



Original software publication

BEARING-FDD: An early detection and diagnosis tool for bearing faults in rotating machinery

L. Magadán^{a,*}, C. Ruiz-Cárcel^b, J.C. Granda^a, F.J. Suárez^a, A. Menéndez-González^a, A. Starr^b^a Department of Computer Science and Engineering, University of Oviedo, Gijón, 33204, Spain^b School of Aerospace, Transport and Manufacturing, University of Cranfield, Bedford, MK43 0AL, United Kingdom

ARTICLE INFO

Keywords:

Fault diagnosis
Fault classification
Rotating machinery
Interpretable AI
Explainable AI

ABSTRACT

This paper presents the design and implementation of a web tool offering an innovative method for detecting, diagnosing and classifying bearing faults in rotating machinery under limited data conditions, providing explainability and interpretability of the results obtained. The tool uses a machine learning model to detect and diagnose bearing faults. A monotonic smoothed stacked autoencoder builds a health indicator without requiring feature extraction, making the tool useful without the need for specialized staff. The tool generates explainability and interpretability reports with a correlation analysis between the health indicator and well-known engineering features and easily interpretable details on the diagnosed faults. The tool includes the option to use preloaded state-of-the-art datasets, while also allowing users to upload their own datasets to analyze vibration data from real industrial equipment.

Code metadata

Current code version
Permanent link to code/repository used for this code version
Permanent link to Reproducible Capsule
Legal Code License
Code versioning system used
Software code languages, tools, and services used
Compilation requirements, operating environments
Support email for questions

v1.0.0
<https://github.com/SoftwareImpacts/SIMPAC-2025-290>
MIT License
Git
Python, Spring, SQL
Python v3.11, Java 17
magadanluis@uniovi.es

1. Motivation and significance

In recent years, the increasing digitalization of industrial companies has driven demand for more sophisticated, efficient and safe machinery. Most industrial processes depend on mechanical and electromechanical systems, with rotating machinery (RM) accounting for approximately 40% of these systems. These components are critical to maintaining the effectiveness and safety of industrial processes, as they are especially prone to failure due to extended operating times and challenging working conditions [1].

In electric motors, various faults affect their overall health. Bearing faults account for about 50% of faults, while the stator contributes around 30%, and the rotor approximately 10%. The remaining faults

are caused by gear issues and looseness. Bearing faults can occur in different parts of the bearing which are explained below. The severity and manifestation of the damage depend on the location of the fault. Understanding and effectively diagnosing bearing faults is crucial for ensuring the reliable operation of electric motors [2].

- Outer-race bearing faults: they correspond to faults that occur in the outer ring of the bearing, which encloses the rolling elements [3]. This fault can be detected by analyzing the spectrum of the vibration signal, which will show harmonic peaks in the Ball Pass Frequency Outer (BPFO) or sidebands separated by the BPFO from a central frequency component due to modulation. BPFO is computed as indicated in Eq. (1), where N_B corresponds to the number of balls of the bearing, B_D is the ball diameter, P_D is the

DOI of original article: <https://doi.org/10.1016/j.aei.2024.102909>.

* Corresponding author.

E-mail address: magadanluis@uniovi.es (L. Magadán).

pitch diameter, β is the contact angle and F is the shaft frequency.

$$BPFO = F \cdot \frac{N_B}{2} \cdot \left(1 - \frac{B_D}{P_D} \cdot \cos(\beta)\right) \quad (1)$$

- Inner-race bearing faults: in this case, the fault occurs in the inner ring of the bearing, which is in contact with the rolling elements for smooth rotation [3]. The spectrum of a signal where an inner-race bearing fault exists is characterized by several harmonic peaks in the Ball Pass Frequency Inner (BPFI), which is calculated as shown in Eq. (2).

$$BPFI = F \cdot \frac{N_B}{2} \cdot \left(1 + \frac{B_D}{P_D} \cdot \cos(\beta)\right) \quad (2)$$

- Ball bearing faults: they refer to defects in the metal balls located between the inner and the outer races [3]. The spectrum of a ball bearing faulty signal is characterized by harmonics in the Ball Spin Frequency (BSF) or sidebands separated by the BSF from a central frequency component due to modulation. Eq. (3) shows how to calculate BSF. This fault is usually accompanied by outer- or inner-race faults.

$$BSF = F \cdot \frac{P_D}{B_D} \cdot \left[1 - \left(\frac{B_D}{P_D} \cdot \cos(\beta)\right)^2\right] \quad (3)$$

- Cage bearing faults: these imply malfunctions in the cage of the bearings. Its function is to maintain the proper spacing and alignment of rolling elements between the inner and outer races [3]. The spectrum of a cage bearing faulty signal is characterized by the existence of harmonics in the Fundamental Train Frequency (FTF), and is usually accompanied by outer- or inner-race faults. FTF can be calculated as shown in Eq. (4).

$$FTF = F \cdot \frac{1}{2} \cdot \left(1 - \frac{B_D}{P_D} \cdot \cos(\beta)\right) \quad (4)$$

Traditionally, fault detection has relied on classic signal processing techniques. However, the continuous advancement of machine learning (ML) algorithms, particularly their efficiency in early fault diagnosis using deep learning techniques, has made fault diagnosis and classification based on ML a highly popular research field in recent years [4]. Convolutional Neural Networks (CNNs) are widely used for fault diagnosis and classification in RM [5,6]. The primary limitation of these approaches, however, is that deep learning algorithms typically require large amounts of data for accurate fault classification and diagnosis. In addition, data should be balanced, which is quite complex in industrial facilities, as faulty samples are difficult to be gathered. To address this issue, techniques such as Generative Adversarial Networks (GANs) are often used, generating synthetic data to train models [7,8]. Despite their potential, these methods remain sensitive to hyperparameter configuration. An improper balance between the generator and discriminator can either hinder learning or result in the continuous generation of similar samples.

While ML algorithms achieve good results, many operate as black-box models, making it difficult to understand the logic behind their decisions. This often renders their outputs impractical for practitioners. Interpretable machine learning focuses on developing models that are inherently understandable, while explainable machine learning seeks to provide post-hoc explanations for black-box models [9]. Prioritizing explainability in the development of new ML algorithms is crucial for enhancing model transparency and identifying errors or biases [10, 11]. This ensures that even non-specialists can easily comprehend how results are produced. Additionally, explainability is essential in highly regulated, safety-critical industries, where all processes must be transparent and traceable for certification purposes.

In the work presented in [12], an explainable bearing-fault classification and diagnosis algorithm built around a Monotonic Smoothed

Stacked Autoencoder (MS2AE) is designed. This neural architecture receives raw vibration signals and is trained using only healthy data, without requiring faulty data for training and handcrafted features, as the feature extraction is directly performed by the MS2AE. The MS2AE encoder extracts monotonic latent features, and then applies a smoothing layer to produce a clean, stable Health Index (HI) for each sample. After training, the system computes HI values for all healthy and unlabeled samples, determines a statistical threshold based on the 95th percentile of healthy HI values, and identifies the first faulty point once five consecutive samples exceed that threshold. Correlation matrices between the HI and traditional time- and frequency-domain features provide interpretability, ensuring that the learned representation aligns with known engineering indicators of degradation.

Once a sample is classified as faulty, the system performs a multistage diagnostic procedure. First, it applies Dynamic Time Warping (DTW) recursively across healthy samples to generate a single denoised healthy signal, then computes the DTW difference between this baseline and the faulty sample to isolate the fault signature. A kurtogram-guided bandpass filter identifies the optimal frequency band containing impulsive fault information, and a Butterworth filter followed by envelope analysis extracts the relevant modulated components. The algorithm then performs FFT analysis on both the filtered and unfiltered isolated fault signals, searching for characteristic harmonics. Based on harmonic presence patterns and the HI magnitude, the system determines both the fault type (outer race, inner race, ball, or cage) and the degradation stage (early, medium, or late).

A web tool has been developed to implement the aforementioned work [12] for the early detection, diagnosis, and classification of bearing faults. This application allows users to either utilize preloaded public vibration datasets or upload their own, requiring only basic information such as bearing frequencies, sampling frequency, load, and bearing type, along with two CSV files containing vibration acceleration samples from both normal motor operation and the samples to be analyzed. The application performs automated analysis without human intervention by employing the MS2AE, which processes the raw vibration data and compresses it into a HI that determines the machinery's health status. Furthermore, it provides explainability reports by comparing the automatically generated HI with features that could be manually extracted by qualified staff. It also offers interpretability reports by detailing the reasons behind the diagnosed fault, ensuring that even users with limited expertise can understand the results.

The remainder of the paper is organized as follows. Section 2 provides a detailed description of the software, including its architecture and functionalities. Section 3 presents an illustrative example of using the proposed software for bearing fault detection, diagnosis and classification, showcasing various approaches with state-of-the-art datasets. A quantitative analysis of the results obtained and a performance comparison with other state-of-the-art methods is presented in Section 4. Section 5 justifies the impact of the proposed web tool in diagnosing bearing faults in RM. Finally, the concluding remarks and future directions are discussed in Section 6.

2. Software description

The developed web tool provides explainable and interpretable bearing fault detection, diagnosis and classification. Users can use preloaded datasets or upload their own to perform bearing fault analysis. Uploaded raw vibration data is compressed into a HI that determines the stage of degradation of the bearing. The primary goal of the software is to detect, diagnose and classify faults in rotating machinery without any human intervention (e.g., manually extracting features), while providing users with explainable and interpretable results. The robustness of the proposed tool, along with the explainability and interpretability of the results, makes it not only usable by non-specialists but also reliable for industries where strict regulations must be followed.

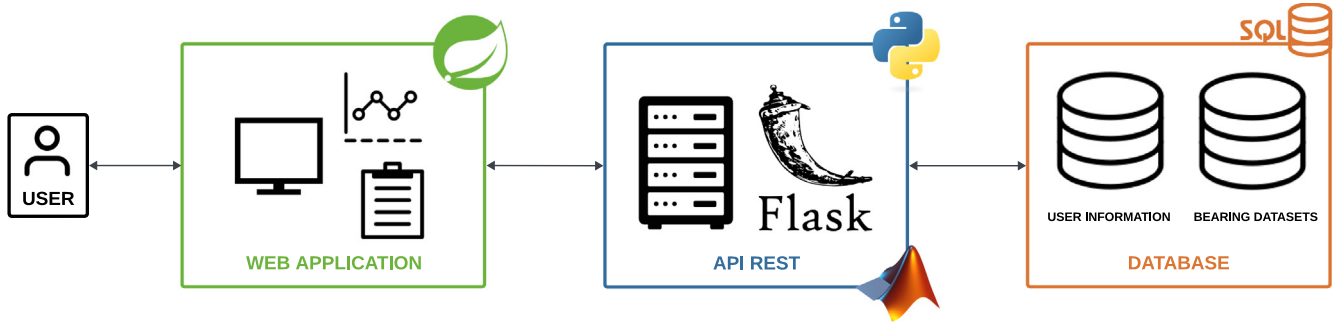


Fig. 1. Software architecture.

2.1. Software architecture

The architecture of the proposed software is shown in Fig. 1, consisting of three main components: a web application developed with Java Spring, a REST API implemented in Python and a relational database.

The web application, with its user-friendly interface, allows users to analyze datasets extracted from rotating machinery. Since the goal of the application is to enable users to conduct their own analyses, they can upload their own datasets for fault detection, diagnosis and classification. Additionally, users can generate detailed reports on the faults detected. These reports focus on providing both explainability and interpretability of the results.

The REST API is responsible for providing the web application with all the necessary functions for fault detection, diagnosis, and classification. It implements the core methods from the system described in [12], utilizing a combination of machine learning and signal processing techniques. For early fault detection, the API uses a MS2AE model. In addition, it leverages methods such as Dynamic Time Warping, Kurtograms, and Fast Fourier Transforms for comprehensive fault detection, diagnosis, and classification. The API also handles file management, enabling users to upload their own datasets. This feature is essential, allowing the software to be used with custom data for more tailored fault analysis.

2.2. Software functionalities

The functions provided by the components can be grouped as follows:

- Web Application: User authentication and visual interface.
- API Rest: User management, file management, fault detection, diagnosis and classification, and report generation.
- Database: storage of user data and bearing datasets.

Considering the functions offered by the set of components and emphasizing the robustness of the three components working together, the following functionalities are achieved:

- ML model training: Machine learning model training using only healthy bearing data.
- Execution of dataset analysis: Machine learning and signal processing techniques for detecting, diagnosing and classifying bearing faults in public well-known datasets (IMS and XJTU-SY).
- User uploads: Users are allowed to upload their own datasets for fault detection, diagnosis and classification.
- Sample selection: Users may select the number of samples to be analyzed from the available datasets.
- Report generation and export: Creation and exportation of the explainability and interpretability reports of the obtained results.
- Two different views: The web application provides an administrator view to manage all datasets uploaded by users and their associated information, and a user view that displays only the preloaded datasets and those uploaded by the user.
- Efficient file storage and management for each user.

3. Illustrative examples

This section presents examples of using the developed tool. Once the user logs into the application, the screen shown in Fig. 2 appears. Two important panels are visible: the panel on the left corresponds to preloaded datasets, while the panel on the right shows the datasets uploaded by the user. Each of these panels is explained separately below.

In the case of preloaded datasets, the dataset to be analyzed must be selected and then the OK button pressed. The window shown in Fig. 3 appears, where a panel on the left displays all the relevant information about the selected dataset and several empty graphs are shown on the right panel. In the graphs, four color indicators represent the degradation stage and type of associated bearing faults. In this work, the bearing faults considered are outer-race, inner-race, ball and cage bearing faults.

The color indicators are as follows: green (healthy), yellow (early degradation stage), orange (medium degradation stage), and red (extreme degradation stage). Additionally, at the bottom, there is another panel that allows the user to enter the number of samples from the first samples of the dataset to be used as healthy samples for training, as well as the number of samples to be analyzed. To illustrate its functionality, the IMS-2 dataset is used, specifying the first 300 samples of the dataset as healthy samples and selecting from sample #530 to sample #550 for analysis.

The analysis is carried out after clicking **Run**. Fig. 4 shows the results of the analysis. The tool has detected, diagnosed and classified an outer-race and cage bearing fault at an early degradation stage, indicated by the presence of 1X and 2X BPFO (236 and 472 Hz) harmonics, as well as 1X, 5X, and 6X FTF (15, 75, and 90 Hz) harmonics. Furthermore, the user has the option to download two reports: one that provides explainability regarding the machine learning procedure and another that offers interpretability of the results obtained.

In the explainability report, correlation matrices between the HI obtained and typical engineering features, such as Root Mean Square (RMS), Kurtosis, Skewness or the amplitudes of some relevant bearing harmonics, are presented, both for the time and frequency domains. This is done to ensure that the HI, despite being generated automatically, exhibits a high correlation with traditionally used engineering features that require prior knowledge for manual extraction. In the example shown in Fig. 4, the variables that are highly correlated with the HI are the RMS (68% correlation) and the amplitude of the 1X BPFO harmonic of the filtered signal (50%). These two features are significantly correlated with the type of fault detected. The obtained correlation matrices are shown in Figs. 5(a) and 5(b). The reader may refer to [12] for the description of all the features included in the matrices. The interpretability report compiles the data shown in the graphs in Fig. 4, allowing them to be exported in PDF format and providing a description of the fault identified and the method of detection.

Users also have the option to upload their own datasets (with a maximum of five datasets per user), which will only be visible to the



Universidad de Oviedo



Detection and Diagnosis of Bearing Faults on Electric Motors

Fig. 2. BEARING-FDD: home page.

Fig. 3. BEARING-FDD: preloaded dataset details.

user who uploaded them. To do this, the user must select the **New Dataset** button in the window shown in Fig. 2. This action will open a window where the user must enter both the name of the dataset and various known values of the analyzed rotating machinery, such as the sampling frequency, BPFO, or BPFI, as well as the type of bearing (see Fig. 6). Once saved, the dataset will appear in the **Saved Datasets** section (see Fig. 2). In this case, a dataset named **user** is shown.

Once this new dataset is selected and the **OK** button pressed, a window opens, displaying the dataset information and the option to load vibration data (see Fig. 7). After loading the data, the page view changes to show a new layout that includes a functionality to let users load new data for analysis (see Fig. 8). In the example shown, a file with new 50 samples to be analyzed is uploaded and the user selects the number of previously loaded samples corresponding to healthy conditions of the analyzed bearing. The analysis process is similar to the

previously described, and the results obtained are analogous to those shown in Fig. 4.

Apart from these illustrative examples, Table 1 summarizes the results obtained for each dataset, analyzing three different set of samples, each at a distinct degradation level. The most relevant results are presented, including the detected faults, the level of degradation and the diagnosis that reveals the occurring faults.

4. Performance evaluation

To complement the qualitative examples presented in Section 3, a quantitative evaluation of the proposed tool was performed using all the considered datasets. Table 2 reports the accuracy, precision, recall, and F1-score obtained for each dataset when using the MS2AE model for fault detection and diagnosis. These results show that the proposed tool achieves consistent and robust diagnostic performance

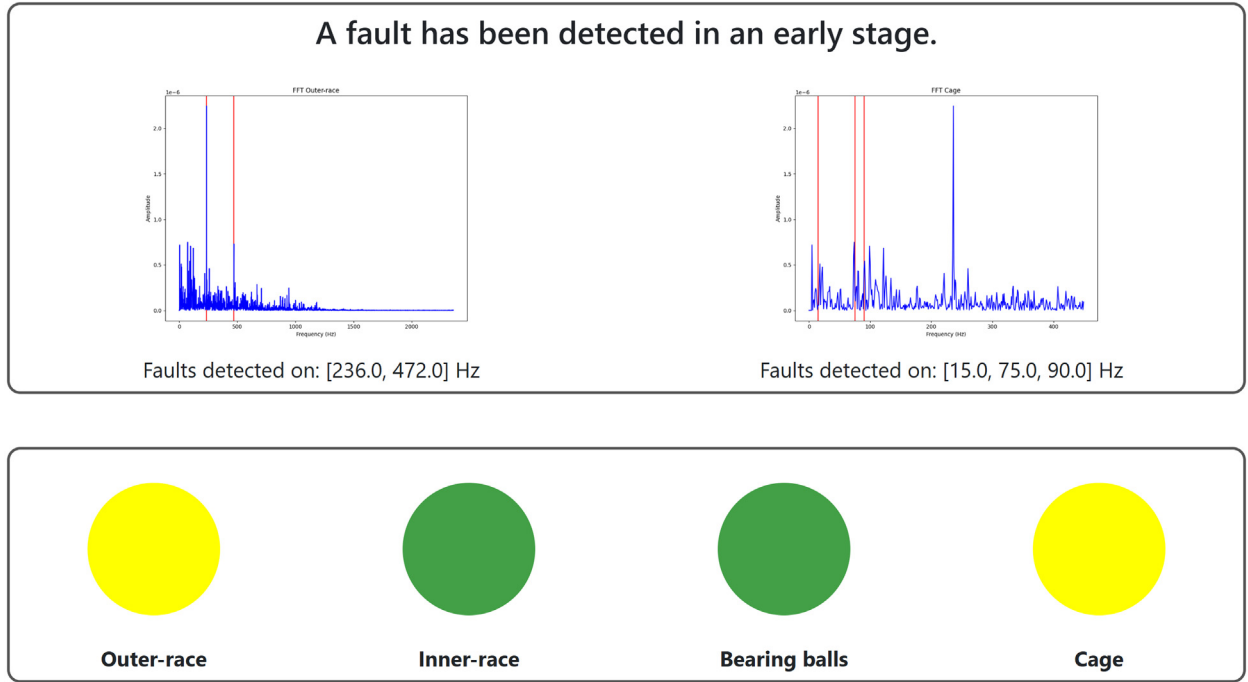


Fig. 4. BEARING-FDD: diagnosis of IMS-2.

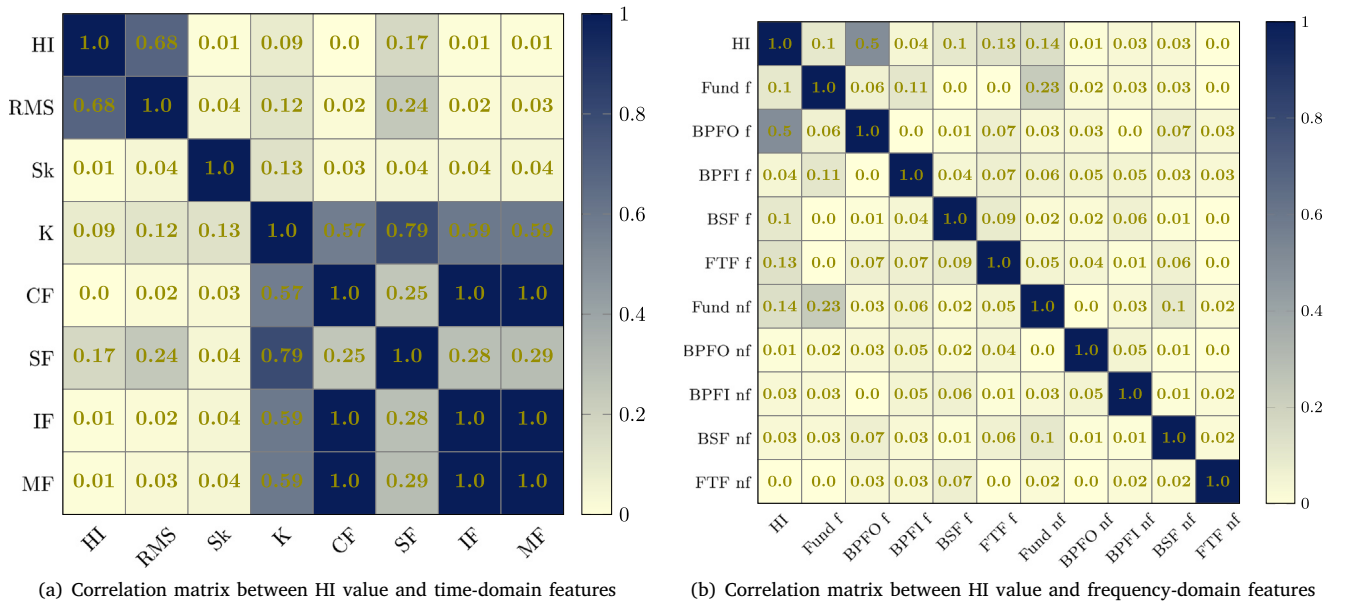


Fig. 5. Correlation matrices between HI value and engineering features in IMS-2 dataset.

across datasets of different sizes and operating conditions. These metrics confirm that the MS2AE model, despite requiring only healthy samples for training, generalizes reliably to multiple degradation stages and fault types.

In order to also evaluate the early fault-detection capabilities of the proposed approach, the First Faulty Point (FFP) detected by the tool was compared with several state-of-the-art methods reported in the literature. FFP is essential for predictive maintenance. Identifying the FFP is the first step in RUL prediction; following its detection, several

algorithmic approaches can be utilized to forecast the degraded state of the component. Table 3 summarizes the results across all datasets. The proposed tool consistently identifies the onset of degradation earlier or at a comparable point to previous works, demonstrating its suitability for real-world predictive maintenance scenarios.

These results indicate that the proposed tool performs competitively, providing earlier detection in several cases while requiring substantially less data and no manual feature extraction. Moreover, it is consistent while predicting under different datasets, as most of the

Table 1
Summary of fault detection, diagnosis and classification results for well-known datasets.

Dataset	Sample	Detected faults	Degradation stage	Diagnosis
IMS-1	#1850 - #1880	Inner-race - Cage	Early stage	1X BPFI - 1X FTF
	#2135 - #2145	Inner-race - Cage	Medium stage	1X BPFI - 2X, 4X FTF
	#2150 - #2156	Outer-race - Inner-race - Cage	Last stage	1X BPFO - 1X BPFI - 3X FTF
IMS-2	#530 - #550	Outer-race - Cage	Early stage	1X, 2X BPFO - 1X, 5X, 6X FTF
	#865 - #875	Outer-race	Medium stage	1X, 2X, 3X, 4X, 5X BPFO
	#970 - #980	Outer-race	Last stage	1X, 2X, 3X, 4X BPFO
IMS-3	#5960 - #5990	Cage	Early stage	2X, 3X, 4X, 5X FTF
	#6170 - #6190	Outer-race	Medium stage	1X, 2X, 3X, 4X BPFO
	#6314 - #6324	Outer-race - Cage	Last stage	1X, 2X, 3X, 4X BPFO - 4X FTF
XJTU 2-1	#450 - #457	Inner-race - Ball - Cage	Medium stage	1X, 2X BPFI - 1X BSF - 5X FTF
	#455 - #460	Inner-race - Ball - Cage	Medium stage	1X, 2X BPFI - 1X BSF - 5X FTF
	#460 - #470	Inner-race - Ball - Cage	Last stage	1X BPFI - 1X BSF - 2X, 5X FTF
XJTU 2-3	#301 - #311	Ball - Cage	Early stage	1X BSF - 1X, 2X, 4X, 5X, 6X FTF
	#410 - #420	Outer-race - Inner-race - Cage	Medium stage	1X BPFO - 1X BPFI - 1X, 2X, 4X FTF
	#525 - #532	Outer-race - Inner-race - Ball - Cage	Medium stage	1X, 5X, 6X BPFO - 2X BPFI - 1X, 3X BSF - 5X FTF
XJTU 3-1	#2340 - #2360	Outer-race - Ball - Cage	Early stage	1X, 2X BPFO - 1X, 3X BSF - 3X FTF
	#2440 - #2460	Outer-race - Inner-race - Ball - Cage	Medium stage	2X, 3X BPFO - 1X BPFI - 1X, 3X BSF - 1X, 3X, 5X FTF
	#2520 - #2540	Outer-race - Inner-race - Ball - Cage	Last stage	1X, 2X, 3X, 5X BPFO - 5X BPFI - 2X, 3X BSF - 6X FTF
XJTU 3-4	#1410 - #1425	Inner-race - Ball - Cage	Early stage	1X BPFI - 1X BSF - 5X FTF
	#1445 - #1460	Outer-race - Inner-race - Ball - Cage	Medium stage	1X BPFO - 1X BPFI - 1X, 3X BSF - 2X, 5X FTF
	#1495 - #1510	Outer-race - Inner-race - Ball - Cage	Last stage	1X BPFO - 1X, 2X BPFI - 1X BSF - 2X, 5X FTF

Insert relevant information

Shaft Freq.

33.33

Hz

Sampling Freq.

20480.0

Hz

Load

26.7

Kn

Bearing Type

Rexnord

BPFO

236.0

Hz

BPFI

297.0

Hz

BSF

278.0

Hz

FTF

15.0

Hz

Fig. 6. BEARING-FDD: adding new dataset information.

Load Data

Choose File

No file chosen

Upload

Fig. 7. BEARING-FDD: loading a new dataset.

Table 2
Quantitative evaluation metrics of MS2AE.

Dataset	Train Samples	Test Samples	Accuracy	Precision	Recall	F1-score
IMS-1	300	1856	0.943	0.947	0.941	0.944
IMS-2	300	684	0.951	0.957	0.946	0.951
IMS-3	300	6024	0.963	0.968	0.96	0.964
XJTU 2-1	300	191	0.942	0.963	0.931	0.947
XJTU 2-3	300	233	0.947	0.934	0.938	0.936
XJTU 3-1	300	2238	0.961	0.965	0.958	0.961
XJTU 3-4	300	1215	0.955	0.968	0.945	0.956

Table 3
Comparison of First Faulty Point.

Dataset	Proposed Solution	[13]	[14]	[15]	[16]	[17]	[18]	[19]
IMS-1	1857	–	1917	1833	1971	–	–	–
IMS-2	536	934	632	533	530	–	–	–
IMS-3	5967	–	–	5952	–	–	–	–
XJTU 2-1	451	–	458	–	–	455	452	450
XJTU 2-3	301	–	–	–	–	327	302	313
XJTU 3-1	2347	2348	2407	–	–	2344	2346	2376
XJTU 3-4	1416	1418	1446	–	–	1418	1417	1430

related works analyzed only make predictions in one dataset (IMS or XJTU), where the rotating machinery analyzed is always placed in the same testbed. Only Guo et al. [14] analyze both datasets under the same modeling framework. However, the model proposed in this paper achieves superior performance in FFP identification across all tests.

Table 4
Three most correlated engineering features with HI.

Dataset	Three most correlated features			Present fault
IMS-1	RMS (0.97)	Shape factor (0.82)	BPFI frequency (0.80)	Inner-race bearing fault
IMS-2	RMS (0.95)	BPFO frequency (0.88)	FTF (0.83)	Outer-race bearing fault
IMS-3	RMS (0.97)	BPFO frequency (0.85)	Fundamental frequency (0.84)	Outer-race bearing fault
XJTU 2-1	RMS (0.95)	BPFI frequency (0.88)	Shape factor (0.72)	Inner-race bearing fault
XJTU 2-3	RMS (0.98)	BPFI frequency (0.76)	Fundamental frequency (0.76)	Cage bearing fault
XJTU 3-1	RMS (0.92)	BPFO frequency (0.55)	Skewness (0.53)	Outer-race bearing fault
XJTU 3-4	RMS (0.98)	Shape factor (0.93)	BPFI (0.92)	Inner-race bearing fault

To support the credibility of the explainability module, the correlations automatically produced between the Health Indicator and classical engineering features were reviewed against previously published studies in the bearing-fault diagnosis literature.

Table 4 lists the three features that show the highest correlation with the HI for each dataset. The selected features and their associations with specific fault types are supported by classical vibration analysis principles.

Across the analyzed datasets, clear patterns emerge in the relationship between the Health Indicator and frequency-domain features associated with specific bearing fault types. In all datasets where an outer-race fault is present (IMS-2, IMS-3, XJTU 3-1), the BPFO harmonic amplitude consistently appears among the most correlated features. This behavior aligns with classical vibration theory, where outer-race faults generate strong spectral components at the BPFO and its harmonics [20–23].

Conversely, in datasets containing inner-race faults (IMS-1, XJTU 2-1, XJTU 3-4), the BPFI harmonic amplitude systematically shows high correlation with the HI. This is also consistent with established diagnostic principles, since inner-race defects typically excite dominant spectral components at the BPFI [20–23]. The agreement between the automatically learned HI and these well-known characteristic frequencies supports the physical consistency of the explainability module.

Another relevant observation is the consistent role of the shape factor in datasets containing inner-race faults. For IMS-1, XJTU 2-1, and XJTU 3-4, the shape factor appears among the top correlated features. The shape factor is a classical time-domain indicator sensitive to waveform deformation and impulsive components, which are typically intensified when defects occur on the inner race due to its rotational motion relative to the load zone [24–26]. The strong relationship between the HI and the shape factor in these cases reinforces that the HI captures not only spectral fault signatures (BPFI) but also their corresponding time-domain impulsiveness characteristics.

Finally, the RMS shows extremely high correlation with the HI across all datasets, with values between 0.92 and 0.98. This behavior is expected and well documented in the literature, as the RMS is one of the most widely used global indicators of vibrational energy and tends to increase consistently when bearing degradation progresses [27–29]. The fact that the RMS is the most correlated feature in all experiments suggests that the HI captures the overall growth in vibrational amplitude associated with general fault severity, while still differentiating specific fault types through more specialized features such as BPFO, BPFI, and shape factor. This reinforces the interpretability of the HI as both a global health metric and a fault-sensitive indicator.

5. Impact

Detecting, classifying, and diagnosing bearing faults in rotating machinery have become critical activities in modern industrial maintenance systems. While practical software tools are highly desirable to support these tasks, very few are mentioned in the literature. BearTools [30] is a MATLAB-based interactive tool designed for studying bearing fault diagnosis through vibration analysis and envelope detection, but it is primarily intended for educational use. SIDMA [31] is a tool for detecting and diagnosing bearing faults in induction motors

using Motor Current Signature Analysis (MCSA) and machine learning algorithms. The tool features an API that allows users to upload motor current measurements and obtain bearing condition assessments; however, it is not publicly available.

The proposed BEARING-FDD software tool is designed based on the aforementioned work in [12], implementing all the functionality in a user-friendly web interface. It features an API and a web front-end. The API enables seamless integration with various industrial systems and applications, supporting real-time monitoring and diagnostics across different platforms. This significantly enhances maintenance strategies and operational efficiency by enabling automated fault detection and predictive maintenance. The web front-end demonstrates the API's capabilities, providing practitioners with a practical interface to test the API using their own datasets, and allowing them to download both explainability and interpretability reports.

The BEARING-FDD tool is publicly available and suitable for use in environments with limited data from the monitored RM. It only requires human input to provide healthy samples, samples to be analyzed, and some key information about the RM being analyzed, such as shaft frequency, bearing-related frequencies, and sample frequency. The diagnosed faults are clearly explained without the need for machine learning expertise, eliminating the necessity for highly specialized personnel.

Because the MS2AE model is trained exclusively on healthy data, it avoids the need for balanced datasets, which is especially valuable in real settings where faulty data is scarce, inconsistent, or unavailable. The recursive use of DTW to merge healthy samples effectively isolates stable patterns and removes noise, producing a clean baseline signal that significantly enhances fault-signature extraction when comparing against faulty samples. Additionally, the kurtogram-driven filtering and harmonic analysis provide an interpretable link between the extracted fault components and physically meaningful bearing frequencies. However, the method inherently relies on the availability of sufficiently high-frequency vibration data: if the sampling frequency falls below 2 kHz, the Nyquist limit restricts the FFT to below 1 kHz, making many bearing-fault harmonics (BPFO, BPFI, BSF, FTF and their multiples) inaccessible. As a result, low sampling rates can mask key spectral indicators and degrade diagnostic performance, highlighting the importance of using vibration acquisition systems capable of capturing the higher-frequency impulsive content associated with early-stage bearing faults.

6. Conclusions

A novel web tool for bearing fault detection, diagnosis and classification providing explainable and interpretable results has been presented. The software allows users to use preloaded datasets or upload their own datasets. The MS2AE model addresses two major challenges in fault analysis. First, it can be effectively applied with limited and imbalanced datasets, as the model requires only healthy data from the monitored machinery for training. Second, no manual feature extraction is necessary, as the model processes raw vibration data without any preprocessing. The model consolidates all features into a health indicator value, which is then used to identify the initial point of fault.

The results obtained from relevant bearing fault datasets demonstrate that the proposed fault detection, diagnosis and classification



Fig. 8. BEARING-FDD: saved dataset.

tool performs effectively under various operating conditions and motor loads. The system operates autonomously, requiring human intervention only to provide healthy samples, unlabeled samples, and essential information about the rotating machinery being analyzed, such as shaft frequency, bearing-related frequencies, and sampling frequency. This autonomy makes the tool well-suited for real-world applications, even when limited data is available from the monitored machinery. Additionally, no specialized expertise is required, as the tool clearly explains the diagnosed faults without necessitating knowledge of machine learning. Furthermore, the tool can be easily applied to any rotating machinery due to its flexibility. It can also be adapted to support the diagnosis of other common faults, such as unbalance, misalignment, and gear faults.

Summarizing, the primary contributions of the developed software are:

- Operates effectively with limited data, as only healthy data is required for training the MS2AE model.
- Eliminates the need for manual feature extraction, as the MS2AE model compresses the raw vibration data into a single health indicator value.
- Demonstrates robustness under varying operating conditions and motor loads.
- Provides explainability of the health indicator obtained using well-known engineering features.
- Offers interpretability, making the system accessible to non-specialized staff, as diagnosed faults are clearly described.

Future work will focus on the applicability of this software to other scenarios, such as gears or combustion engines. Moreover, the integration of additional ML models for RUL prediction, as well as the software's effectiveness in diagnosing other common faults, such as unbalance, misalignment, and gear faults, will be explored.

CRedit authorship contribution statement

L. Magadán: Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Data curation, Conceptualization. **C. Ruiz-Cárcel:** Writing – review & editing, Supervision, Resources, Methodology, Investigation, Formal analysis, Conceptualization. **J.C. Granda:** Writing – review & editing, Visualization, Supervision, Software, Methodology, Investigation. **F.J. Suárez:** Writing –

review & editing, Supervision, Project administration, Methodology, Funding acquisition, Formal analysis. **A. Menéndez-González:** Writing – review & editing, Validation, Software, Methodology, Data curation. **A. Starr:** Writing – review & editing, Supervision, Formal analysis, Conceptualization.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Francisco Jose Suarez reports financial support was provided by Spanish National Plan of Research, Development, and Innovation, project EDNA (PID2021-124383OB-I00). Luis Magadan reports financial support was provided by Severo Ochoa Program, Principado de Asturias (PA-22-BP21-120). If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This research was partially funded by the Spanish National Plan of Research, Development, and Innovation under project EDNA (PID2021-124383OB-I00), the European Union, the University of Oviedo and the University of Cranfield. L. Magadán is supported by the Severo Ochoa program (PA-22-BP21-120).

References

- [1] A.G. Nath, S.S. Udmale, D. Raghuvanshi, S.K. Singh, Improved structural rotor fault diagnosis using multi-sensor fuzzy recurrence plots and classifier fusion, *IEEE Sensors J.* 21 (19) (2021) 21705–21717.
- [2] H. Hashemian, Wireless sensors for predictive maintenance of rotating equipment in research reactors, *Ann. Nucl. Energy* 38 (2–3) (2011) 665–680.
- [3] T. Williams, X. Ribadeneira, S. Billington, T. Kurfess, Rolling element bearing diagnostics in run-to-failure lifetime testing, *Mech. Syst. Signal Process.* 15 (5) (2001) 979–993.
- [4] S. Zhang, S. Zhang, B. Wang, T.G. Habetler, Deep learning algorithms for bearing fault diagnostics—a comprehensive review, *IEEE Access* 8 (2020) 29857–29881.
- [5] Y. Xu, K. Feng, X. Yan, R. Yan, Q. Ni, B. Sun, Z. Lei, Y. Zhang, Z. Liu, Cfcnn: A novel convolutional fusion framework for collaborative fault identification of rotating machinery, *Inf. Fusion* 95 (2023) 1–16.

- [6] Z. Meng, W. Cao, D. Sun, Q. Li, W. Ma, F. Fan, Research on fault diagnosis method of ms-cnn rolling bearing based on local central moment discrepancy, *Adv. Eng. Inform.* 54 (2022) 101797.
- [7] Z. Meng, Q. Li, D. Sun, W. Cao, F. Fan, An intelligent fault diagnosis method of small sample bearing based on improved auxiliary classification generative adversarial network, *IEEE Sensors J.* 22 (20) (2022) 19543–19555.
- [8] S. Liu, H. Jiang, Z. Wu, Y. Liu, K. Zhu, Machine fault diagnosis with small sample based on variational information constrained generative adversarial network, *Adv. Eng. Inform.* 54 (2022) 101762.
- [9] R. Marcinkevičius, J.E. Vogt, Interpretability and explainability: A machine learning zoo mini-tour, 2020, arXiv preprint [arXiv:2012.01805](https://arxiv.org/abs/2012.01805).
- [10] N. Costa, L. Sánchez, Variational encoding approach for interpretable assessment of remaining useful life estimation, *Reliab. Eng. Syst. Saf.* 222 (2022) 108353.
- [11] S. Li, T. Li, C. Sun, R. Yan, X. Chen, Multilayer grad-cam: An effective tool towards explainable deep neural networks for intelligent fault diagnosis, *J. Manuf. Syst.* 69 (2023) 20–30.
- [12] L. Magadán, C. Ruiz-Cárcel, J. Granda, F. Suárez, A. Starr, Explainable and interpretable bearing fault classification and diagnosis under limited data, *Adv. Eng. Inform.* 62 (2024) 102909.
- [13] H. Wang, X. Zhang, X. Guo, T. Lin, L. Song, Remaining useful life prediction of bearings based on multiple-feature fusion health indicator and weighted temporal convolution network, *Meas. Sci. Technol.* 33 (10) (2022) 104003.
- [14] W. Guo, X. Li, X. Wan, A novel approach to bearing prognostics based on impulse-driven measures, improved morphological filter and practical health indicator construction, *Reliab. Eng. Syst. Saf.* 238 (2023) 109451.
- [15] T. Yan, D. Wang, J.-Z. Kong, T. Xia, Z. Peng, L. Xi, Definition of signal-to-noise ratio of health indicators and its analytic optimization for machine performance degradation assessment, *IEEE Trans. Instrum. Meas.* 70 (2021) 1–16.
- [16] Y. Zhou, B. Chen, X. Huang, R. Jiang, X. Wang, X. Xie, Q. Xu, W. Hu, W. Zhang, H.-Z. Song, Study on strain effects on the electronic and optical properties of hexagonal boron phosphide monolayer, in: *International Conference on Optoelectronic Information and Functional Materials (OIFM 2023)*, vol. 12781, SPIE, 2023, pp. 198–204.
- [17] H. Wei, Q. Zhang, Y. Gu, Remaining useful life prediction of bearings based on self-attention mechanism, multi-scale dilated causal convolution, and temporal convolution network, *Meas. Sci. Technol.* 34 (4) (2023) 045107.
- [18] Y. Deng, G.-C. Sun, L. Han, Y. Wang, Y. Zhang, M. Xing, 2-d wavenumber domain autofocusing for high-resolution highly squinted sar imaging based on equivalent broadside model, *IEEE Trans. Geosci. Remote Sens.* 61 (2023) 1–15.
- [19] Z. Li, P. Xu, X.-B. Wang, Online anomaly detection and remaining useful life prediction of rotating machinery based on cumulative summation features, *Meas. Control.* 56 (3–4) (2023) 615–629.
- [20] S. Khot, S. Gurav, Investigation of effect of surface failures on inner and outer race of bearing on vibration spectrum, in: *2015 International Conference on Nascent Technologies in the Engineering Field, ICNTE, IEEE*, 2015, pp. 1–4.
- [21] O. Das, D.B. Das, D. Birant, Machine learning for fault analysis in rotating machinery: A comprehensive review, *Heliyon* 9 (6) (2023).
- [22] P. Nivesrangsan, D. Jantarajirokul, Bearing fault monitoring by comparison with main bearing frequency components using vibration signal, in: *2018 5th International Conference on Business and Industrial Research, ICBIR, IEEE*, 2018, pp. 292–296.
- [23] O. Duque, M. Perez, D. Morinigo, Detection of bearing faults in cage induction motors fed by frequency converter using spectral analysis of line current, in: *IEEE International Conference on Electric Machines and Drives*, vol. 2005, IEEE, 2005, pp. 17–22.
- [24] S. Bhatti, A.A. Shaikh, A. Mansoor, M. Hussain, New approaches of stochastic models to examine the vibration features in roller bearings, *Appl. Sci.* 14 (4) (2024) 1616.
- [25] S. Bhatti, A.A. Shaikh, A. Mansoor, M. Hussain, Applying stochastic models to analyze vibration features of roller bearings, 2023.
- [26] L. Zhan, F. Ma, J. Zhang, C. Li, Z. Li, T. Wang, Fault feature extraction and diagnosis of rolling bearings based on enhanced complementary empirical mode decomposition with adaptive noise and statistical time-domain features, *Sensors* 19 (18) (2019) 4047.
- [27] A. Nabhan, M. Nouby, A. Sami, M. Mousa, Multiple defects detection in outer race of gearbox ball bearing using time domain statistical parameters, *Int. J. Veh. Struct. Syst. (IJVSS)* 8 (3) (2016).
- [28] ISO, Mechanical vibration — measurement and evaluation of machine vibration (ISO 20816), 2016, <https://www.iso.org/standard/63180.html>. (Last accessed 12 December 2025).
- [29] P.H. Jain, S.P. Bhosle, Analysis of vibration signals caused by ball bearing defects using time-domain statistical indicators, *Int. J. Adv. Technol. Eng. Explor.* 9 (90) (2022) 700.
- [30] M. Rebiai, B. Bengherbia, M.O. Zmirli, S. Lachenani, An interactive educational tool for studying the diagnosis of bearing faults using vibration analysis, in: *2022 2nd International Conference on Advanced Electrical Engineering, ICAEE, IEEE*, 2022, pp. 1–6.
- [31] G. Gil-Prieto, J. Antonino-Davui, D. Tarín-Caballero, P. Mullor-Ruiz, A. Quijano-López, Automated tool for bearing fault diagnosis in induction motors, based on mcsa technique and machine learning algorithm, in: *2022 CIGRE Paris Session*, 2022.



DSPACE

<https://dspace.org/>

BEARING-FDD: an early detection and diagnosis tool for bearing faults in rotating machinery

Magadán, Luis; Ruiz-Cárcel, Cristobal; Granda, Juan C.; Suárez, Francisco J.; Menéndez-González, A ...

2026-04

Elsevier

<https://doi.org/10.1016/j.simpa.2025.100810>

Magadán L, Ruiz-Cárcel C, Granda JC, et al., (2026) BEARING-FDD: an early detection and diagnosis tool for bearing faults in rotating machinery. Software Impacts, Volume 27, April 2026, Article number 100810