integrates Support Vector Machine (SVM) and K-Nearest Neighbor (K-NN) techniques to monitor rotating machinery using vibration data. Another technique that can be found in the literature is Decision Trees (DT). It has been effectively utilized in the diagnosis of bearing faults in rotating machinery [24, 25]. This method, which offers a clear visualization of the decision-making process, has been used to enhance the interpretability of diagnostic models. The Naive Bayes classifier, which relies on the assumption of independence, has proven to be effective and common in diagnosing bearing faults [26]. Some drawbacks limiting the use of this method are the conditional independence hypothesis and the accuracy of the estimate [27, 28].

Deep learning models, such as Convolutional Neural Networks (CNNs), have shown promising results in bearing fault diagnosis by automatically extracting features from raw data, leading to improved diagnostic accuracy [29, 30]. The effectiveness of these methods is likely to depend on the specific characteristics of the bearing faults and the quality of the data available. It is also common to use a combination of these methods to improve the accuracy of the diagnosis [31–34].

This paper offers an exhaustive review of the most recent developments in the application of Artificial Intelligence (AI) techniques for the diagnosis of faults in rotating machinery, with a particular emphasis on bearing defects. The paper begins by exploring the AI techniques that have been employed, providing a clear explanation of various methods, and discussing their primary advantages and disadvantages. Subsequently, it provides a detailed description of the test bench and explains the strategy developed for preprocessing the database. The paper concludes by presenting the results and engaging in a comparative discussion of the methods used.

2 Bearing Faults Diagnosis

2.1 Traditional Methods for Bearing Diagnosis

Bearing diagnosis is an essential component of predictive maintenance in industrial systems. Bearings play a crucial role in rotating machinery, and their failure can lead to costly downtime and significant equipment damage. Historically, the diagnosis of bearing faults has relied heavily on classical methods such as vibration analysis and acoustic emission techniques.

2.1.1 Vibration Analysis

Vibration analysis has been one of the most widely used techniques for bearing fault detection. It involves measuring the vibrations emitted by machinery to identify patterns indicative of faults. Traditional vibration analysis methods rely on time-domain and frequency-domain features to detect anomalies in the vibration signals. Techniques such as Fast Fourier Transform (FFT) are commonly used, where characteristic fault frequencies can be identified [35]. However, while effective, these methods require expert knowledge for accurate interpretation and can be time-consuming, as they often involve manual inspection of vibration signatures [36].

2.1.2 Acoustic Emission Techniques

Acoustic emission (AE) techniques have also been employed for bearing diagnosis, capturing high-frequency stress waves generated by defects in bearings. AE sensors are capable of detecting crack formation and surface defects at an early stage. Nonetheless, acoustic emission methods can be susceptible to environmental noise and require sophisticated signal processing techniques to extract meaningful information [37].

2.1.3 Oil Analysis

Oil analysis is another traditional approach that involves examining lubricant samples for particles or contaminants indicative of wear and tear in the bearing components. This method provides insights into the internal condition of bearings and can help predict failures before they occur. However, it typically requires laboratory testing and may not be practical for real-time monitoring [38].

2.2 The Shift Toward Intelligent Systems

As industries moved toward more complex and automated systems, the limitations of classical diagnostic methods became apparent. These traditional techniques often require significant manual intervention and are limited in their ability to handle large volumes of data. Consequently, the focus shifted toward more intelligent and automated approaches, leveraging advancements in signal processing and data analysis.

Signal processing techniques such as wavelet transform and Hilbert–Huang Transform (HHT) have been developed to address the limitations of traditional methods. Wavelet transform, in particular, offers a multi-resolution analysis capability, allowing for the detection of transient features in non-stationary signals commonly associated with bearing defects [39]. These advanced techniques improve fault

detection accuracy but still require expert interpretation and manual setup.

2.3 The Rise of Artificial Intelligence in Bearing Diagnosis

The advent of Artificial Intelligence (AI) has revolutionized the field of fault diagnosis by introducing automated and more accurate methods for bearing monitoring [1, 3, 4]. AI techniques such as machine learning and deep learning have shown great promise in enhancing the diagnosis process by automating feature extraction and classification tasks. The integration of AI in fault diagnosis not only enhances accuracy and reliability but also reduces the dependence on human expertise. As AI technology continues to advance, further improvements in the efficiency and effectiveness of bearing fault diagnosis are anticipated, paving the way for smarter and more resilient industrial systems.

Combining AI with traditional techniques has led to hybrid approaches that capitalize on the strengths of both worlds. These hybrid models offer a balanced approach, leveraging the precision of signal processing with the automation and scalability of AI [13].

3 Artificial Intelligence Methods Applied in Bearing Faults Diagnosis

In the domain of machine learning, various techniques necessitate distinct data preprocessing requirements. Support Vector Machines (SVM), K-Nearest Neighbors (K-NN), and Decision Trees (DT) are methods that typically require various forms of data preprocessing. The latter may involve tasks such as data normalization or standardization, handling missing values, or feature extraction and selection. These steps are crucial to ensure that the data fed into these models are in a suitable format and of a quality, enabling the models to learn effectively [40].

On the other hand, 1D Convolutional Neural Networks (1D-CNNs), a type of deep learning model, can often work with raw data without the need for extensive preprocessing. They are capable of automatically learning and extracting useful features from raw data during the training process. This makes them particularly useful for tasks such as time-series analysis or natural language processing, where the raw data can be fed directly into the model. However, it is important to note that while 1D-CNNs can work with raw data, some levels of preprocessing such as data cleaning or reshaping might still be required depending on the specific task or dataset [41].

3.1 Methods with Data Preprocessing: SVM, K-NN, and DT

In the field of bearing fault detection, machine learning algorithms such as Support Vector Machines (SVM), K-Nearest Neighbors (K-NN), and Decision Trees are widely utilized. These algorithms excel at analyzing data, learning from it, and subsequently applying the acquired knowledge to make informed decisions about the presence of bearing faults. The application of these machine learning techniques has consistently produced satisfactory results in this domain. They have significantly contributed to the advancement of predictive maintenance strategies by enhancing the accuracy and efficiency of fault detection processes. This, in turn, helps in reducing downtime and maintenance costs, thereby improving the overall operational efficiency of the machinery.

3.2 Method Without Data Preprocessing: 1D-CNN

The application of deep learning methods, such as 1D-CNNs, has transformed the field of bearing diagnosis by allowing machines to autonomously learn complex patterns from vibration data. This advancement has significantly enhanced the accuracy and efficiency of fault detection in rotating machinery. Recently, there has been a substantial increase in the use of deep learning, which is a type of machine learning that is competent not only at sorting things into categories with more accuracy, both in broad and in specific terms, but also at handling lots of data [42].

Deep learning methods are more efficient and accurate compared to traditional machine learning techniques because the former often solve the whole problem at once, while the latter break the problem down into smaller parts before putting the pieces together at the end [43].

Figure 1 illustrates the process of fault diagnosis using Machine Learning (ML) and Deep Learning (DL). It reveals the steps from data collection to fault identification and classification and clearly compares the two processes.

Table 1 displays the advantages and limitations of the different machine learning algorithms used in this paper, namely SVM, K-NN, DT, and 1D-CNN. It offers insights into their strengths and weaknesses in addressing various data and usage challenges.

4 Experiment Setup and Signal Acquisition

4.1 Data Description

The dataset utilized in this section was obtained from the Bearing Data Center at Case Western Reserve University (CWRU) [44]. Many researchers have employed this



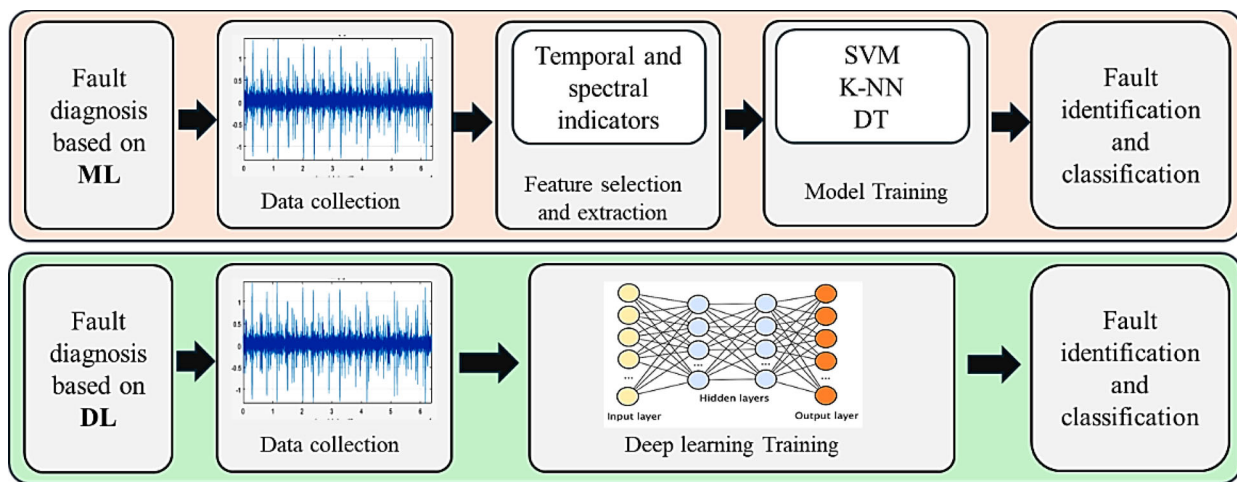


Fig. 1 Process of ML and DL for fault diagnosis

Table 1 List of AI algorithms and their advantages and limitations

Algorithm	Advantages	Limitations
SVM	Ability to handle multi-dimensional spaces, adaptation to nonlinear boundaries, minimizing generalization errors	Need for data preprocessing, less efficiency in managing missing data
K-NN	Ease of implementation, ability to handle complex data, fast learning	Influence of Euclidean distance, sensitivity to k-parameter selection
DT	Relative ease to interpret, robustness to outliers, handling missing values	Not good for regression, overfitting problem, instability, predictions are not smooth or continuous
1D-CNN	Robustness to noise and variability, faster training speed, efficient feature extraction, effective for 1D signals	Limited data augmentation, feature extraction challenges: sensitivity to hyperparameters

database in their investigations [1, 2, 40, 45–48]. The website provides access to test data related to both functional and defective ball bearings. These tests were conducted using a Reliance Electric 2 HP electric motor, with acceleration measurements taken at locations near and far from the motor bearings.

Figure 2 shows a 2 HP motor (positioned on the left), a torque transducer/encoder (center), a dynamometer (right), and control electronics (not shown). The motor shaft is supported by the test bearings. Point defects were intentionally

introduced into these bearings using EDM, resulting in diameters ranging from 0.178 to 0.533 mm at the inner raceway, ball, and outer raceway. Bearings with defects were then reinstalled in the test motor, and vibration data were recorded for motor loads varying from 0 to 3 HP. Different sensor positions were used for the outer race, with motor speeds oscillating between 1797 and 1730 rpm. The data collection for drive-end bearing faults occurred at a rapid rate of 48,000 samples per second. This high sampling frequency allowed precise monitoring and analysis of the bearing behavior, helping in the fault detection and diagnosis.

4.2 The Developed Strategy for Data Processing

Figure 3 shows the working steps and comprehensive preprocessing of vibration signal data from different operating conditions.

4.2.1 Initialization and Data Loading

The initial phase focuses on initializing the working environment and loading the data. Besides, there is an extraction of signals from «.mat» format files (MATLAB), which contain crucial information from sensors. These signals, representing measurements in the time-domain, are organized into a matrix named *M*. This *M* matrix constitutes an essential data structure, serving as a basis for subsequent analyses.

4.2.2 Signals Filtrating

Filtering is a critical process that helps in discarding unnecessary elements from vibration signals. By employing a finite impulse response (FIR) filter along with a Hamming window, the masking effect can significantly be diminished.

Fig. 2 Bearing components and the experimental configuration of the ball bearing system on the CWRU bearing test rig

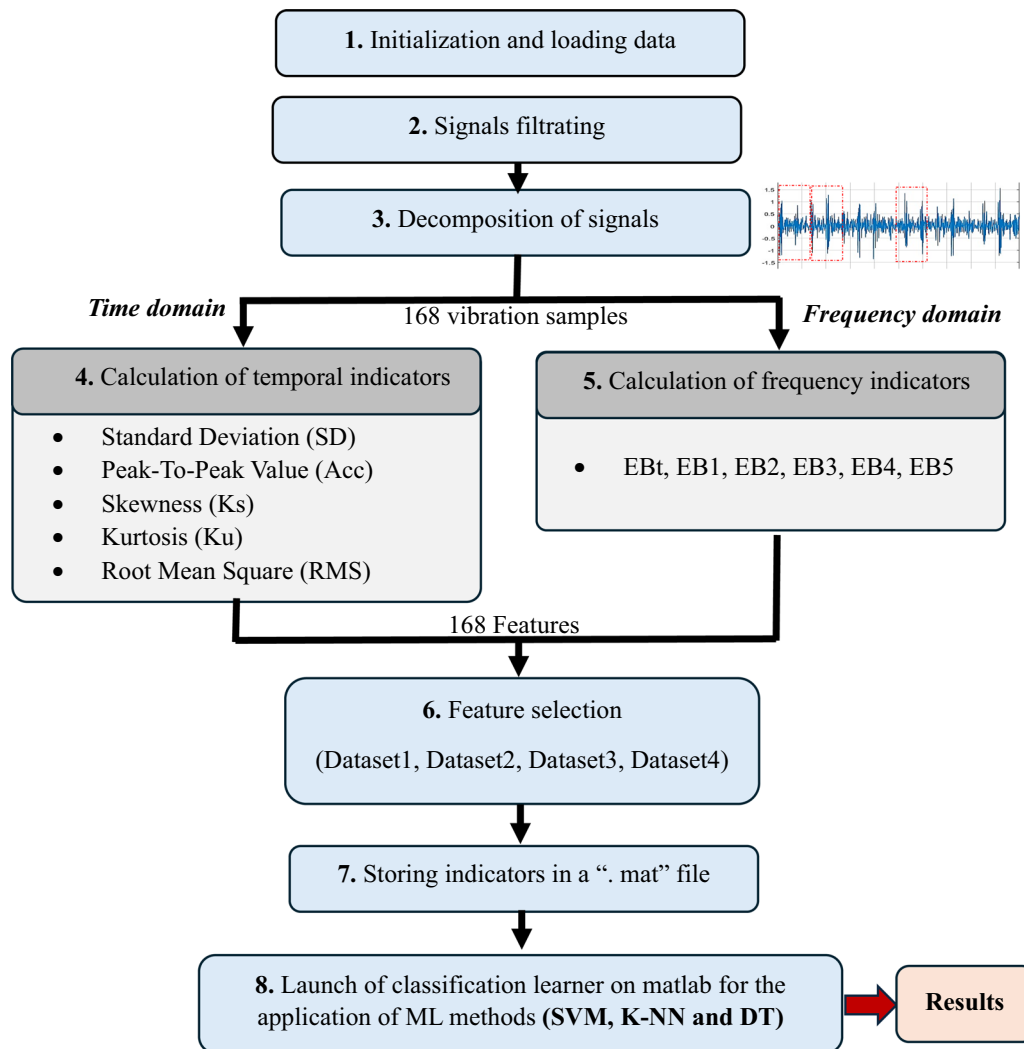
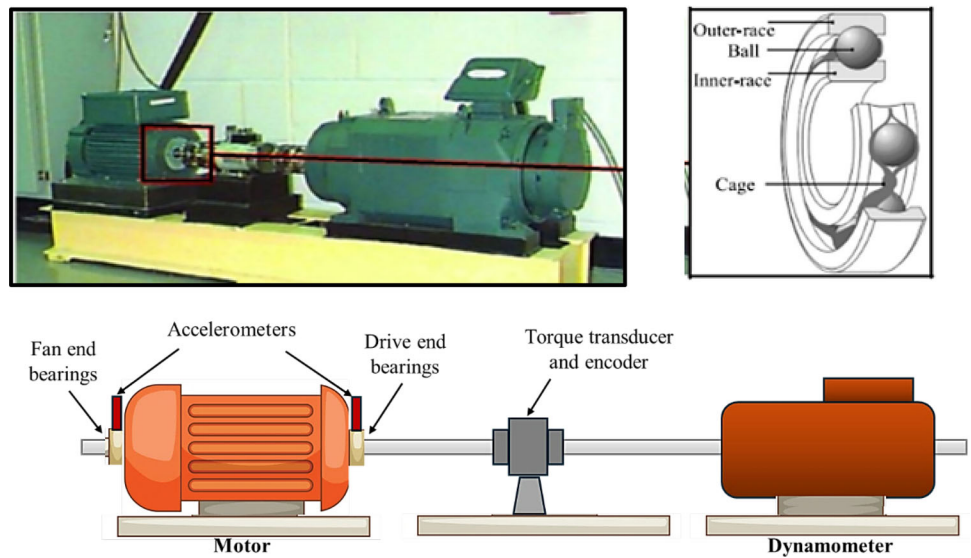


Fig. 3 The framework of the proposed strategy

4.2.3 Signals Decomposition

Signal decomposition is an effective method that entails the division of the signal into multiple components or frequency bands to collect crucial information for fault detection. By independently examining each component, this method facilitates the computation of time and frequency indicators for each component, assisting in the determination of signal attributes across diverse frequency ranges.

In this step, the time signals are segmented into 28 samples of fixed length, thus allowing a more in-depth analysis of variations and local behaviors. Each signal is divided into 28 samples, which in turn are then organized and stored.

4.2.4 Feature in Time-Domain

Statistical indicators are frequently employed among the time-domain features because of their robust association with early bearing damage [49]. Root Mean Square (RMS) quantifies the overall vibration energy in a signal. It offers a measurement tool for the signal's amplitude, which is crucial for detection of bearing defects. While high RMS values indicate increased vibration levels, the kurtosis indicator assesses the distribution of vibration peaks. Furthermore, elevated kurtosis values suggest localized faults, such as pitting or spalling, which create sudden impact forces during rotation. Peak-to-peak (A_{cc}) is another significant feature that measures the difference between the highest and lowest points in a vibration waveform. The large values of this feature indicate significant variations in vibration amplitudes, often linked to bearing defects like looseness or cracked races. Besides, Standard Deviation (SD) measures the dispersion or spread of vibration data points around the mean and reflects the variability of the signal. High SD values indicate inconsistent vibration patterns, which can be indicative of bearing defects such as looseness, misalignment, or early-stage wear. The last temporal indicator used in this paper is the skewness which evaluates the symmetry of the vibration distribution. Positive skewness indicates a longer tail on the right side of the distribution, while negative skewness reveals a longer left tail. Abnormal skewness values may suggest specific fault conditions. For instance, positive skewness could be associated with localized defects like spalling, while negative skewness might be accredited to outer raceway wear.

Hence, the collection of five time-domain features is generated utilizing the equations provided in Table 2. The parameter x_i represents a sample in the acquired signal, and N defines the total number of samples.

4.2.5 Feature in Frequency-Domain

Frequency-domain analysis, commonly used for monitoring bearing conditions, empowers the distinct frequencies linked

Table 2 Equations of the extracted time-domain features [50]

Feature	Equation
Root mean square	$RMS = \sqrt{\frac{1}{N} \sum_{i=1}^N x_i^2}$
Standard deviation	$SD = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - \text{mean})^2}$
Kurtosis	$Kurtosis = \frac{\sum_{i=1}^N (x_i - \text{mean})^4}{N \cdot SD^4}$
Peak-to-peak	$Acc = \max(x) - \min(x)$
Skewness	$Skewness = \frac{1}{N} \sum_{i=1}^N \left(\frac{x_i - \text{mean}}{\text{std}} \right)^3$

to bearing defects in vibration and current signals. Energies from the envelope spectrum are utilized for this purpose.

This section implements the Hilbert transform to obtain the envelope spectrum of each signal slice. Then, the energy $E^{frequency}$ is calculated in specific frequency ranges, namely [0–1500 Hz], [1500–3000 Hz], [3000–4500 Hz], [4500–6000 Hz], and [0–6000 Hz]. These energy measurements are used to evaluate five frequency indicators, including the total energy “EBt” in the band [0–6000 Hz] and the energy in specific sub-bands “EB1, EB2, EB3, and EB4,” thus providing a detailed analysis of the distribution of frequency energy in each signal sample. The energy equation is defined by:

$$E^{frequency} = \sum_{-\infty}^{+\infty} |X(f)|^2$$

where $X(f)$ is the signal equation in the frequency-domain, and s represents the sample size.

4.2.6 Feature Selection

To evaluate the effectiveness and preference of different classification methods, it is necessary to select a database. Additionally, in order to select sets of indicators that will be used as inputs for machine learning methods, it is essential to study the impact of the defect's severity on the indicators' reactivity.

Figure 4 presents the evolution of the five temporal indicators as a function of the number of observations, totaling 168 observations corresponding to the total number of slices between the state without load and that with 3 HP. The first 42 observations correspond to the defect with a diameter of 0.533 mm (14 without load and 28 with 3 HP). Starting from the 43rd observation, the signal has a defect of 0.355 mm, then from the 85th to 126 observations, the defect decreases to 0.178 mm, finally reaching the normal state (the last 42 observations). From the evolution curve of the RMS values, a significant increase in amplitude, surpassing 2.5, is observed

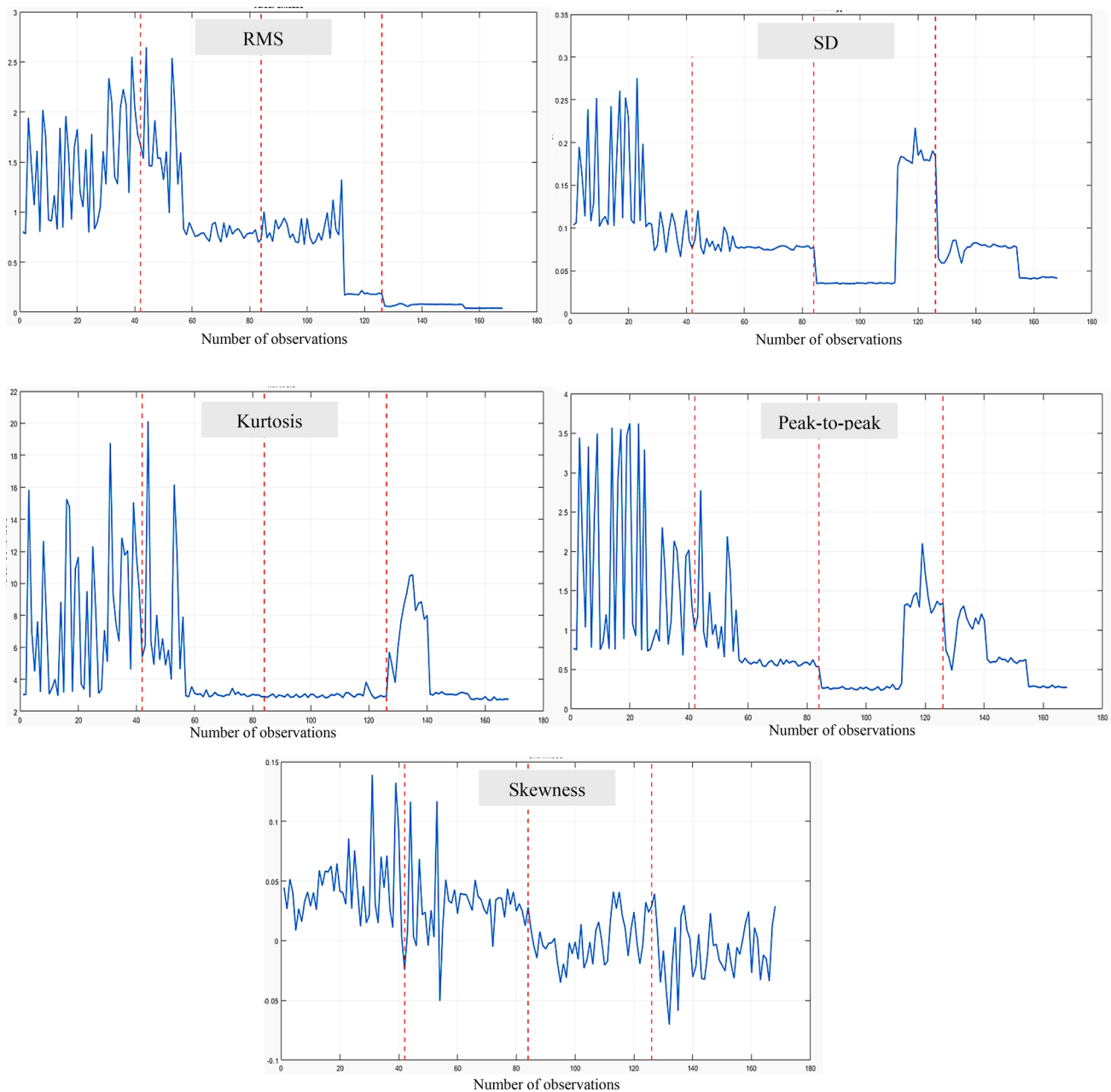


Fig. 4 Time-domain features distribution of the data

as the diameter of the defect on the highest rolling element grows.

Notably, this indicator exhibits high sensitivity to variations in the fault. The Standard Deviation (SD) value for a bearing in the case without defect, under normal conditions, generally does not exceed 0.1. However, an increase exceeding 0.25 is observed as the defect diameter on the rolling element increases. This indicator turns out to be sensitive to variations in the default. The variation of the kurtosis values and the peak-to-peak value also shows that these two indicators are sensitive to variations in the fault. The variation of

the skewness indicator can be explained by the fact that the distinction between the fault-free state of the bearing and the state with fault can be difficult because the peak values are close.

Frequency-domain indicators, extracted from vibration signal envelope spectra, play a crucial role in bearing fault diagnosis. Figure 5 illustrates the evolution of these indicators of signal energy in various fault scenarios. The indicators EB1, EB2, EB3, and EBt exhibit remarkable discriminatory power between healthy bearings and those with faults. As defects develop, the amplitude of these indicators increases

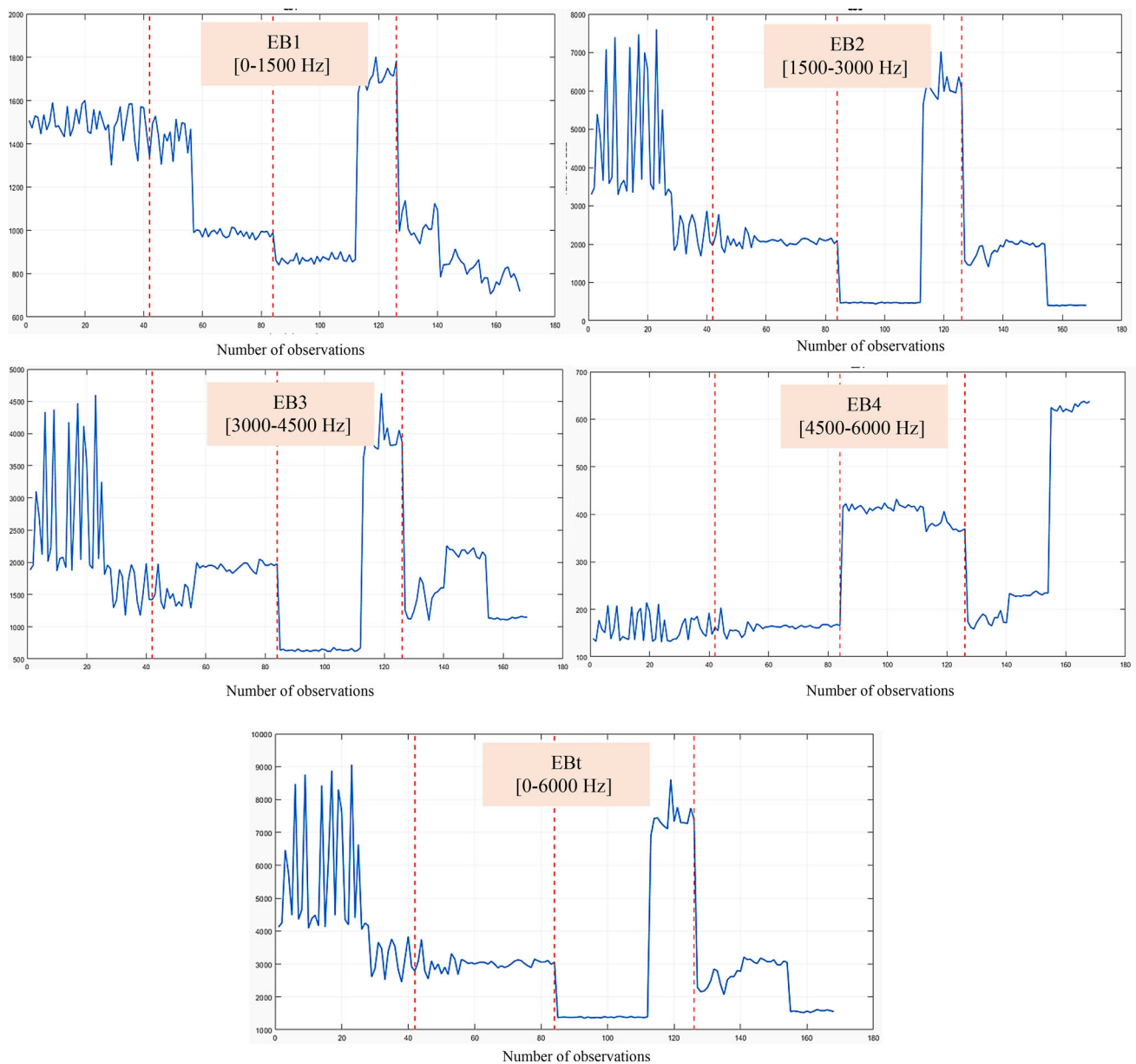


Fig. 5 Energy features distribution of the data

significantly. Notably, even minor faults cause noticeable deviations from the healthy baseline. Their sensitivity to fault variations makes them valuable tools for early detection. In contrast, the behavior of EB4 is less straightforward. Its variation does not yield a clear and interpretable curve shape.

To sum up, the relevant indicators to be used for defect classification include the standard deviation, root mean square, peak-to-peak, kurtosis, and total energy (EBt), as well as the energy in bands EB1, EB2, and EB3.

4.3 Storage of Indicators in a .mat File

This step ensures the preservation of both temporal and frequency-domain results (10 indicators) in a matrix. This matrix serves to systematically group the indicators extracted from each signal slice. Subsequently, these results are saved in a .mat file, which contains a main matrix that consolidates all the results. This recording process enables organized storage of the resultant data for future reference or in-depth analysis.

Based on the shape of the indicators depicted in Figs. 4 and 5, four sets of indicators have been selected as follows:

- Set 1: All indicators,
- Set 2: Standard deviation, peak-to-peak, EB1, EB2,
- Set 3: Kurtosis, root mean square, EBt, EB3,
- Set 4: Standard deviation, root mean square, EBt, EB1.

5 Results and Discussion

After introducing the test bench, this section explores the application of artificial intelligence methods on the vibration data from the drive-end bearing, considering loads of 0 and 3 HP. A method that does not require data preprocessing will be implemented alongside other methods that do require preprocessing. Finally, a comprehensive comparative study will be conducted.

5.1 Methods with Data Preprocessing

5.1.1 Support Vector Machine (Medium Gaussian SVM)

The SVM version with a medium Gaussian kernel refers to the application of a Gaussian kernel function (also known as RBF—Radial Basis Function) with a bandwidth parameter set at a medium level.

The Gaussian kernel function is among the most commonly used techniques that perform transformation from the original feature space to a higher-dimensional space, thereby facilitating the separation of nonlinear classes. This kernel is influenced by a parameter called "gamma," which governs the configuration of the Gaussian function. While a high gamma tightens the Gaussian function, a low gamma widens it. By selecting an intermediate gamma, a balanced decision boundary, neither overly flexible nor overly rigid, can be established, thus helping to prevent overfitting to the training data.

- Set 1

The implementation of this method on the entire dataset resulted in a remarkable accuracy of 98.2% for the unloaded set with a 3 HP load. Figures 6 and 7 display the corresponding scatter plots for both sets, along with the prediction outcomes.

Figures 8 and 9 represent the confusion matrices for both cases, together with the associated prediction results.

- Set 2

The implementation of this method on the entire dataset 2 resulted in a remarkable accuracy of 98.2% for the unloaded set and 99.1% for the set with a 3 HP load. Figures 10 and

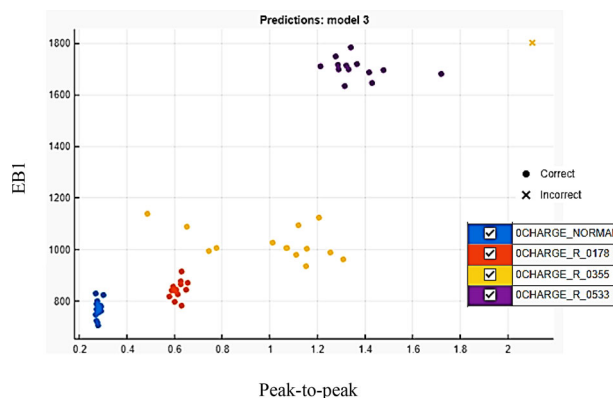


Fig. 6 Scatter plot for 0 charge (Set 1)

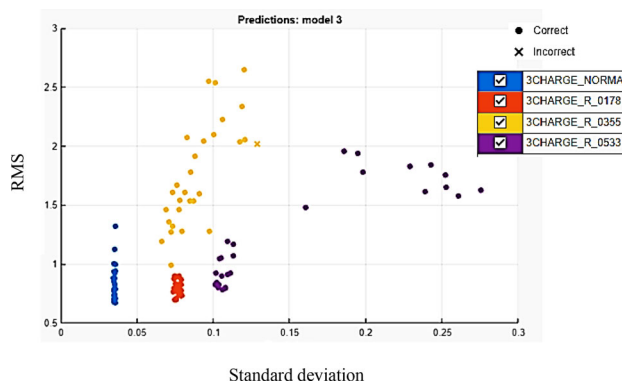


Fig. 7 Scatter plot for 3 HP charge (Set 1)

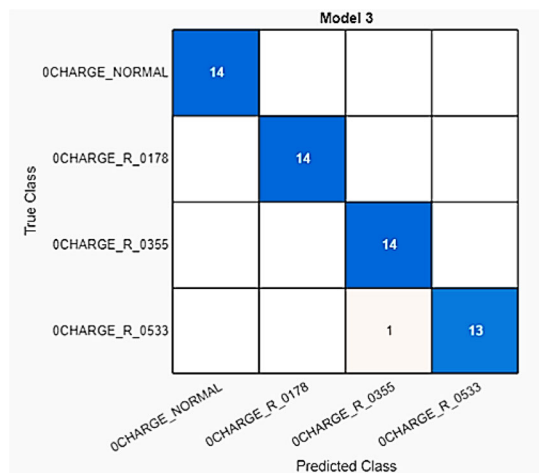


Fig. 8 Confusion matrix without load (Set 1)

11 display the corresponding scatter plots for the second set, alongside the prediction results.

The confusion matrices for both loads are presented in Figs. 12 and 13, accompanied by the associated prediction results.

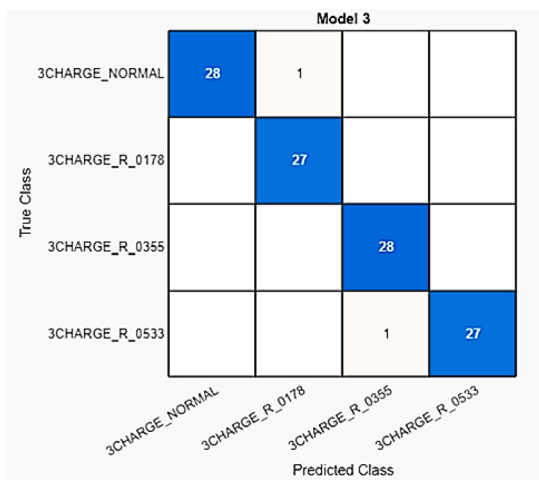


Fig. 9 Confusion matrix with 3 HP charge (Set 1)

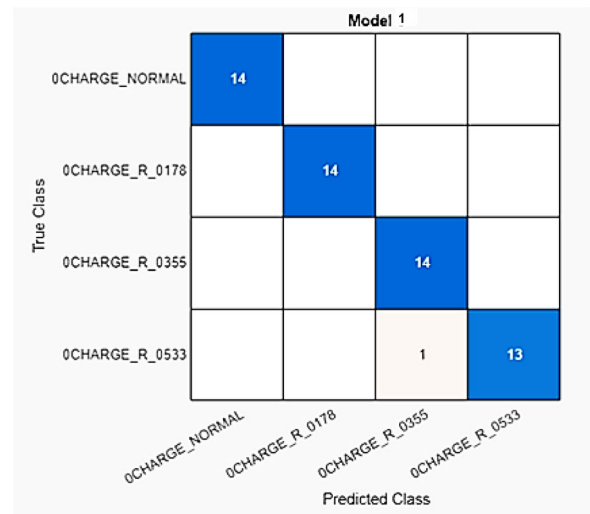


Fig. 12 Confusion matrix with 0 charge (Set 2)

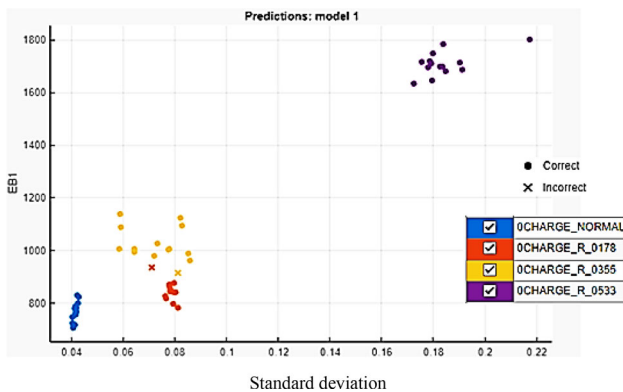


Fig. 10 Scatter plot for 0 charge (Set 2)

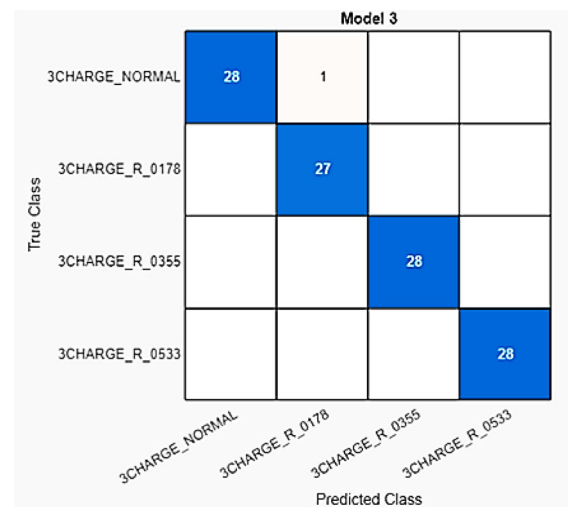


Fig. 13 Confusion matrix for 3 HP charge (Set 2)

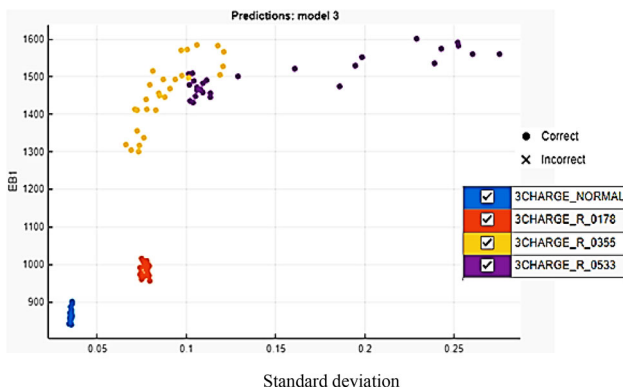


Fig. 11 Scatter plot for 3 HP charge (Set 2)

• Set 3

Applying this method across all dataset 3 yielded an impressive accuracy rate of 98.2% for the set without load and for the set with a 3 HP load. The scatter plots for these two sets, along with their respective prediction outcomes, are

exhibited in Figs. 14 and 15. The confusion matrices of the third set are represented in Figs. 16 and 17.

• Set 4

An impressive accuracy of 96.4% for the unloaded set and 98.2% for the set with a 3 HP load was achieved through the implementation of this method on the entire dataset 4. Figures 18 and 19 present the corresponding scatter plots for both sets, along with the prediction outcomes.

The confusion matrices for both sets are illustrated in Figs. 20 and 21, along with their corresponding prediction results.

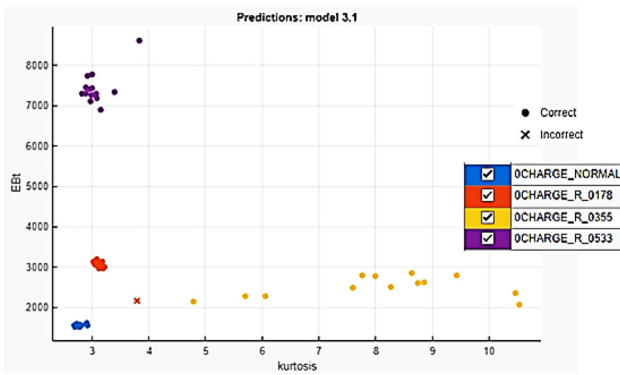


Fig. 14 Scatter plot for 0 charge (Set 3)

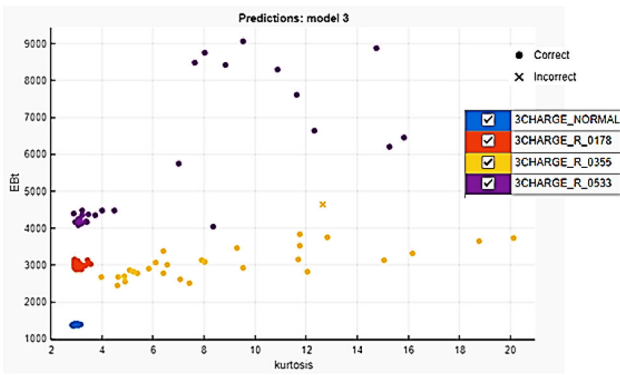


Fig. 15 Scatter plot for 3 HP charge (Set 3)

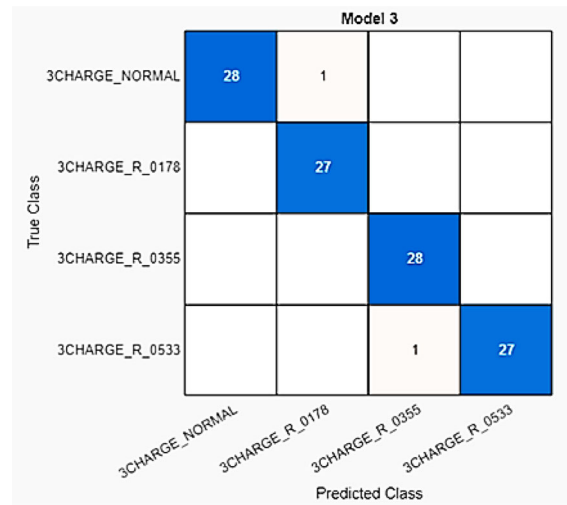


Fig. 17 Confusion matrix for 3 HP charge (Set 3)

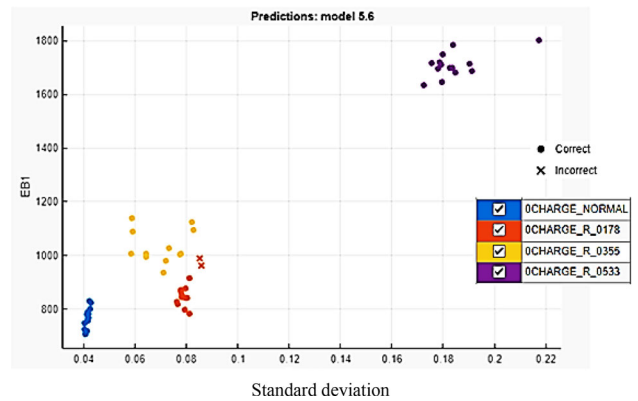


Fig. 18 Scatter plot for 0 charge (Set 4)

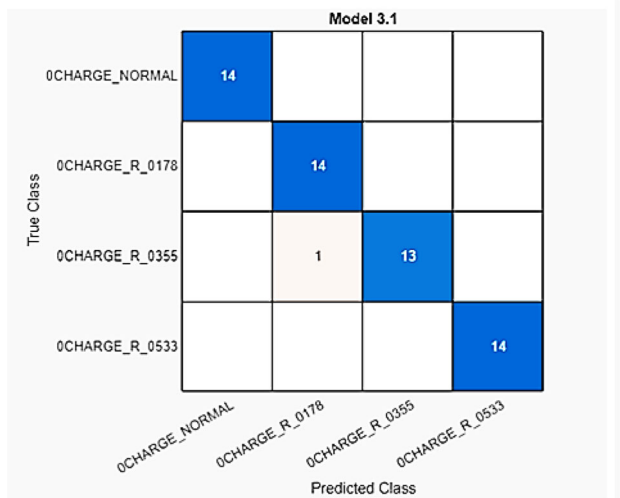


Fig. 16 Confusion matrix with 0 charge (Set 3)

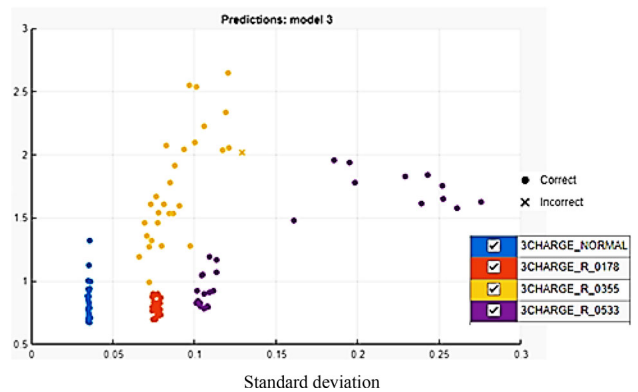


Fig. 19 Scatter plot for 3 HP charge (Set 4)

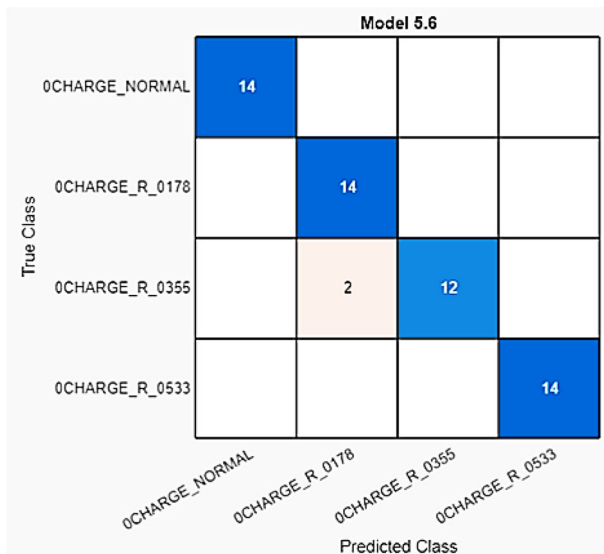


Fig. 20 Confusion matrix with 0 charge (Set 4)

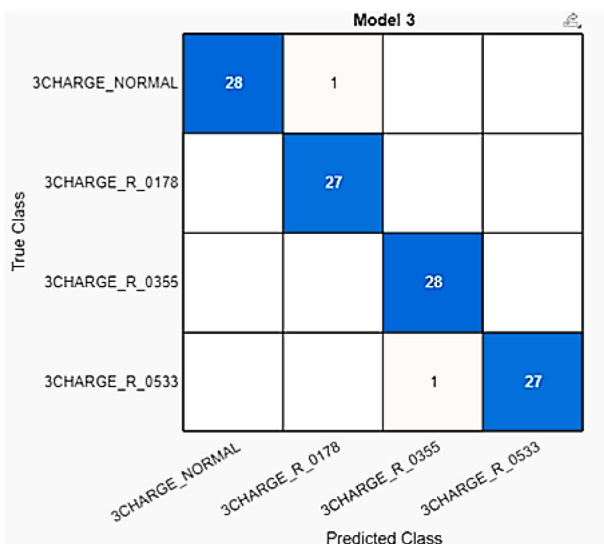


Fig. 21 Confusion matrix for 3 HP charge (Set 4)

5.1.2 Decision Tree (Fine Tree)

The “Fine Tree” model is a variant of the Decision Tree that is characterized by the creation of more detailed and precise trees. It excels in handling complex datasets by breaking down decisions at multiple levels with finesse.

This method is applied to the same database and indicator sets as the SVM method. After obtaining similar results (scatter plots and confusion matrices), only the percentage accuracy of this method for the four sets of indicators is presented in this section.

- Set 1: The application of this method to the complete dataset resulted in an accuracy of 98.2% for both cases, without load and with a load of 3 HP.
- Set 2: When this method was applied to dataset 2, it achieved a significant accuracy of 96.4% for the set without load and 99.1% for the set with a 3 HP load.
- Set 3: The implementation of this method on dataset 3 led to a significant accuracy of 98.2% for the set without load and 95.5% for the set with a load of 3 HP.
- Set 4: Utilizing this method on dataset 4 achieved a noteworthy accuracy of 98.2% for the unloaded set and 99.1% for the set with a 3 HP charge.

5.1.3 Nearest Neighbor (Fine K-NN)

The fine K-NN represents a step forward from the standard K-NN by factoring in the distance between points during the classification process. In contrast to the standard K-NN method where each neighbor has an equal influence on the classification decision, fine K-NN assigns varying weights to each neighbor based on their proximity to the observation under assessment. As a result, neighbors that are closer have a greater influence on the classification, while those further away have a lesser effect. This method of weighting based on distance allows for more accurate decision-making by focusing on the most relevant neighbors, leading to a comprehensive enhancement of the model’s performance.

For this method, it is imperative to provide the accuracy percentage for each indicator set, avoiding the presentation of scatter plot curves and confusion matrices.

- Set 1: The application of this method to the entire of the data resulted in a notable accuracy of 96.4% for the set without charge and 95.5% for the set with a load of 3 HP.
- Set 2: The utilization of this method on the whole dataset 2 led to a noteworthy precision of 92.9% for the set without load and an impressive 98.2% for the set carrying a charge of 3 HP.
- Set 3: The deployment of this method on the full dataset 3 resulted in a significant accuracy of 96.4% for the set without load and 95.5% for the set with a load of 3 HP.
- Set 4: The execution of this method on the comprehensive dataset 4 achieved a noteworthy precision of 91.1% for the set without load and a remarkable 98.2% for the set bearing a load of 3 HP.

5.2 Method Without Data Preprocessing: 1D-CNN

Unlike some traditional methods that require meticulous preparation to extract meaningful features from vibration

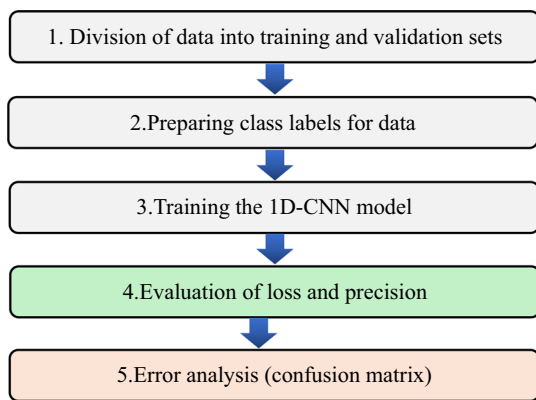


Fig. 22 Steps of training 1D-CNN model

signals, one-dimensional convolutional neural networks (1D-CNN) can autonomously learn from raw data. They can independently identify patterns and significant traits, thus eliminating the need for specific preprocessing procedures.

5.2.1 Choice of 1D-CNN Model Hyperparameters

Hyperparameters play a crucial role in setting up a one-dimensional convolutional neural network (1D-CNN) model. In this specific configuration, the stride is set to 1, determining the distance between each convolution operation applied to the one-dimensional signal. Pooling is performed with the MaxPooling1D method, which reduces the dimensionality of the signal by retaining the maximum values. The loss function is defined as categorical cross-entropy, suitable for multi-class classification tasks. The optimizer used is Adam, which represents a popular optimization algorithm. The batch size is set to 300, determining the number of samples processed in each training iteration. The model is trained over 10 periods, each of which represents a complete pass through the dataset. Finally, the Rectified Linear Unit (ReLU) activation function is applied, promoting nonlinearity in the model by introducing positive activations. This hyperparameter configuration aims to optimize the performance of the 1D-CNN for the specific task at hand.

5.2.2 Application of the 1D-CNN Method to All Data

Figure 22 presents the different stages of training the 1D-CNN model in detail.

5.2.3 Division of Data Into Training and Validation Sets

- Training set

This set is used to train the model by providing it with a large amount of data. It allows the model to acquire the

ability to recognize patterns and understand the relationships between features and different classes. The training set is used to adjust the weights and biases of the model through successive iterations, aiming to minimize the error on these data.

- Validation set

This set is dedicated to evaluating the model's performance during the training process. It plays a crucial role in adjusting the model's parameters, thus preventing overfitting. At each training iteration (or period), the model is evaluated on the validation set to obtain an estimate of its performance on the data it has not yet encountered. These evaluations are used to adjust the parameters, thus optimizing the model's performance while ensuring its generalization to new data.

5.2.4 Class Label Preparation for Data

It is imperative to characterize the various categories or classes to which the data are affiliated. For instance, in the context of bearing fault detection, these classes could be defined as "Normal" or "0CHARGE_R_0178." Subsequently, the data and labels are paired so that each data element is matched with the appropriate label.

5.2.5 Training the 1D-CNN Model

The training set was used to adjust the model's weights, including parameters and biases. This process involves exposing the training data to the various layers of the model and using gradient backpropagation to successively adjust the weights to optimize the model's performance. Throughout this training phase, the regular evaluations of the model's performance were carried out on the validation set. This approach allowed for close monitoring of the model, prevention of overfitting, and, where necessary, adjustments to the parameters to achieve optimal performance.

5.2.6 Monitor Learning: Evaluation of Loss and Accuracy

To evaluate the performance of the 1D-CNN model, two curves were generated: one for the loss and the other for the precision.

- The loss curve

The evolution of the loss function reveals how the gap between the model predictions and the actual labels changes throughout the training epochs. The main objective is to ensure a constant decrease in the loss in each period. A steady

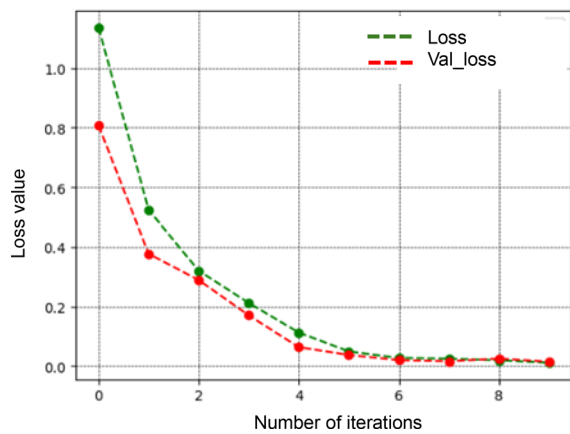


Fig. 23 Loss curve for 0 HP

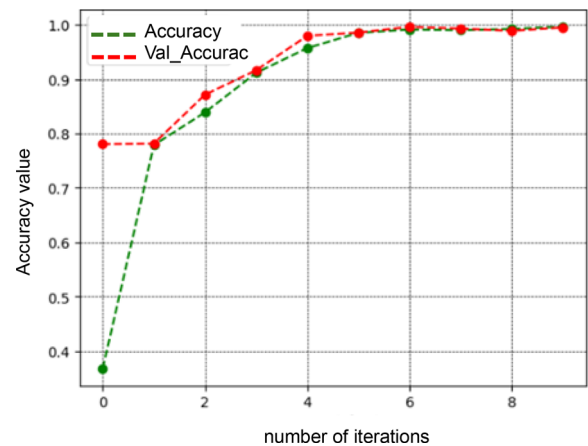


Fig. 25 Accuracy curve for 0 HP

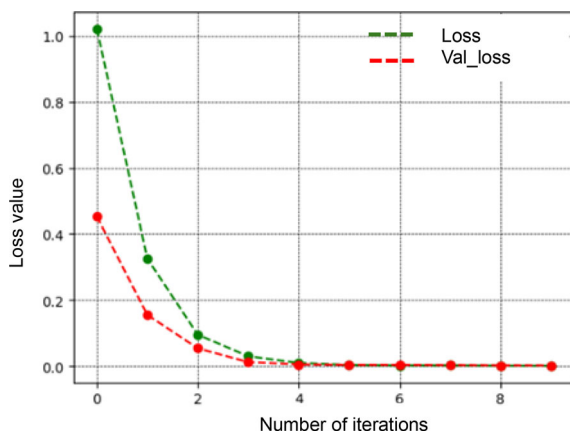


Fig. 24 Loss curve for 3 HP

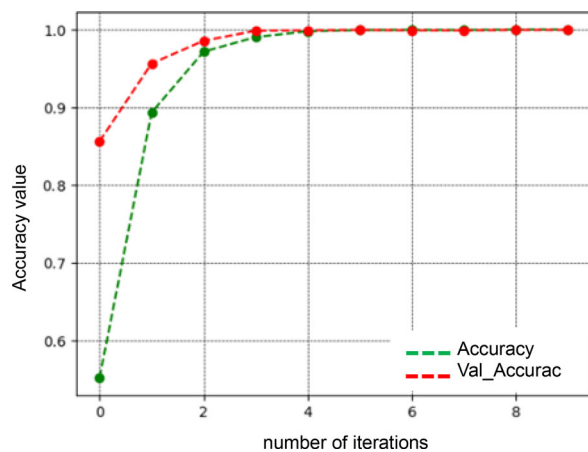


Fig. 26 Accuracy curve for 3 HP

decrease indicates that the model is appropriately incorporating the patterns inherent in the data, as illustrated in Figs. 23 and 24.

"Loss" measures the error between the model predictions and actual labels on the training set, while "loss_val" evaluates the error on a separate validation set not encountered by the model during training. The objective is to minimize both "Loss" and "loss_val" values. However, if the "loss_val" increases while the "Loss" decreases, this could indicate overfitting, indicating that the model fails to generalize effectively to unseen data.

- Precision curve

The trajectory of the precision curve reveals how the model's accuracy, calculated as the number of correct predictions divided by the total number of samples, progresses over epochs. The primary goal is to observe a constant improvement in accuracy over time, thus signaling an improvement in the model's data classification, as evidenced by Figs. 25 and 26. Accuracy measures the proportion of correct predictions

relative to the total number of samples in the training set, while val_accuracy evaluates the proportion of correct predictions on the validation set. The major goal is to achieve high accuracy, both for overall accuracy and for val_accuracy.

5.2.7 Error Analysis

In order to investigate the errors, the confusion matrix was used, a fundamental instrument for studying the inaccuracies of a classification model. This matrix provides a detailed overview of the model's performance for each class, exposing how the model correctly or incorrectly assigns data to each of the classes, as demonstrated in Figs. 27 and 28.

While the vertical axis of the matrix represents the actual classification labels, the horizontal axis represents the prediction labels. The elements of the confusion matrix near or equal to 1 on the main diagonal indicate that the model has successfully made predictions for these classes. More specifically, these values reflect the rates of correct predictions, suggesting that the model has appropriately assigned

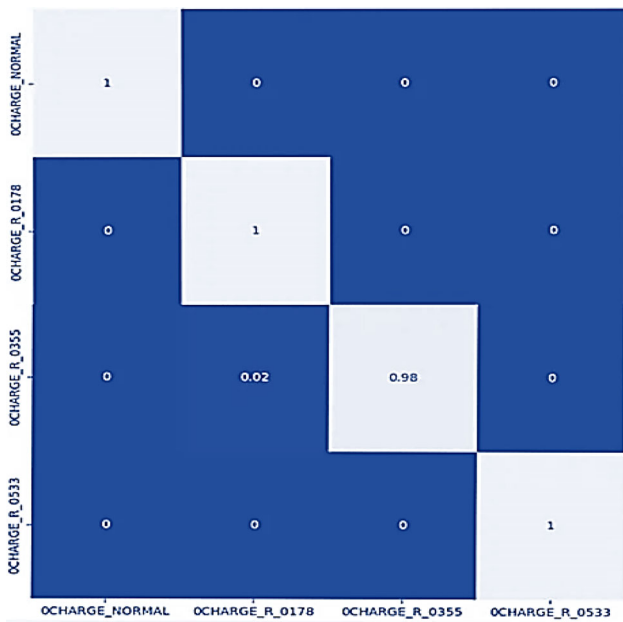


Fig. 27 Confusion matrix for 0 HP

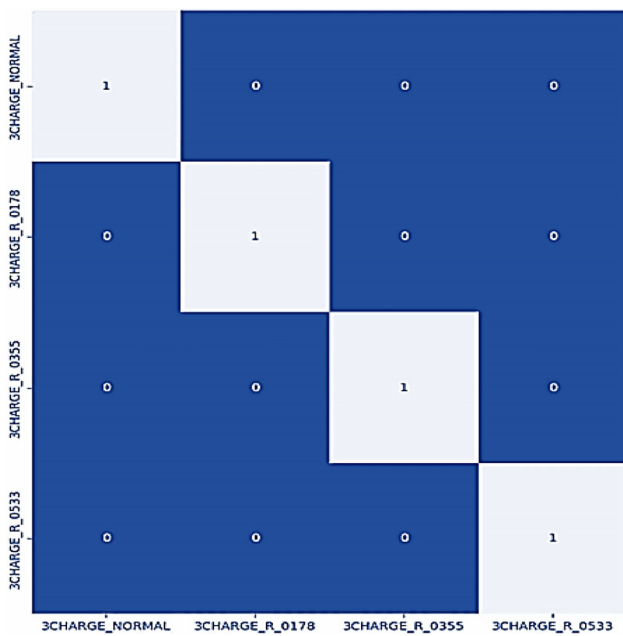


Fig. 28 Confusion matrix for 3 HP

the correct class to the corresponding samples. On the other hand, values outside the main diagonal, equal to 0, reveal that the model has not generated confusion between these classes, as no incorrect prediction has been observed for these specific class pairs. Values different from zero outside the main diagonal indicate classification errors, demonstrating situations in which the model has made incorrect class predictions. These values provide insight into how often the model confuses certain classes with others.

This evaluation underscores the proven benefits of the 1D-CNN method, emphasizing its pivotal role in enhancing maintenance practices. By enabling early fault detection, 1D-CNNs reorganize the adoption of predictive maintenance strategies. The potential integration of 1D-CNNs into predictive maintenance systems confirms a promising future for operational efficiency and industrial equipment reliability.

5.3 Discussions

Table 3 represents the classification accuracies of different methods.

The analysis of the results from the application of the methods on the four datasets, followed by the classification using the 1D-CNN method for detecting bearing defects in the rolling element, highlights significant trends in terms of classification accuracy:

- The 1D-CNN model displays remarkable performance in both scenarios, demonstrating exceptionally high accuracy, and even maintains perfect accuracy in the presence of the maximum load.
- Set 1 (All indicators): The results show high stability of accuracy, demonstrating that the use of the full set of indicators maintains consistent performance, whether without load or with 3 HP. The robustness of this set could be attributed to the diversity of information provided by the full set of indicators, thus reducing the impact of load variation on accuracy.
- Set 2 (Standard deviation, peak-to-peak, EB1, EB2): Although accuracy is generally high, there is a slight variation between sets with and without load. The reduced variance of the indicators in this set may explain the stability of accuracy, but the slight variation suggests that the load can moderately influence performance.
- Set 3 (Kurtosis, effective value, EBt, EB3): Accuracy slightly decreases for some algorithms, indicating that the specific selection of indicators may be more sensitive to the load. The increased variance of the indicators in this set could explain this variation in accuracy, showing that the diversity of features may be more sensitive to load changes.
- Set 4 (Standard deviation, effective value, EBt, EB1): Performance generally remains high with a slight variation between sets with and without load. The relative stability of this set suggests that the selection of specific indicators maintains robust accuracy, even with load variations.
- Influence of the load: Overall, the addition of the 3 HP load seems to have a positive impact on accuracy in most cases, indicating that load variation can help improve the classification ability of bearing defects.

Table 3 Classification accuracy

	Set 1		Set 2		Set 3		Set 4	
	0 HP	3 HP	0 HP	3 HP	0 HP	3 HP	0 HP	3 HP
Fine Tree	98.2%	98.2%	96.4%	99.1%	98.2%	95.5%	98.2%	99.1%
Medium Gaussian SVM	98.2%	98.2%	98.2%	99.1%	98.2%	98.2%	96.4%	98.2%
Fine K-NN	96.4%	95.5%	92.9%	98.2%	96.4%	95.5%	91.1%	98.2%
	0 HP				3 HP			
1D-CNN	99.4%				100%			

The analysis of the results demonstrates that the diversity of indicator sets significantly influences the stability of accuracy in bearing fault classification. The full set of indicators maintains consistent performance, while narrower sets may exhibit varying sensitivity to the load. Generally speaking, the addition of the 3 HP load commonly improves accuracy, underlining its positive impact on the ability of the 1D-CNN method to distinguish bearing faults.

6 Conclusion

The present research work has demonstrated the considerable potential of artificial intelligence techniques, specifically Machine Learning and Deep Learning, in diagnosing bearing faults. A key aspect of this study is the comparative analysis conducted between the different methods, providing valuable insights into their respective strengths and limitations in the context of bearing fault diagnosis.

The study has revealed that both methods that require data preprocessing and those that do not can achieve high precision in fault detection. Notably, the 1D-CNN model achieved an impressive accuracy rate of 100% for 3 HP, demonstrating its resilience and high accuracy even under varying load conditions (99.4% for 0 HP). Regarding the comparison among the three methods that require preprocessing, none can be definitively deemed more accurate than the others. Their accuracy percentages are very similar, and the results depend on the engine load and the database selected. The incorporation of four distinct datasets, chosen based on the distribution curves of the indicators relative to the number of observations, has added depth to the study. The results emphasize the critical role of feature selection and the influence of load variations on the accuracy of fault diagnosis.

This study contributes to the ongoing advancements in improving the accuracy and efficiency of bearing fault diagnosis, thus facilitating more dependable and effective maintenance strategies across various sectors. In the future, the focus will be on developing more efficient indicators and

incorporating other datasets to further enhance the robustness and applicability of these methods in bearing fault diagnosis.

Funding The author(s) received no financial support for the research, authorship, and/or publication of this article.

Declarations

Conflict of interest The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

References

- Hasan, M.J.; Sohaib, M.; Kim, J.-M.: An explainable AI-based fault diagnosis model for bearings. *Sensors* **21**, 4070 (2021). <https://doi.org/10.3390/s21124070>
- Kibrete, F.; Woldemichael, D.E.: Applications of artificial intelligence for fault diagnosis of rotating machines: A review. In *International Conference on Advances of Science and Technology*, pp. 41–62. Springer Nature Switzerland, Cham, November 2022
- Sawaqed, L.S.; Alrayes, A.M.: Bearing fault diagnostic using machine learning algorithms. *Prog. Artif. Intel.* **9**, 341–350 (2020). <https://doi.org/10.1007/s13748-020-00217-z>
- Iqbal, M.; Madan, A.K.: Artificial intelligence-based bearing fault diagnosis of rotating machine to improve the safety of power system. In *International Conference on Renewable Power*, pp. 933–942. Springer Nature Singapore, Singapore, March 2023
- Lu, Y.; Xie, R.; Liang, S.Y.: CEEMD-assisted kernel support vector machines for bearing diagnosis. *Int. J. Adv. Manuf. Technol.* **106**, 3063–3070 (2020). <https://doi.org/10.1007/s00170-019-04858-w>
- Hu, W.; Gu, F.; Chen, S.: Large data and ai analysis based online diagnosis system application of steel ladle slewing bearing. In: Ball, A.; Gelman, L.; Rao, B. (Eds.) *Advances in Asset Management and Condition Monitoring: COMADEM*, pp. 1519–1527. Springer International Publishing, Cham (2020)
- Toma, R.N.; Prosvirin, A.E.; Kim, J.M.: Bearing fault diagnosis of induction motors using a genetic algorithm and machine learning classifiers. *Sensors* **20**(7), 1884 (2020)
- Goyal, D.; Choudhary, A.; Pabla, B.S., et al.: Support vector machines based non-contact fault diagnosis system for bearings. *J. Intell. Manuf.* **31**, 1275–1289 (2020). <https://doi.org/10.1007/s10845-019-01511-x>



9. Sharma, A.: Fault diagnosis of bearings using recurrences and artificial intelligence techniques. *J. Nondestruct. Eval.* **5**(3), 031004 (2022). <https://doi.org/10.1115/1.4053773>
10. Samanta, B.; Al-Balushi, K.R.; Al-Araimi, S.A.: Artificial neural networks and genetic algorithm for bearing fault detection. *Soft. Comput.* **10**, 264–271 (2006)
11. Unal, M.; Onat, M.; Demetgul, M.; Kucuk, H.: Fault diagnosis of rolling bearings using a genetic algorithm optimized neural network. *Measurement* **58**, 187–196 (2014)
12. Ettefagh, M.M.; Ghaemi, M.; Asr, M.Y.: Bearing fault diagnosis using hybrid genetic algorithm K-means clustering. In: 2014 IEEE International Symposium on Innovations in Intelligent Systems and Applications (INISTA) Proceedings, pp. 84–89. IEEE, June 2014
13. Xiong, J.; Zhang, Q.; Liang, Q.; Zhu, H.; Li, H.: Combining the multi-genetic algorithm and support vector machine for fault diagnosis of bearings. *Shock. Vib.* **2018**, 3091618 (2018)
14. Jayaswal, P.; Verma, S.N.; Wadhvani, A.K.: Application of ANN, fuzzy logic and wavelet transform in machine fault diagnosis using vibration signal analysis. *J. Qual. Maint. Eng.* **16**(2), 190–213 (2010)
15. Zheng, J.; Cheng, J.; Yang, Y.: A rolling bearing fault diagnosis approach based on LCD and fuzzy entropy. *Mech. Mach. Theory* **70**, 441–453 (2013)
16. Wang, Z.; Zhang, Q.; Xiong, J.; Xiao, M.; Sun, G.; He, J.: Fault diagnosis of a rolling bearing using wavelet packet denoising and random forests. *IEEE Sens. J.* **17**(17), 5581–5588 (2017)
17. Vakharia, V.; Gupta, V.K.; Kankar, P.K.: Efficient fault diagnosis of ball bearing using relief and random forest classifier. *J. Braz. Soc. Mech. Sci. Eng.* **39**(8), 2969–2982 (2017)
18. Han, T.; Jiang, D.; Zhao, Q.; Wang, L.; Yin, K.: Comparison of random forest, artificial neural networks and support vector machine for intelligent diagnosis of rotating machinery. *Trans. Inst. Meas. Control.* **40**(8), 2681–2693 (2018)
19. Sharma, A.; Jigyasu, R.; Mathew, L.; Chatterji, S.: Bearing fault diagnosis using weighted K-nearest neighbor. In: 2018 2nd International Conference on Trends in Electronics and Informatics (ICOEI), pp. 1132–1137. IEEE, May 2018
20. Wang, Q.; Liu, Y.B.; He, X.; Liu, S.Y.; Liu, J.H.: Fault diagnosis of bearing based on KPCA and KNN method. *Adv. Mater. Res.* **986**, 1491–1496 (2014)
21. Kumar, H.S.; Manjunath, S.H.: Use of empirical mode decomposition and K-nearest neighbour classifier for rolling element bearing fault diagnosis. *Materials Today: Proceedings*, **52**, 796–801, (2022)
22. Vishwendra, M.A.; Salunkhe, P.S.; Patil, S.V.; Shinde, S.A.; Shinde, P.V.; Desavale, R.G.; Dharwadkar, N.V.: A novel method to classify rolling element bearing faults using K-nearest neighbor machine learning algorithm. *ASCE-ASME J. Risk Uncertain. Eng. Syst. Part B: Mech. Eng.* **8**(3), 031202 (2022)
23. Andre, A.B.; Beltrame, E.; Wainer, J.: A combination of support vector machine and k-nearest neighbors for machine fault detection. *Appl. Artif. Intell.* **27**(1), 36–49 (2013)
24. Amarnath, M.; Sugumaran, V.; Kumar, H.: Exploiting sound signals for fault diagnosis of bearings using decision tree. *Measurement* **46**(3), 1250–1256 (2013)
25. Demetgul, M.: Fault diagnosis on production systems with support vector machine and decision trees algorithms. *Int. J. Adv. Manuf. Technol.* **67**, 2183–2194 (2013)
26. Yi, X.J.; Chen, Y.F.; Hou, P.: Fault diagnosis of rolling element bearing using Naïve Bayes classifier. *Vib. Proced.* **14**, 64–69 (2017). <https://doi.org/10.21595/vp.2017.19153>
27. Zhang, N.; Wu, L.; Yang, J.; Guan, Y.: Naive bayes bearing fault diagnosis based on enhanced independence of data. *Sensors* **18**(2), 463 (2018)
28. Muralidharan, V.; Sugumaran, V.: A comparative study of Naïve Bayes classifier and Bayes net classifier for fault diagnosis of monoblock centrifugal pump using wavelet analysis. *Appl. Soft Comput.* **12**(8), 2023–2029 (2012)
29. Guo, X.; Chen, L.; Shen, C.: Hierarchical adaptive deep convolution neural network and its application to bearing fault diagnosis. *Measurement* **93**, 490–502 (2016)
30. Chen, Z.; Deng, S.; Chen, X.; Li, C.; Sanchez, R.V.; Qin, H.: Deep neural networks-based rolling bearing fault diagnosis. *Microelectron. Reliab.* **75**, 327–333 (2017)
31. Zhang, J.; Yi, S.; Liang, G.U.O.; Hongli, G.A.O.; Xin, H.O.N.G.; Hongliang, S.O.N.G.: A new bearing fault diagnosis method based on modified convolutional neural networks. *Chin. J. Aeronaut.* **33**(2), 439–447 (2020)
32. Sun, Y.; Li, S.: Bearing fault diagnosis based on optimal convolution neural network. *Measurement* **190**, 110702 (2022)
33. Sinitsin, V.; Ibryaeva, O.; Sakovskaya, V.; Eremeeva, V.: Intelligent bearing fault diagnosis method combining mixed input and hybrid CNN-MLP model. *Mech. Syst. Signal Process.* **180**, 109454 (2022)
34. Li, H.; Huang, J.; Ji, S.: Bearing fault diagnosis with a feature fusion method based on an ensemble convolutional neural network and deep neural network. *Sensors* **19**(9), 2034 (2019)
35. Althubaiti, A.; Elasha, F.; Teixeira, J.A.: Fault diagnosis and health management of bearings in rotating equipment based on vibration analysis—a review. *J. Vib.* **24**(1), 46–74 (2022)
36. Mohd Ghazali, M.H.; Rahiman, W.: Vibration analysis for machine monitoring and diagnosis: a systematic review. *Shock. Vib.* **2021**(1), 9469318 (2021)
37. Van Hecke, B.; Qu, Y.; He, D.: Bearing fault diagnosis based on a new acoustic emission sensor technique. In: Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability, 229(2), 105–118, (2015)
38. Nabhan, A.; Ghazaly, N.; Samy, A.; Mousa, M.O.: Bearing fault detection techniques—a review. *Turk. J. Eng. Sci. Technol.* **3**(2), 1–18 (2015)
39. Peng, Z.K.; Peter, W.T.; Chu, F.L.: A comparison study of improved Hilbert-Huang transform and wavelet transform: application to fault diagnosis for rolling bearing. *Mech. Syst. Signal Process.* **19**(5), 974–988 (2005)
40. Neupane, D.; Seok, J.: Bearing fault detection and diagnosis using case western reserve university dataset with deep learning approaches: a review. *Ieee Access* **8**, 93155–93178 (2020)
41. Kiranyaz, S.; Avci, O.; Abdeljaber, O.; Ince, T.; Gabbouj, M.; Inman, D.J.: 1D convolutional neural networks and applications: a survey. *Mech. Syst. Signal Process.* **151**, 107398 (2021)
42. Eren, L.: Bearing fault detection by one-dimensional convolutional neural networks. *Math. Probl. Eng.* **2017**, 1–9 (2017)
43. Yang, K.; Zhao, L.; Wang, C.: A new intelligent bearing fault diagnosis model based on triplet network and SVM. *Sci. Rep.* **12**(1), 5234 (2022)
44. Loparo, K.: Bearings Vibration Data Set, CaseWestern Reserve University. Available online: http://www.eecs.case.edu/laboratory/bearing/welcome_overview.htm (accessed on 20 July 2012)
45. Smith, W.A.; Randall, R.B.: Rolling element bearing diagnostics using the Case Western Reserve University data: a benchmark study. *Mech. Syst. Signal Process.* **64**, 100–131 (2015)
46. Zhang, R.; Peng, Z.; Wu, L.; Yao, B.; Guan, Y.: Fault diagnosis from raw sensor data using deep neural networks considering temporal coherence. *Sensors* **17**(3), 549 (2017)



47. Boudiaf, A.; Moussaoui, A.; Dahane, A.; Atoui, I.: A comparative study of various methods of bearing faults diagnosis using the case Western Reserve University data. *J. Fail. Anal. Prev.* **16**(2), 271–284 (2016)
48. Liu, W.; Zhang, Z.; Zhang, J.; Huang, H.; Zhang, G.; Peng, M.: A novel fault diagnosis method of rolling bearings combining convolutional neural network and transformer. *Electronics* **12**(8), 1838 (2023)
49. Jaber, A.A.: Diagnosis of bearing faults using temporal vibration signals: a comparative study of machine learning models with feature selection techniques. *J. Fail. Anal. Prev.* **24**, 752–768 (2024)
50. Liu, J.; Xu, Z.; Zhou, L.; Yu, W.; Shao, Y.: A statistical feature investigation of the spalling propagation assessment for a ball bearing. *Mech. Mach. Theory* **131**, 336–350 (2019)

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.