Sound & Vibration 2026, 60(6), 4916.
https://doi.org/10.59400/ sv4916

Conceptual Paper

Explainable Machine Learning for Vibration-Based Fault Diagnosis and Predictive-Maintenance Decision Support in Rotating Machinery

Susmita Suresh Pagare[1], Tarun Madan Kanade[2, *], Nitin Raut[3], Jonathan Sudhir Joseph[4], P. William[5, 6], Koyel Roy[7], Pallavi Yarde[8], Yogeesh N[9], Tushar Savale[10]

[1]Research Scholar, Symbiosis Institute of Operations Management, Symbiosis International (Deemed University), India
[2]Faculty of Management, Symbiosis Institute of Operations Management, Symbiosis International (Deemed University), India
[3] Faculty of Engineering, Sohar University, Sohar, Oman
[4] Faculty of Management, Thakur College of Engineering and Technology, Mumbai, India
[5] School of Computer Science and Technology, Karunya Institute of Technology and Sciences (Deemed University), Coimbatore, India
[6] Centre for Research and Consultancy, UNITAR International University, Selangor, Malaysia
[7] Faculty of Law, Symbiosis Law School Hyderabad, Symbiosis International (Deemed University), India
[8] Dr. D. Y. Patil School of Science & Technology, Dr. D. Y. Patil Vidyapeeth, Pimpri, Pune, India
[9] Mathematics & Natural Sciences, Gulf University for Science & Technology, Kuwait
[10] Faculty of Management, Balaji Institute of Modern Management, Sri Balaji University, Pune, India

[*]**Corresponding author:** Tarun Madan Kanade, tarun.kanade@siom.in

CITATION

Pagare SS, Kanade TM, Raut N, et al . Explainable Machine Learning for Vibration-Based Fault Diagnosis and Predictive-Maintenance Decision Support in Rotating Machinery. Sound & Vibration. 2026; 60(6): 4916. https://doi.org/10.59400/ sv4916

COPYRIGHT

Copyright © 2026 by author(s).
Sound & Vibration is published by Piscomed Publishing Pte. Ltd. This work is licensed under the Creative Commons Attribution (CC BY) license.
https://creativecommons.org/licenses/by/4.0/

Abstract: This conceptual study proposes an explainable machine-learning framework for vibration-based fault diagnosis and predictive-maintenance decision support in rotating machinery. The architecture integrates vibration preprocessing, time- and frequency-domain feature engineering, machine-learning and deep-learning classification, SHAP-based interpretation, and maintenance decision support. The proposed framework is designed for vibration datasets comprising 48 constant-speed and 60 variable-speed processes, recorded using two triaxial accelerometers at a sampling frequency of 20 kHz and rotational speeds of up to 5,000 rpm. The rolling-element bearing dataset includes three conditions: NoFault (healthy), Ball Pass Frequency Outer (BPFO), and Ball Spin Frequency (BSF). For future empirical evaluation, the proposed framework considers 64 training, 22 validation, and 22 testing processes. Time- and frequency-domain features, including RMS, standard deviation, skewness, kurtosis, peak, crest factor, impulse factor, and shape factor, are considered for fault characterization. Previous studies have reported classification accuracies exceeding 98.5% for SVM and k-NN models, 87.2% using raw acceleration data, 93.7% for CNN-BiLSTM under variable-speed conditions, 99.8% for ResNet under constant-speed conditions, and 99.3% for a deep-learning ensemble. These values are presented as reference benchmarks rather than experimental results of the present study. The proposed framework incorporates SHAP-based explanations to interpret the contributions of vibration features, including skewness, form factor, and SMRA. The effectiveness of the framework in supporting inspection prioritization and maintenance decisions requires empirical validation. The framework prioritizes early fault identification and maintenance support over RUL estimation because degradation and maintenance-event data are unavailable.

Keywords: Explainable Artificial Intelligence; Vibration-Based Fault Diagnosis; Predictive Maintenance; Rotating Machinery; Machine Learning; SHAP; Condition Monitoring

1. Introduction

1.1. Background

Mechanics, motors, couplings, gears, shafts, and bearings all contribute to the dependability of modern industrial systems. Machinery failures reduce production efficiency, increase maintenance requirements, and affect equipment availability. To enhance maintenance planning, approaches can identify unusual machine behavior and support predictive maintenance [1,2].

Vibration monitoring is necessary for evaluating the health of rotating equipment. The amplitude, frequency, and statistics of vibration are affected by misalignment, friction, bearings, and imbalance. Vibration fluctuations may indicate mechanical faults during machine operation [3]. Vibration-based diagnostics is used in data-driven condition monitoring.

Advances in sensing and signal processing have facilitated the use of ML and DL in machinery fault diagnosis. ML methods typically rely on manually extracted features, whereas DL methods can learn patterns directly from vibration signals. Previous studies have reported a classification accuracy of 99.3% using a CNN-BiLSTM ensemble for bearing-fault detection [2] and 99.4% using deep-learning techniques based on vibration images [4]. Direct comparisons are not possible due to differences in datasets, operational circumstances, and models.

Classification accuracy alone cannot explain why a model produces a particular prediction. XAI helps interpret model behavior, while SHAP can explain the contributions of input features to predictions at both global and individual levels [3]. These explanations reflect model attribution rather than physical causation or mechanical fault mechanisms.

This conceptual framework integrates ML/DL-based fault diagnosis, vibration analysis, SHAP explanations, and engineering interpretation to support maintenance decision-making. It provides a basis for linking diagnostic predictions with engineering assessment and inspection prioritization, while prioritizing early fault identification and inspection support over RUL estimation.

1.2. Research Problem

Fault diagnosis using vibration still faces three obstacles.

First, diagnostic performance and model interpretability are two distinct aspects of fault diagnosis. Although ML models can classify fault conditions, their predictions may not clearly explain which features contribute to each classification. This may make it difficult for engineers to evaluate and apply model outputs.

Second, changes in rotational speed, load, environmental conditions, and measurement settings may affect vibration characteristics and diagnostic performance. Models developed under specific operating conditions may not generalize well to other conditions. Therefore, operating conditions should be considered when designing and evaluating vibration-based fault-diagnosis approaches [2,5].

Third, fault classification alone does not determine maintenance priorities. Maintenance decisions also depend on operating conditions, equipment criticality, fault severity, and engineering knowledge. Therefore, model predictions and explanations should be considered together with engineering assessment when planning inspections and maintenance activities.

This study proposes a conceptual framework that integrates vibration-based feature analysis, machine-learning and deep-learning-based fault diagnosis, SHAP-based explanations, and engineering interpretation to support maintenance-oriented decision-making. Since independent experimental validation and industrial deployment have not been conducted, the framework

focuses on fault identification and maintenance decision support rather than remaining useful life (RUL) prediction.

1.3. Research Gap

Four gaps exist in the literature.

First, existing studies have examined fault diagnosis and explainable AI for rotating machinery, but the integration of model predictions, feature explanations, engineering interpretation, and maintenance decision support remains limited [3,6]. This study extends beyond feature attribution by connecting diagnostic explanations with engineering interpretation and maintenance decisions.

Second, variations in rotational speed and load can affect vibration characteristics and diagnostic performance. Comparisons across different models and studies are limited by differences in datasets, experimental designs, and evaluation procedures [2,5]. This study considers operating-condition variations but does not aim to validate model robustness.

Third, results from the lab may not be indicative of real-world business practices. Sensor noise, loads, environment, and interactions may all have an impact on performance. As discussed in [1,7], laboratory-based results alone are insufficient to establish the readiness of a diagnostic approach for industrial application.

Fourth, model explanations are not always directly connected to maintenance decision-making. RMS reflects the overall vibration level, while kurtosis characterizes the peakedness of the amplitude distribution and may indicate impulsive vibration patterns. SHAP values explain model predictions rather than directly indicating fault severity or physical causes.

To address these gaps, this study proposes a conceptual and methodological framework integrating vibration data preparation, feature engineering, ML/DL-based fault diagnosis, SHAP-based explanation, engineering interpretation, and maintenance decision support. Its conceptual contribution lies in linking diagnostic predictions and feature explanations with engineering assessment and maintenance decision-making. The framework distinguishes between prediction, explanation, and maintenance actions. As independent model training, performance evaluation, and SHAP analysis have not been conducted, the proposed framework remains conceptual and requires empirical validation.

1.4. Research Objectives

In order to aid in predictive maintenance decision-making and vibration-based problem detection, this research intends to provide an integrated framework.

RO1: To examine how vibration-based features can support fault identification in rotating machinery.

RO2: To describe how ML and DL methods can be incorporated into vibration-based fault diagnosis.

RO3: To identify appropriate performance measures for future evaluation of fault-diagnosis models.

RO4: To describe how SHAP can be used to explain the contributions of vibration features to model predictions.

RO5: To explain how model-attributed features can be interpreted using established engineering knowledge of rotating-machinery faults.

RO6: To develop a framework connecting explainable fault diagnosis with inspection prioritisation and maintenance decision support.

1.5. Research Questions

RQ1: How can vibration-based features differentiate fault types in rotating

machinery?

RQ2: How can ML and DL methods be incorporated into vibration-based fault diagnosis?

RQ3: Which vibration features and performance measures are relevant to fault-diagnosis evaluation?

RQ4: How can SHAP be applied to explain predictions from vibration-based fault-diagnosis models?

RQ5: How can model explanations be interpreted using established engineering knowledge?

RQ6: How can explainable fault diagnosis support inspection prioritisation and maintenance decision-making?

1.6. Contributions of the Study

This paper makes four conceptual and methodological advancements.

First, this study proposes an integrated conceptual framework combining vibration preprocessing, feature engineering, ML/DL-based diagnosis, explainability, and engineering interpretation. Beyond classification and feature attribution, the framework connects diagnostic outputs with engineering interpretation to support maintenance-oriented decision-making.

Second, the framework incorporates SHAP to explain model predictions at both global and individual levels. These explanations represent model feature attributions rather than direct evidence of physical fault causation.

Third, engineering interpretation links RMS, kurtosis, and characteristic frequency to vibration-engineering data. Mechanical failure should be interpreted using fault characteristics and operating circumstances, not feature rankings.

Fourth, maintenance-oriented decision support integrates engineering review, inspection prioritisation, and planning based on diagnostic predictions and explanations. It supports further monitoring and inspection without assuming that SHAP values directly determine maintenance decisions.

This study contributes a conceptual framework integrating vibration-based fault diagnosis, explainability, and engineering decision support. Its practical effectiveness and applicability require further empirical validation, including model evaluation and comparison with existing methodologies.

2. Literature Review And Proposed Research Framework

2.1. Vibration-Based Fault Diagnosis

Mechanical problems modify vibration amplitude, frequency, and statistical features, hence vibration-based fault diagnosis is used to monitor rotating machinery. BPFO and BSF may create distinct vibration patterns. Their identification depends on rotational speed, bearing shape, sensor placement, signal quality, and operation. Time-domain metrics like RMS, variance, skewness, and kurtosis characterize signal amplitude and distribution. Frequency-domain analysis identifies dominant spectral components and fault-related frequencies [8,9].

The proposed framework is designed with reference to the rotating-machinery bearing dataset described in [2], which includes normal and faulty bearings operating under constant-speed and variable-speed conditions.

Performance figures reported in previous studies should be treated as reference benchmarks rather than experimental results of the present study. The dataset provides a reference for describing the proposed vibration-based fault-diagnosis framework.

2.2. Machine Learning and Deep Learning for Fault Diagnosis

Machine-learning methods can identify bearing conditions using

vibration-based features. SVM, k-NN, and Random Forest are commonly used with structured feature data, whereas CNNs, RNNs, and residual networks can learn patterns from vibration signals or their representations. Data volume, signal representation, processing needs, and deployment environment influence suitability [1,10,11].

Study [3] reported accurate fault-diagnosis results using conventional machine-learning methods. Some have used convolutional networks and transfer learning to diagnose bearings [12,13]. While data-driven diagnosis shows promise, it is difficult to compare accuracy due to differences in datasets, fault classifications, train-test methodologies, and operational circumstances.

The proposed framework considers various machine-learning and deep-learning models, including SVM, k-NN, Random Forest, CNN, LSTM-based networks, ResNet, and CNN-BiLSTM ensembles. Model training, hyperparameter selection, and performance evaluation using accuracy, precision, recall, and F1-score, supported by confusion matrices, are proposed as steps for future empirical validation and have not been conducted in the present study.

2.3. Diagnosis under Variable Operating Conditions

When the vibration patterns change due to variations in rotational speed and load, models trained under limited laboratory conditions may no longer be reliable. Rather than learning regime-specific properties, classifiers may learn speed-dependent ones [5,12,14]. Several methods, including transfer learning, domain adaptation, and signal resampling, deal with distribution shifts.

These methods should be considered in relation to the specific operating conditions and evaluation design. The proposed framework considers both constant- and variable-speed operating scenarios. However, cross-condition evaluation is required to establish diagnostic robustness, which remains subject to future empirical validation.

2.4. Explainable Machine Learning and SHAP

High classification accuracy does not necessarily explain why a model assigns a particular fault label. Explainable artificial intelligence (XAI) can help connect model outputs with features that engineers recognise. Shapley Additive Explanations (SHAP) attributes a model prediction to individual input features relative to a reference distribution. Global explanations summarise feature contributions across observations, whereas local explanations describe contributions to an individual prediction.

Research has explored unsupervised fault detection in rotating machinery [6]. Studies have also demonstrated the relevance of feature-contribution analysis for industrial condition monitoring [3]. Recent research has applied explainable AI to bearing fault diagnosis [14] and physics-informed self-supervised learning to rotating-machinery diagnosis [15], while review studies have examined emerging approaches in this field [16] and transfer-learning-based approaches have been explored for explainable rotating-machinery fault diagnosis [17]. A comparison of these studies is presented in **Table 1**.

Table 1. Comparison of selected studies relevant to explainable vibration-based fault diagnosis

Study	Research focus	Relevance to the proposed framework
[14]	Explainable AI-based bearing fault diagnosis;	Demonstrates the application of explainable AI to bearing fault diagnosis
[6]	Unsupervised fault detection and diagnosis	Extends consideration beyond conventional supervised classification
[17]	Transfer learning-based fault diagnosis using augmented synthetic data;	Highlights the use of transfer learning in explainable rotating-machinery fault diagnosis

Continued table:

Study	Research focus	Relevance to the proposed framework
[3]	Feature contributions in machine fault diagnosis	Supports interpreting features in condition monitoring
[16]	Review of explainable AI in rotating-machinery fault diagnosis;	Provides an overview of existing methods and emerging research directions
[18]	Review of vibration signal analysis for intelligent rotating-machinery diagnosis and prognosis;	Provides background on vibration-based diagnosis and prognosis in rotating machinery.

Source: Authors' synthesis of the cited studies. The relevance column describes their relationship to the proposed framework; it does not claim that the listed limitations were explicitly identified by the original authors.

The proposed framework places SHAP between fault classification and technical interpretation. In the present study, global and local SHAP plots are schematic illustrations rather than outputs generated from a fitted model, and should not be interpreted as empirical results. SHAP identifies feature contributions to model predictions rather than establishing causal relationships or the physical causes of machinery faults.

2.5. Predictive Maintenance and Decision Support

Current operational data can support predictive maintenance through condition monitoring, inspection, and maintenance planning. Sorting defects into categories may identify unusual events and concerns that require additional examination. Different diagnostic and maintenance solutions exist. Predicting a fault class doesn't indicate damage is guaranteed, action is needed immediately, or the optimum maintenance plan is set.

Vibration-based predictive-maintenance research increasingly combines sensor signals with data-driven methods [1,9]. When taken into account with operational context and engineering judgment, the predictions and explanations provided by the models might help with the prioritization of inspections under the proposed framework. An example of this would be a targeted bearing examination that is triggered by an aberrant classification and a validated fault-related spectral pattern. This is a proposed decision pathway, not a validated decision rule or an observed outcome of the present study.

Remaining useful life (RUL) estimation is outside the framework's current scope. Reliable RUL modelling generally requires suitable degradation trajectories, failure-time information or other longitudinal evidence. A fault classifier alone cannot establish the remaining operating life of a component.

2.6. Research Gaps

Four research gaps have been identified in the existing literature. Since diagnostic performance and explanation are usually handled independently, model predictions need to be connected to interpretable feature contributions. Second, variations in rotational speed and operating conditions may affect model generalizability and therefore require explicit consideration. Third, sensor noise, improper installation, and other practical factors may limit the applicability of controlled test-bench findings to industrial environments. Fourth, the connection between diagnostic outputs and maintenance decision-making remains insufficiently addressed. These research gaps motivate the proposed integrated conceptual framework, whose conceptual contribution lies in connecting fault diagnosis, explainability, engineering interpretation, and maintenance decision support. Its effectiveness in addressing these gaps remains subject to future empirical validation.

2.7. Proposed Research Framework

The proposed framework integrates an existing vibration dataset, preprocessing, feature engineering, machine-learning and deep-learning-based

fault diagnosis, SHAP-based explanation, engineering interpretation, and maintenance decision support. The dataset serves as a reference for the proposed framework rather than as the basis for independent experimental validation.

Machine-learning and deep-learning models are proposed for bearing fault classification. Global and local SHAP explanations are proposed to illustrate feature contributions to model predictions, supporting engineering interpretation of vibration patterns and potential bearing faults. These interpretations may assist engineers in prioritizing inspections and planning maintenance activities.

In a future implementation, feedback mechanisms may include model updates, data integration, and performance monitoring. Signal waveforms, frequency spectra, and SHAP visualizations are included as illustrative examples to demonstrate the proposed framework. They do not represent actual experimental results or independently calculated model outputs from the present study. A holistic framework links vibration-based fault detection, explainability, and engineering interpretation to facilitate maintenance choices. Its scope is maintenance-oriented fault detection and decision support rather than remaining useful life (RUL) prediction.

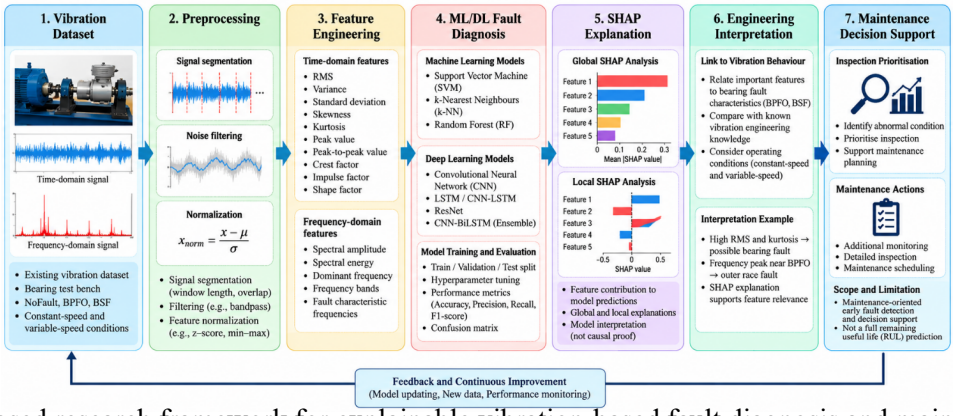


Figure 1. Proposed research framework for explainable vibration-based fault diagnosis and maintenance decision support in rotating machinery.

The framework shown in **Figure 1** links an existing vibration dataset with candidate preprocessing and feature-engineering procedures, fault classification, SHAP-based explanation, engineering interpretation and potential maintenance actions. The feedback loop represents a proposed future extension. The inclusion of a stage does not establish that the procedure has been implemented or empirically validated. **Figure 1** is a conceptual illustration of the proposed framework and does not present empirical findings from the present study.

3. Research Methodology

This study proposes a conceptual and methodological framework for vibration-based fault diagnosis in rotating machinery. The proposed framework outlines vibration data analysis, preprocessing, feature extraction, feature screening, machine-learning and deep-learning classification, model evaluation, SHAP-based explainability, engineering interpretation, and predictive-maintenance decision support.

The proposed framework follows the sequence: Vibration Dataset → Preprocessing → Feature Engineering → Feature Screening → ML/DL Modelling → Fault Diagnosis → Model Evaluation → SHAP-Based Explainability → Engineering Interpretation → Predictive-Maintenance Decision Support.

The framework outlines a proposed analytical process for connecting fault classification with model interpretability and maintenance decision support. Its diagnostic performance and practical effectiveness require independent empirical

validation.

The framework considers diagnostic performance and model interpretability as key aspects of vibration-based fault diagnosis. However, the proposed analytical procedures and their effectiveness have not yet been independently validated.

3.1. Research Design

This study proposes a conceptual research design for vibration-based fault diagnosis using an existing rotating-machinery bearing dataset. The framework considers three bearing conditions: NoFault, Ball Pass Frequency Outer Race (BPFO), and Ball Spin Frequency (BSF) fault. It comprises five stages: operating-condition and dataset description, vibration preprocessing, time- and frequency-domain feature extraction, ML/DL-based fault diagnosis, and SHAP-based explainability with engineering interpretation. Model training, evaluation, and SHAP analysis are proposed components of the framework rather than completed experimental procedures.

3.2. Machinery and Dataset

This study considers an existing rotating-machinery bearing dataset reported in [2] as the reference dataset for the proposed framework. The dataset includes three bearing conditions: NoFault, BPFO fault, and BSF fault.

Two triaxial accelerometers recorded vibration signals at a sampling frequency of 20 kHz. The dataset covers rotational speeds ranging from 0 to 5,000 rpm under constant- and variable-speed conditions. Raw acceleration signals, signal envelopes, and time- and frequency-domain features are considered as possible data representations.

The dataset contains 108 processes, including 48 constant-speed processes and 60 variable-speed processes. The constant-speed subset comprises 28 training, 10 validation, and 10 testing processes, while the variable-speed subset comprises 36 training, 12 validation, and 12 testing processes. Overall, the dataset includes 64 training, 22 validation, and 22 testing processes, accounting for 59.3%, 20.4%, and 20.4%, respectively.

The original dataset contains approximately twice as many BPFO samples as BSF and NoFault samples, indicating class imbalance. Class-balancing methods, such as downsampling, may be considered in a future implementation of the proposed framework. The original class distribution and any subsequent data-balancing procedures should be reported when empirical model training and evaluation are conducted.

3.3. Data Partitioning and Leakage Prevention

The proposed framework uses vibration signals for fault diagnosis. The dataset description includes training, validation, and testing subsets defined at the process level. Process-level partitioning helps maintain independence between the subsets and reduces the risk of data leakage.

Signal windows should be generated within each process after process-level partitioning. Scaling, feature selection, and resampling procedures should be fitted using training data only and then applied to the validation and testing subsets as appropriate. This strategy is intended to prevent information leakage during model development.

The proposed evaluation strategy involves dividing the dataset into training, validation, and testing sets at the process level. Process independence should be maintained to prevent data leakage during preprocessing, feature selection, and model evaluation.

3.4. Data Pre-processing

Vibration signal preprocessing prepares the raw signals for subsequent feature

extraction and fault diagnosis. The proposed procedures include signal-quality assessment, noise filtering, segmentation, normalization, and frequency-domain transformation.

Fourier analysis of spectral characteristics is based on vibration signals sampled at 20 kHz. Processing choices should preserve fault information and be consistent across data subsets.

The proposed preprocessing procedures for future empirical implementation include noise filtering, signal segmentation, normalization, and FFT-based frequency-domain transformation. Specific parameter settings, including filter cut-off frequencies, filter order, segmentation length, overlap, and FFT window functions, should be defined and reported during future empirical implementation.

3.5. Feature Extraction

Feature extraction considers complementary time-domain and frequency-domain characteristics.

3.5.1. Time-Domain Features

The extracted features include mean, RMS, variance, SD, peak, peak-to-peak, skewness, kurtosis, crest factor, impulse factor, and form factors. These features characterize signal magnitude, variability, distribution asymmetry, and impulsive behavior.

RMS is expressed as:

$$RMS = \sqrt{\frac{1}{N}\sum_{i=1}^{N} x_i^2} \quad (1)$$

Kurtosis is expressed as:

$$K = \frac{\frac{1}{N}\sum_{i=1}^{N}(x_i-\overline{x})^4}{\sigma^4} \quad (2)$$

Skewness is expressed as:

$$S = \frac{\frac{1}{N}\sum_{i=1}^{N}(x_i-\overline{x})^3}{\sigma^3} \quad (3)$$

Here, (x_i) is the i-th observation, (N) is the number of observations, (x) is the signal mean, and (σ) is its standard deviation. Brusa et al. [3] contextualize statistical feature contributions in explainable machine failure detection. However, these statistical features do not necessarily indicate a specific bearing fault.

3.5.2. Frequency-Domain Features

Frequency-domain analysis examines dominant frequency, spectral amplitude, frequency-band energy, and other spectral parameters. The discrete Fourier transform is:

$$X(k) = \sum_{n=0}^{N-1} x(n)\, e^{-j2\pi kn/N} \quad (4)$$

Here, *x(n)* represents the time-domain signal, *X(k)* represents its frequency-domain representation, *(N)* is the number of samples and *(k)* is the frequency-bin index.

The vibration-energy distribution across frequencies is described via spectral analysis. Signal processing, bearing geometry, and rotation speed are needed to interpret BPFO or BSF components. A spectral peak alone does not confirm the presence of a bearing defect.

3.5.3. Multi-Domain Representation

Time-domain features describe signal magnitude and statistical behaviour, whereas frequency-domain features describe spectral characteristics. Their combined use provides a potential basis for classification and subsequent explanation. The time- and frequency-domain features described above represent candidate features within the proposed framework. Their actual extraction, selection, and use as model inputs require further empirical implementation and validation.

3.6. Feature Screening and Selection

Feature screening identifies redundant or uninformative variables using descriptive statistics, correlation analysis, and multicollinearity assessment. Low-information or redundant features may be removed before model classification.

Feature selection and SHAP serve different purposes. Feature selection determines the input features used for model training, whereas SHAP explains their contributions to model predictions. Feature screening should be performed using training data only to prevent information leakage. A high SHAP contribution indicates a stronger influence on the model prediction but does not necessarily imply greater physical significance or establish a fault cause.

3.7. Machine-Learning and Deep-Learning Models

The candidate conventional models are Support Vector Machines (SVM), k-Nearest Neighbours (k-NN), Random Forest (RF) and XGBoost. The candidate deep-learning architecture is CNN-BiLSTM. SVM and k-NN use different classification principles, while Random Forest and XGBoost employ tree-based ensemble learning. CNN-BiLSTM combines convolutional feature extraction with bidirectional recurrent learning. Their suitability depends on the input representation, dataset characteristics and evaluation conditions.

[2] reported 93.7% accuracy for CNN-BiLSTM under variable-speed conditions, 99.8% accuracy for ResNet under constant-speed conditions and 99.3% overall accuracy for a deep ensemble. These are published benchmarks, not independently obtained results of the present study. They also concern different evaluation settings and should not be interpreted as a controlled comparison of the candidate models.

The available methodological record does not specify kernel parameters, tree counts, learning rates, network layers, dropout, batch size or training epochs. These settings are not assumed. Future empirical implementation of the framework should identify the models to be trained, their configurations, and the data used for training, validation, and testing. These procedures have not been conducted in the present study and remain part of future empirical implementation.

3.8. Model Evaluation

The proposed evaluation measures include accuracy, precision, recall, F1-score, and the confusion matrix. The F1-score is calculated as:

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (5)$$

The F1-score is the harmonic mean of precision and recall. The confusion matrix summarizes the correct and incorrect classifications for the NoFault, BPFO, and BSF conditions. Given the class imbalance in the dataset, class-wise evaluation measures are important.

Rotational speed can influence vibration patterns and classification. Therefore, constant-speed and variable-speed conditions should be evaluated separately. While process-level splitting minimizes data leakage, it does not guarantee diagnostic robustness across different operating conditions.

ROC-AUC may also be considered as an additional evaluation metric in future empirical implementation. No independently generated model evaluation metrics or confusion matrices are reported in the present study, as candidate models have not been independently trained and evaluated.

3.9. SHAP-Based Explainability

SHAP is suggested for feature contributions to model predictions. The additive representation is:

$$f(x) = \varnothing_0 + \sum_{i=1}^{M} \varnothing_i \quad (6)$$

where $f(x)$ represents model output for observation x, ϕ_0 represents the baseline output and ϕ_i represents the contribution of **feature i.**

3.9.1. Global and Local Analysis

Global SHAP analysis ranks features according to their mean absolute SHAP values, indicating their overall contributions to model predictions. Local SHAP analysis explains the feature contributions to individual predictions relative to the baseline output. The proposed framework is designed to classify bearing conditions as NoFault, BPFO, or BSF. However, no independently generated SHAP explanations, observation-level results, or model output parameters are available in the present study. Therefore, no specific feature contributions are assumed.

3.9.2. Engineering Interpretation

SHAP contributions can help identify features that influence model predictions. However, diagnostic hypotheses based on these explanations should be assessed against characteristic fault frequencies, operating conditions, and relevant engineering knowledge.

The interpretation process follows: Model Prediction → Feature Contribution → Engineering Interpretation

SHAP explains feature contributions to model predictions rather than physical fault mechanisms or causal relationships. Empirical SHAP visualizations require trained models, defined input data, and reproducible analysis procedures.

3.10. Engineering Interpretation and Predictive-Maintenance Decision Support

The proposed decision-support process combines fault categorisation, engineering interpretation, and maintenance-focused decision assistance: Vibration Features → Fault Prediction → SHAP Explanation → Engineering Interpretation → Inspection Priority → Maintenance-Oriented Decision Support

Feature-level explanations may assist maintenance engineers in locating potential problems. If a characteristic fault-related spectral pattern is identified, targeted bearing inspection may be considered after technical verification. This represents a proposed decision pathway rather than a validated maintenance rule.

The applicability of the proposed framework may be limited by variations in operating conditions, sensor noise, the three bearing conditions considered, and the laboratory test-bench setting. These factors may affect its generalizability to industrial environments and require further validation. The framework prioritizes fault identification and maintenance decision support over RUL prediction because failure records, maintenance events, and degradation trajectories are not available for RUL modelling. The proposed framework represents a conceptual approach and should be distinguished from published benchmarks and independently validated experimental findings.

4. Framework Illustration and Literature-Based Evidence

4.1. Dataset Characteristics

This section describes the vibration dataset reported in [2], which concerns rolling-element bearings under three conditions: NoFault (healthy), BPFO, and BSF. The dataset includes both constant-speed and variable-speed operating conditions and provides a reference for the proposed framework for vibration-based fault diagnosis and predictive-maintenance decision support.

The vibration signals were recorded using two triaxial accelerometers at a sampling frequency of 20 kHz on a rotating-machinery test bench. The dataset covers rotational speeds of up to 5,000 rpm under constant-speed and variable-speed conditions. Available data representations include raw acceleration

signals, signal envelopes, and time- and frequency-domain features. The dataset characteristics and data partitioning are summarized in **Table 2**.

Table 2. Dataset Characteristics

Dataset characteristic	Data
Machine type	Rotating machinery / bearing test bench
Monitored component	Rolling-element bearings
Classification classes	3
Class 1	NoFault (Healthy)
Class 2	BPFO
Class 3	BSF
Sensor type	Triaxial accelerometers
Number of sensors	2
Sampling frequency	20 kHz
Maximum rotational speed	5000 rpm
Operating conditions	Constant-speed and variable-speed
Data representation	Raw acceleration signals and signal envelopes
Time-domain information	Statistical vibration descriptors
Frequency-domain information	Spectral and fault-frequency characteristics
Constant-speed processes	48
Variable-speed processes	60
Total processes	108
Constant-speed training processes	28
Constant-speed validation processes	10
Constant-speed testing processes	10
Variable-speed training processes	36
Variable-speed validation processes	12
Variable-speed testing processes	12
Overall training processes	64
Overall validation processes	22
Overall testing processes	22
Approximate overall split	59.3% / 20.4% / 20.4%

Source: [2].

The dataset described in [2] comprises 108 processes, including 48 constant-speed and 60 variable-speed processes. For future empirical evaluation, the proposed framework considers a process-level data partition of 64 training, 22 validation, and 22 testing processes, corresponding to approximately 59.3%, 20.4%, and 20.4%, respectively. This partition is proposed for future model development and evaluation and does not represent an independently conducted experimental procedure in the present study.

The original class distribution should be verified against the dataset documentation. To address the class imbalance, downsampling of the BPFO class may be considered in future empirical implementation of the proposed framework. Since this procedure has not been performed in the present study, no post-downsampling class distribution is reported.

Future empirical evaluation should use consistent process-level data partitions. Preprocessing, scaling, and feature selection should be fitted using training data only, while the test set should be reserved for final evaluation. These procedures are intended to reduce data leakage and support a more reliable assessment of model generalization.

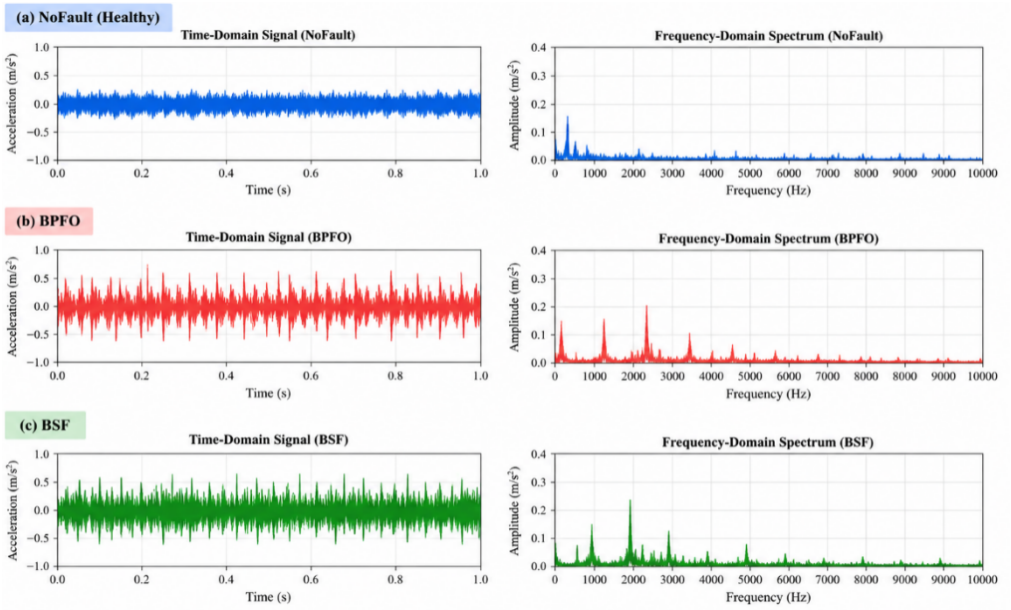


Figure 2. Representative Vibration Signals for NoFault (Healthy), BPFO, and BSF Conditions. (These waveforms are illustrative and do not represent measured signals.)

Figure 2 presents schematic illustrations of vibration patterns under three bearing conditions. These waveforms are intended to illustrate conceptual differences and are not derived from measured signals or used as experimental evidence.

4.2. Vibration-Feature Analysis

For vibration-based flaw diagnostics, time-domain and frequency-domain features work together. Frequency-domain characteristics describe spectral patterns and potential fault-related frequency components, whereas time-domain characteristics describe signal amplitude and statistical distribution.

Candidate time-domain features in the proposed framework include RMS, variance, standard deviation, skewness, kurtosis, peak value, peak-to-peak value, crest factor, impulse factor, and form factor. These features characterize vibration amplitude, dispersion, and waveform shape. For example, skewness describes distributional asymmetry, while shape factor is defined as the ratio of the root mean square value to the mean absolute value of a signal.

[3] demonstrated the diagnostic relevance of statistical vibration features through SHAP-based interpretation. SMRA, form factor, skewness, and frequency were identified as important features in that study. However, the importance of these features may vary across datasets and models.

Frequency-domain features include dominant frequency, spectral amplitude, and frequency-band energy. In 2025, [2] reported bearing classification using raw acceleration signals, signal envelopes, and time- and frequency-domain representations. These representations may provide complementary information for bearing-condition characterization. Since operating speed can affect characteristic frequency components, spectral interpretation should account for rotational speed.

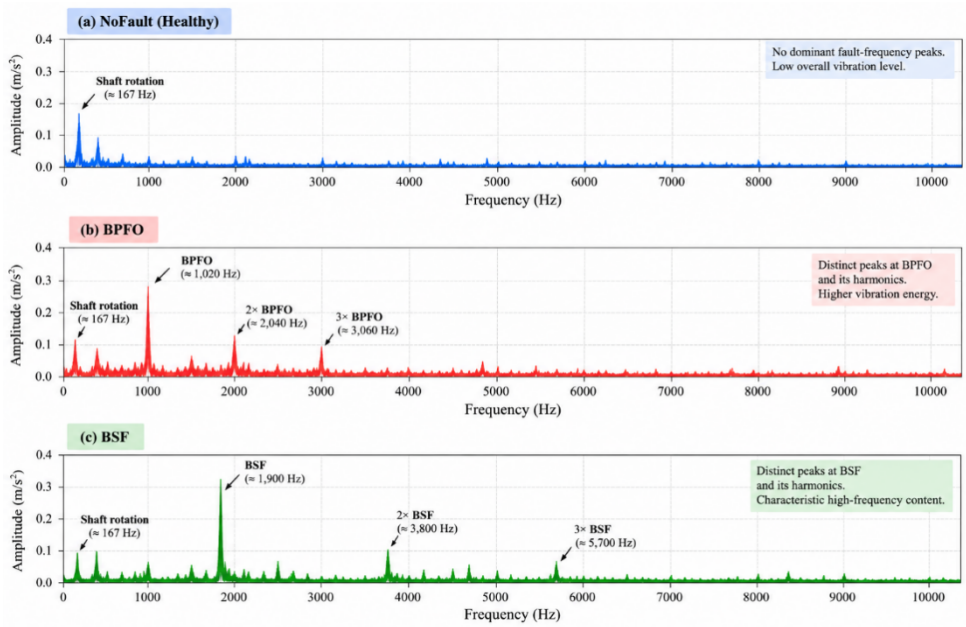


Figure 3. Illustrative FFT Spectra for NoFault (Healthy), BPFO, and BSF Conditions. (These spectra are conceptual examples and do not represent results calculated from measured signals in the present study.)

Figure 3 presents schematic illustrations of FFT spectra under NoFault, BPFO, and BSF conditions. These spectra are illustrative and do not represent actual FFT results generated in the present study.

4.3. Published Model Performance Benchmarks and Illustrative SHAP Global Analysis

Previous studies have used traditional machine-learning and deep-learning methods for vibration-based fault diagnosis. However, reported accuracy depends on the dataset, signal representation, model configuration, operating conditions, and evaluation protocol. Results from different studies should therefore be interpreted as published benchmarks rather than as a controlled comparison.

[3] reported classification accuracies above 98.5% for support vector machines (SVM) and k-nearest neighbours (k-NN) in their bearing-diagnosis setting. [2] reported 87.2% accuracy for a raw-acceleration baseline and 99.3% overall accuracy for a CNN-BiLSTM deep ensemble. Study [2] also reported 93.7% accuracy for CNN-BiLSTM under variable-speed conditions and 99.8% for ResNet under constant-speed conditions. All these values are published benchmarks from previous studies and were not independently obtained or verified through model evaluation in the present study.

These reported values indicate that diagnostic performance may vary with model architecture, signal representation, dataset preparation, operating conditions, and evaluation procedures. However, because the studies were conducted under different experimental settings, these results should be interpreted as published benchmarks rather than used to establish a definitive ranking of algorithms. A controlled comparison would require consistent datasets, process-level data partitions, model configurations, and evaluation procedures. The 99.3% reported for the CNN-BiLSTM ensemble represents the overall performance reported in [2], including accuracy, recall, and F1-score. Therefore, these values summarize published evidence rather than presenting an independent performance evaluation of the proposed framework. **Figure 4** illustrates, using hypothetical values, how a confusion matrix summarizes class-wise results for the NoFault, BPFO, and BSF conditions; the preparation of **Figures 4–7** is explained in the note following **Figure 7**.

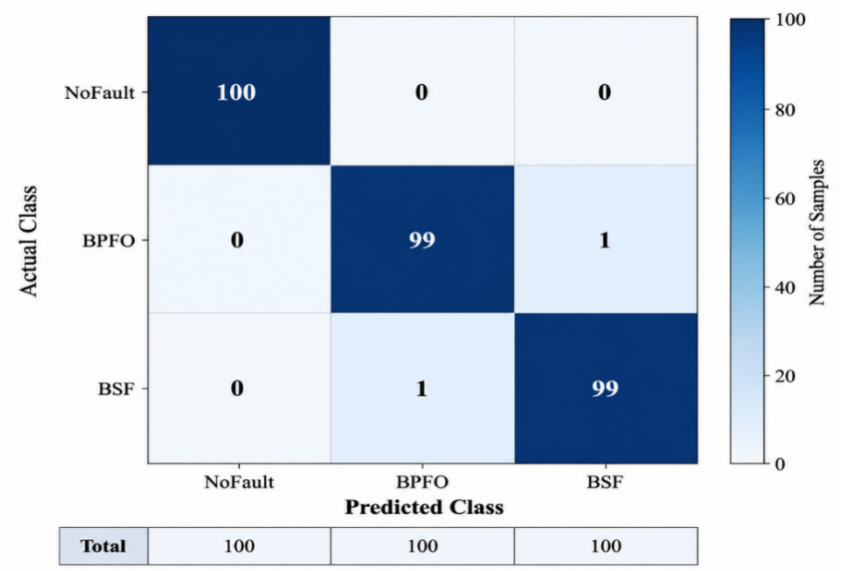


Figure 4. Illustrative Confusion Matrix for NoFault, BPFO, and BSF Classification.

Source: Authors' own illustration.

Beyond classification accuracy, SHAP can be used to explain the contributions of input features to model predictions. Global SHAP analysis uses mean absolute SHAP values to summarize feature contributions across observations, whereas local SHAP analysis focuses on individual predictions.

Study [3] also applied SHAP to SVM- and k-NN-based models for vibration-based fault diagnosis. The study identified SMRA, shape factor, skewness, and frequency-related features as important contributors to model predictions. The reported features and their interpretations are summarized in **Table 3**.

Table 3. SHAP Features Reported by [3]

Feature	Domain	Reported interpretation
SMRA	Time	Important for defect-free-class recognition
Shape factor	Time	Important for fault discrimination
Skewness	Time	Relevant diagnostic feature
RMS frequency	Frequency	Highly important for one fault class
p19	Frequency-related	Greater impact in k-NN

Source: [3]. Interpretations refer to the cited study and are not newly calculated SHAP results.

Feature importance should be interpreted in the context of the model, input representation, and experimental conditions. SHAP attribution explains how features contribute to model predictions but does not establish causal relationships or identify the physical causes of faults. The feature rankings reported in [3] should not be directly transferred to the models or dataset discussed in [2], as they were obtained under different experimental settings. **Figure 5** illustrates the format of a global SHAP feature-importance plot for these features using hypothetical values.

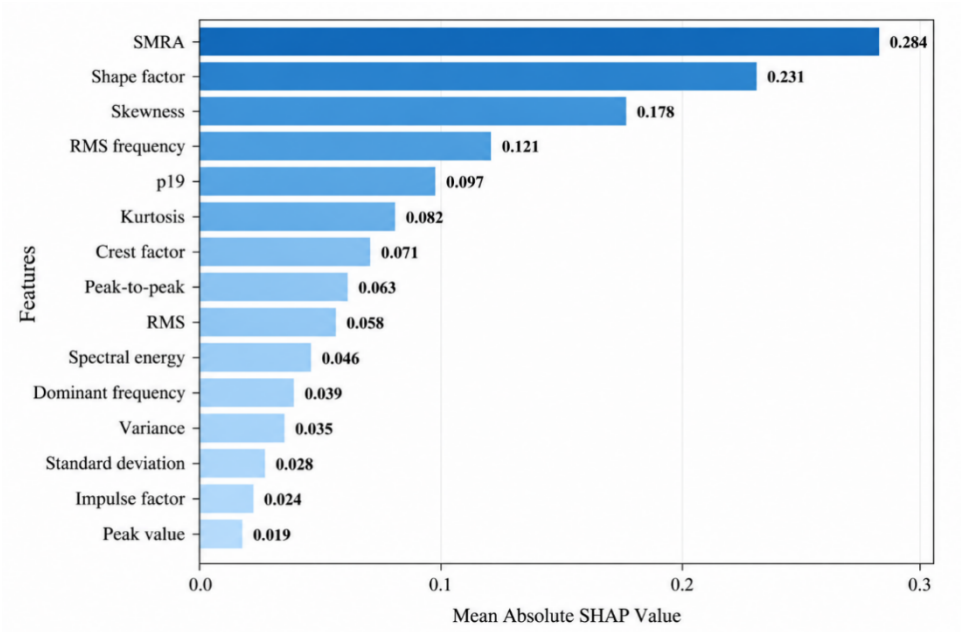


Figure 5. Illustrative Global SHAP Feature Importance for a Fault-Diagnosis Model.
Source: Authors' own illustration.

4.4. Illustrative SHAP Local Analysis

Local SHAP analysis can identify the contributions of individual features to a specific model prediction. Local explanations can be used to examine predictions for the three bearing conditions: NoFault, BPFO, and BSF. **Figure 6** presents an illustrative beeswarm plot, and **Figure 7** presents illustrative waterfall plots for one example of each condition.

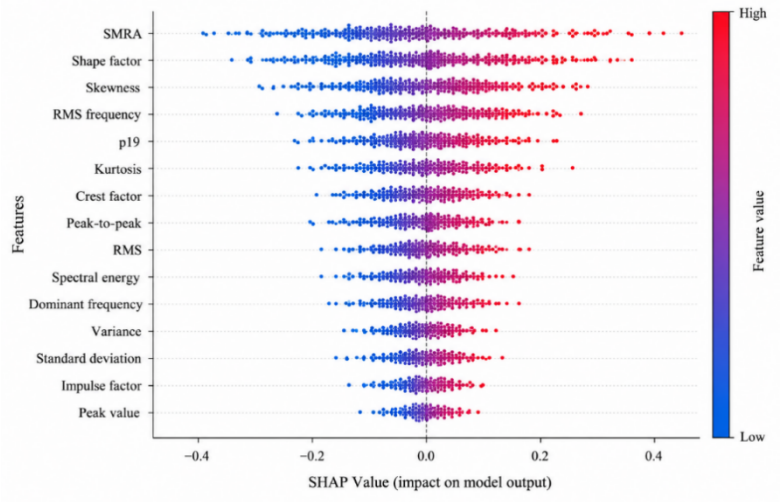


Figure 6. Illustrative SHAP Beeswarm Plot for a Fault-Diagnosis Model.
Source: Authors' own illustration.

A SHAP beeswarm plot typically identifies the model, input features, and output class or model output. It illustrates the distribution of feature contributions across observations and shows how feature values are associated with their contributions to model predictions.

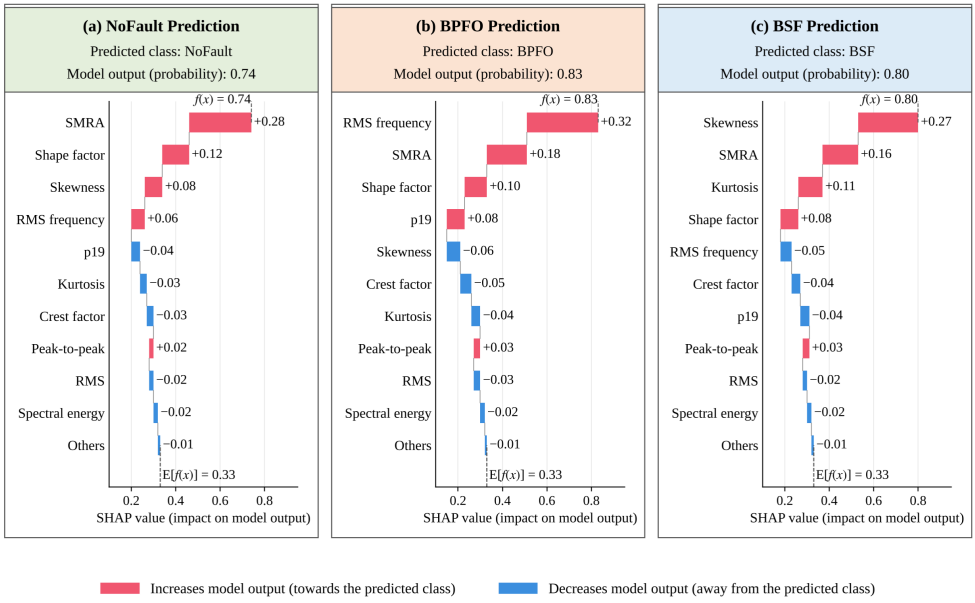


Figure 7. Illustrative SHAP Waterfall Plots for NoFault, BPFO, and BSF Classification Examples.
Source: Authors' own illustration.

SHAP waterfall plots illustrate how individual feature contributions move the model output from its baseline value to the final prediction.

Note on **Figures 4–7**: Figures 4–7 were created independently by the authors as conceptual illustrations; they were not reproduced or adapted from any published figure, and no reproduction permission was therefore required. No model was trained and no SHAP computation was performed to produce them. All numerical values were specified by the authors solely to demonstrate the format and interpretation of each visualization. In **Figure 4**, the confusion-matrix entries assume a hypothetical test set of 100 samples per class with one BPFO–BSF misclassification in each direction, giving an overall accuracy of 298/300 $\approx$ 99.3%, which is of the same order as the published benchmark in [2]; the class-wise entries are not taken from [2]. In Figures 5 and 6, the five leading features (SMRA, shape factor, skewness, RMS frequency, and p19) correspond to the features reported as important in [3] (**Table 3**), and the remaining features are candidate time- and frequency-domain features described in Section 3.5. The mean absolute SHAP values in **Figure 5** were assigned by the authors as decreasing hypothetical magnitudes, and the points in **Figure 6** are synthetic values generated to follow the same feature ordering; no numerical value was taken from [3]. In **Figure 7**, the baseline value E[f(x)] $= 0.33$ corresponds to a hypothetical reference set with equal proportions of the three classes, the feature contributions were specified by the authors, and each prediction probability was obtained as the baseline plus the sum of the feature contributions, consistent with Equation (6): $0.33 + 0.41 = 0.74$ (NoFault), $0.33 + 0.50 = 0.83$ (BPFO), and $0.33 + 0.47 = 0.80$ (BSF). These values do not represent model predictions or SHAP results from the present study.

4.5. Robustness and Generalisation Considerations

The dataset described in [2] includes both constant-speed and variable-speed operating conditions. Study [2] reported different classification accuracies for CNN-BiLSTM and ResNet under constant-speed and variable-speed conditions, as well as an overall accuracy for a CNN-BiLSTM ensemble. These published results are summarized in **Table 4**.

Table 4. Published Model Performance Under Operating Conditions

Operating condition	Model	Accuracy (%)
Constant speed	ResNet	99.8
Variable speed	CNN-BiLSTM	93.7
Constant and variable speed	CNN-BiLSTM ensemble	99.3

Source: [2]. These values are published benchmarks and were not independently obtained in the present study.

These published results indicate that operating conditions and model architectures may influence diagnostic performance. However, the reported accuracy values should not be used to establish a direct comparison of model robustness because the models were evaluated under different experimental settings and input representations. The 93.7% accuracy reported for CNN-BiLSTM refers to its performance under variable-speed conditions and does not represent robustness to load fluctuations.

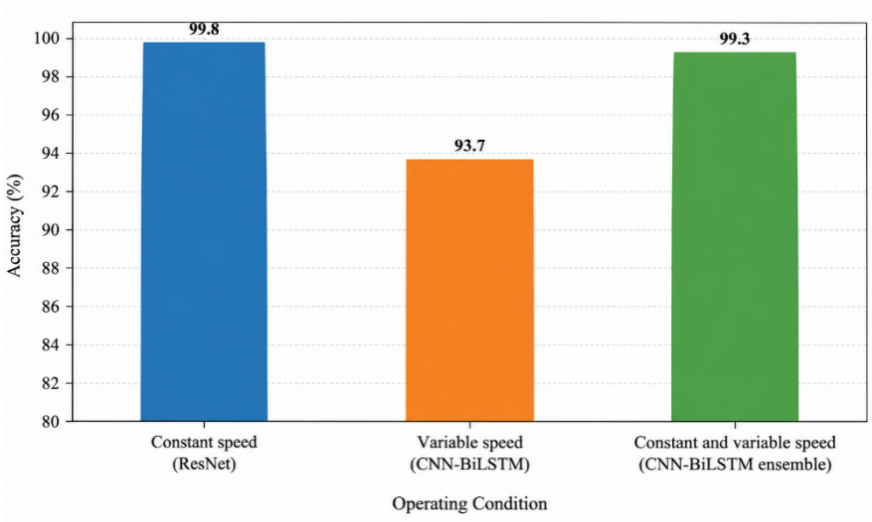


Figure 8. Published Model Performance Benchmarks Across Operating Conditions.

Figure 8 presents an illustrative comparison of the published performance benchmarks summarized in **Table 4**, with each value linked to its source, model, and operating condition. These results provide a descriptive overview rather than a controlled comparison of model robustness. Future empirical validation should evaluate selected models under both constant-speed and variable-speed conditions using consistent data partitions and evaluation metrics. Class imbalance, data leakage, and classification errors should also be considered. Such validation has not been conducted in the present study.

4.6. Predictive-Maintenance Implications

The proposed framework suggests that explainable machine learning and vibration-based fault diagnosis may support maintenance decision-making. Through feature attribution, engineers can better understand the features contributing to model predictions.

Study [2] reported 87.2% accuracy for a raw-acceleration baseline and 99.3% overall accuracy for a CNN-BiLSTM ensemble. These results provide a reference for future investigations of ensemble approaches and different signal representations. Study [3] used SHAP to investigate the contributions of vibration features to diagnostic model predictions. However, the effectiveness of the proposed framework has not been independently validated using additional data.

Possible maintenance path: Vibration Features → Fault Prediction → SHAP Explanation → Engineering Interpretation → Inspection Priority →

Maintenance-Oriented Decision Support

Potential diagnostic issues highlight the need for engineering assessment and confirmatory inspection. Prioritizing maintenance tasks requires consideration of machine criticality, operating conditions, fault severity, and the potential consequences of failure. Although SHAP explanations may provide useful insights, they should not replace engineering judgment or condition-monitoring data in maintenance decision-making.

Rather than focusing on RUL prediction, the proposed framework prioritizes fault classification and maintenance decision support. RUL estimation requires longitudinal data on equipment degradation and failure times. Estimates of remaining useful life are not supported by the findings of the three-class classification framework.

Diagnostic performance observed on laboratory test benches may not fully reflect performance under industrial operating conditions. Factors such as sensor noise, improper installation, load variations, and environmental interference may affect diagnostic performance. Therefore, independent validation under realistic operating conditions is necessary. Such validation should assess diagnostic performance, model explainability, and the practical relevance of engineering interpretations for maintenance decision-making.

5. Discussion, Implications And Limitations

5.1. Diagnostic Performance

The reviewed literature demonstrates the potential of vibration-based machine learning for rotating-machinery and bearing-fault classification. [3] reported accuracies above 98.5% for support vector machines (SVM) and k-nearest neighbours (k-NN) using statistical vibration features. [2] reported 87.2% accuracy for a raw-acceleration baseline and 99.3% overall accuracy for a CNN-BiLSTM ensemble. Their study also reported 99.8% accuracy for ResNet under constant-speed conditions and 93.7% for CNN-BiLSTM under variable-speed conditions.

These results highlight the importance of signal representation and model configuration. Deep-learning approaches offer greater flexibility in learning signal representations, whereas traditional machine-learning methods may also perform well with appropriate data. Studies [12,19] and [9] emphasize feature representation, model development, and generalization. All of the following should be considered when choosing a model: accuracy, recall, explainability, computing requirements, resilience, confusion patterns, and F1-score. Direct comparison of constant-speed and variable-speed results is not appropriate because model configurations and experimental conditions differ. These values represent published benchmarks rather than results produced by the proposed framework.

5.2. Vibration-Feature Interpretation

Time-domain and frequency-domain features provide complementary information for bearing fault diagnosis. Study [2] examined raw acceleration signals, signal envelopes, and vibration descriptors, while study [3] identified skewness and shape factor as important features in SHAP-supported classification.

Most of these are statistical time-domain features. Frequency-domain features include dominant frequency, spectral amplitude, and frequency-band energy. Additionally, study [17] shows that machinery fault diagnosis involves preprocessing, feature extraction, and signal representation.

Engineering interpretation goes beyond statistical significance. The shape factor describes the relationship between the RMS value and the mean absolute value of a waveform, while skewness measures distributional asymmetry.

Frequency-related features may help identify patterns associated with mechanical excitation, but their interpretation depends on rotational speed and signal processing. A feature's contribution to a model prediction does not establish that it caused a fault. Accordingly, explanations should be assessed against expected vibration behaviour, operating conditions and known fault characteristics.

5.3. Explainability

The proposed conceptual framework is designed to support the interpretation of model predictions by integrating explainability into vibration-based fault diagnosis. Study [3] used SHapley Additive exPlanations (SHAP) to estimate feature contributions to model outputs and interpret SVM and k-NN predictions.

Global SHAP analysis aggregates feature contributions across data, whereas local SHAP analysis explains specific predictions. Study [16] emphasizes the importance of reliability, stability, robustness, and physical consistency in explainable AI for industrial applications. Study [6] also discusses the computational demands and feature-ranking aspects of explainable fault-diagnosis approaches.

SHAP values explain model predictions, not physical causes. A substantial SHAP contribution indicates that a feature has a strong influence on a model prediction, but it does not establish the physical cause of a bearing failure. Modeling and datasets may alter feature rankings. Thus, SHAP explanations should be considered alongside engineering verification, physical analysis, and expert judgment in the proposed framework. Empirical SHAP findings require a trained model, defined input data, and reproducible SHAP analysis.

5.4. Maintenance-Oriented Predictive Monitoring

Vibration-based fault classification can identify patterns associated with abnormal machinery conditions and inform inspection decisions. The maintenance pathway proposed in this paper is: Vibration Measurement → Fault Classification → SHAP Explanation → Engineering Interpretation → Inspection Priority → Maintenance-Oriented Decision Support

SHAP can help identify the features contributing to model predictions of BPFO or BSF faults. Engineers may examine these contributions in relation to expected mechanical behaviour and decide whether further measurements or inspections are needed.

Study [20] demonstrates how the integration of vibration and motor-current data, a Composite Health Index, and temporal decision-layer processing can help reduce false alarms in industrial applications. Integrating complementary measurements and decision-making processes may therefore be beneficial.

The proposed framework integrates model explainability and engineering assessment into fault classification. The proposed framework supports fault detection and maintenance-related decision-making but does not constitute a predictive maintenance system. Further evaluation under realistic industrial operating conditions is needed to assess the framework's practical applicability and its potential benefits for maintenance operations.

5.5. Theoretical Implications

The proposed conceptual framework integrates condition monitoring, explainable AI, and machine learning. It emphasizes the relationship between fault classification, feature attribution, and engineering verification.

Model evaluation should consider not only classification performance but also explanation quality. It is important to distinguish statistical feature attribution from physical interpretation, as SHAP explanations do not directly establish the physical causes of machinery faults. Engineering knowledge is therefore necessary to interpret model explanations and assess their relevance to

maintenance decisions [16].

Effective evaluation is crucial. The dataset includes processes under both constant-speed and variable-speed conditions. Studies [5,7] and [1] highlight challenges related to operating-condition variations, limited fault data, and model generalization.

The proposed conceptual pathway is: Diagnostic Performance → Model Explanation → Engineering Verification → Maintenance Decision Support

The proposed framework contributes an integrated approach to diagnostic explanation and maintenance decision support. However, its advantages over existing integrated approaches require further comparative evaluation.

5.6. Practical Implications

The proposed framework has four main practical implications. Vibration-based machine learning can classify NoFault, BPFO, and BSF conditions when supported by appropriately labelled data and systematic evaluation.

Furthermore, SHAP can help engineers understand the contributions of individual features to model predictions. These explanations can support fault diagnosis but do not directly establish the physical causes of machinery faults.

The third consideration is the machine's criticality, the type of fault, and its severity. Maintenance decisions should consider operating conditions, potential failure consequences, and fault severity rather than relying solely on predicted fault classes.

Fourth, practical implementation requires consideration of operating conditions and deployment constraints. The framework should be evaluated under both constant-speed and variable-speed conditions. Study [16] discusses challenges related to computational costs in industrial deployment, explanation stability, and sensor variability. These factors should be considered when assessing the practical applicability of the proposed framework.

5.7. Limitations

A number of factors limit the applicability of the proposed framework. The proposed framework is designed with reference to a bearing test-bench dataset covering three conditions: NoFault, BPFO, and BSF. Laboratory test-bench conditions may not fully reflect real industrial operating environments, even when both constant-speed and variable-speed conditions are considered.

Secondly, the framework is limited to predefined fault categories and labelled samples, which may restrict its ability to identify unknown or unexpected defects. Incomplete descriptions of operating conditions and class imbalance may also affect classification performance. Studies [7] and [1] discuss limitations related to fault-data availability and model generalization.

The dataset does not provide sufficient information on load, temperature, lubrication, sensor location, or environmental interference. External validation using different industrial datasets and equipment is therefore necessary.

Fourth, SHAP shows feature contributions, not causation. The stability and physical consistency of explanation rankings should be assessed across different models and datasets.

Fifth, previous studies [2] and [3] have reported classification accuracies of 87.2%–99.8% and above 98.5%, respectively. The confusion matrices and SHAP graphs presented in this paper are illustrative examples and do not represent model predictions or calculations independently generated in the present study. RUL estimation involves deterioration trajectories, failure-time data, and maintenance histories in addition to three-class classification. Therefore, the proposed framework focuses on fault identification and maintenance decision

support rather than RUL prediction.

5.8. Future Research

Future evaluation should use industrial datasets covering different machines, rotational speeds, loads, operating conditions, sensor combinations, and fault severity levels. Closing the training-deployment gap requires domain adaptation and transfer learning. For example, the CNN-based transfer learning experiment reported in [12] achieved 99.7% classification accuracy in the target domain, providing a reference for future research on domain adaptation.

Another promising direction is zero-shot fault diagnosis, which may enable the identification of previously unseen fault categories under variable-load conditions. Reference [21] provides a basis for further investigation in this area. Multisensor integration may combine motor-current, temperature, vibration, and acoustic signals for more comprehensive machinery monitoring. The industrial example of [20] provides a framework for researching complementary signals and health-index approaches. Fuzzy-neuro models that combine learned vibration signatures with physics-based indicators, such as RMS and kurtosis, to produce interpretable maintenance-risk scores [22], and fuzzy risk interpretation of acoustic signal descriptors [23] also provide a basis for linking complementary measurements with graded maintenance actions.

Future XAI studies should evaluate explanation fidelity, stability, robustness, localization, physical consistency, and computational cost rather than relying solely on visual SHAP rankings. Standardized evaluation and physically consistent explanations are important research priorities in explainable AI for rotating-machinery fault diagnosis [16]. When such evaluations move towards industrial implementation, experience from machine-learning applications in automotive manufacturing and operations [24] and from cloud-based deployment of machine-learning models [25] indicates that scalability and computing infrastructure should also be considered.

Finally, future studies could extend the proposed conceptual framework by connecting explainable fault diagnosis with health-index development, uncertainty quantification [26], degradation modelling, and maintenance-event data.

This may lead to a larger architecture: Fault Detection → Explainable Fault Diagnosis → Health Assessment → Maintenance Decision Support

RUL estimation should be incorporated only when appropriate longitudinal degradation and failure data are available and the resulting models have been independently validated.

6. Conclusion

Complex rotating equipment needs dependable, intelligible condition-monitoring systems. Machine learning and deep learning can identify fault patterns in vibration signals to support bearing fault diagnosis. Model and vibration representation selection is stressed in the literature. Study [3] reported classification accuracies above 98.5% for SVM and k-NN, while study [2] reported 99.3% overall accuracy for a CNN-BiLSTM ensemble. These values are published results from previous studies, not experimental results of the present study.

This conceptual study proposes an integrated framework combining vibration data preprocessing, feature engineering, machine-learning- and deep-learning-based fault diagnosis, SHAP-based explainability, engineering interpretation, and maintenance decision support. Diagnostic predictions, interpretable feature contributions, and maintenance inspection priorities may be linked. The proposed framework has not been independently validated through experiments, and its operational performance and predictive capabilities therefore remain to be established.

Explainability should complement fault diagnosis. SHAP enables engineers to examine feature contributions to model predictions at both global and local levels. However, SHAP explains feature contributions to model predictions but does not establish physical causation. Study [16] emphasizes the importance of physical consistency, stability, robustness, and reliability in explainable AI for industrial applications. Model explanations should therefore be interpreted in the context of vibration engineering principles and operating conditions.

Rather than estimating remaining useful life (RUL), the proposed framework focuses on fault identification and maintenance decision support. RUL estimation is beyond the scope of the present framework and would require additional data, including failure times, degradation trajectories, and maintenance records. It should therefore be considered in future research. Future studies should use industrial datasets covering different machines, operating conditions, and fault severity levels to evaluate the proposed framework. Quantitative explainability, uncertainty estimates, multisensor integration, and transfer learning need further investigation.

The proposed framework suggests that explainable machine learning may support maintenance by identifying fault patterns in vibration signals. In future applications, diagnostic performance, engineering verification, and interpretable explanations may collectively support more informed industrial fault-diagnosis and maintenance decisions. The proposed approach is grounded in findings reported in previous studies and provides a conceptual basis for future empirical investigation.

Author contributions: Conceptualization, S.S.P. and T.M.K.; methodology, S.S.P. and T.M.K.; investigation, S.S.P., J.S.J. and P.W.; resources, T.M.K. and K.R.; writing—original draft preparation, S.S.P. and T.M.K.; writing—review and editing, T.M.K., N.R., J.S.J., P.W., K.R., P.Y., Y.N. and T.S.; visualization, S.S.P., Y.N. and T.S.; supervision, T.M.K.; project administration, T.M.K. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Conflicts of interest: The authors declare no conflicts of interest.

References

[1] Neupane D, Bouadjenek MR, Dazeley R, et al. Data-driven machinery fault diagnosis: A comprehensive review. Neurocomputing. 2025, 627,129588. https://doi.org/10.1016/j.neucom.2025.129588.

[2] Martiri L, Esmaili P, Moschetti A, et al. Deep-Learning-Based Bearing fault Classification using vibration signals under Variable-Speed Conditions. Instruments. 2025, 9(4):33. https://doi.org/10.3390/instruments9040033.

[3] Brusa E, Cibrario L, Delprete C, et al. Explainable AI for Machine Fault diagnosis: Understanding features' contribution in machine learning models for industrial condition monitoring. Applied Sciences. 2023, 13(4):2038. https://doi.org/10.3390/app13042038.

[4] Viana CAA, Alves DS, Machado TH. Predictive method for machinery fault detection using deep learning and vibration images. Journal of Vibration Engineering & Technologies. 2026,14,341. https://doi.org/10.1007/s42417-026-02554-0.

[5] Sowmya S, Saimurugan M, Edinbarough I. Rotational Machine fault diagnosis using Artificial intelligence (AI) strategies for the operational challenges under variable speed condition: A review. IEEE Access. 2024,12: 144870–144889. https://doi.org/10.1109/access.2024.3469212.

[6] Brito LC, Susto GA, Brito JN, et al. An explainable artificial intelligence approach for unsupervised fault detection and diagnosis in rotating machinery. Mechanical Systems and Signal Processing. 2021, 163:108105. https://doi.org/10.1016/j.ymssp.2021.108105.

[7] Matania O, Dattner I, Bortman J, et al. A systematic literature review of deep learning for vibration-based fault diagnosis of critical rotating machinery: Limitations and challenges. Journal of Sound and Vibration. 2024, 590: 118562. https://doi.org/10.1016/j.jsv.2024.118562.

[8] Almutairi K, Wen H, Sinha JK. Standardisation of vibration-based parameters for rotor and bearing for machine faults detection using machine learning model. Journal of Vibration Engineering & Technologies. 2025, 13,7. https://doi.org/10.1007/s42417-025-02041-y.

[9] Tama BA, Vania M, Lee S, et al. Recent advances in the application of deep learning for fault diagnosis of rotating machinery using vibration signals. Artificial Intelligence Review. 2022,56(5): 4667–4709. https://doi.org/10.1007/s10462-022-10293-3.

[10] Yang D, Karimi HR, Gelman L. An explainable intelligence fault diagnosis framework for rotating machinery. Neurocomputing. 2023,541:126257. https://doi.org/10.1016/j.neucom.2023.126257.

[11] Chen J, Lin C, Yao B, et al. Intelligent fault diagnosis of rolling bearings with low-quality data: A feature significance and diversity learning method. Reliability Engineering & System Safety. 2023,237:109343. https://doi.org/10.1016/j.ress.2023.109343.
[12] Mongia C, Sehgal S. Vibration Signal-Based Fault Diagnosis of rotary machinery through convolutional neural network and transfer learning method. Vibration. 2025,8(2):27. https://doi.org/10.3390/vibration8020027.
[13] Ignjatovska AA, Petreski Z, Gavriloski V, et al. Fault Diagnosis of Rotating Machinery Using Supervised Machine Learning Algorithms with Integrated Data-Driven and Physics-Informed Feature Sets. Sensors, 2026,26(6):1876. https://doi.org/10.3390/s26061876.
[14] Hasan MJ, Sohaib M, Kim JM. An explainable AI-Based fault diagnosis model for bearings. Sensors. 2021,21(12):4070. https://doi.org/10.3390/s21124070.
[15] Amin MA, Ahsan MS, Maua J, et al. Physics-informed self-supervised diagnosis of rotating machinery using latent ODEs and transformer encoders. PLoS ONE. 2026,21(2): e0339239. https://doi.org/10.1371/journal.pone.0339239.
[16] Tang S, Ren Z, Zheng L, et al. Explainable Artificial intelligence in Rotating Machinery Fault diagnosis: A Comprehensive review and Emerging Trends. Sensors. 2026,26(17):5379. https://doi.org/10.3390/s26175379.
[17] Brito LC, Susto GA, Brito JN, et al. Fault Diagnosis using eXplainable AI: A transfer learning-based approach for rotating machinery exploiting augmented synthetic data. Expert Systems With Applications. 2023,232:120860. https://doi.org/10.1016/j.eswa.2023.120860.
[18] Bagri I, Tahiry K, Hraiba A, et al. Vibration Signal Analysis for Intelligent Rotating Machinery Diagnosis and Prognosis: A Comprehensive Systematic Literature Review. Vibration. 2024,7(4):1013–1062. https://doi.org/10.3390/vibration7040054.
[19] ElAraby WS, Madian AH, Saad MH. Fault diagnosis for rotating machinery based on deep learning. Noise & Vibration Worldwide. 2025,56(8-9):400–421. https://doi.org/10.1177/09574565251348863.
[20] Pişmişoğlu A, Ozkat EC, Konar M. Machine Learning-Based predictive maintenance of a wire drawing machine using vibration and motor current signals. Eksploatacja I Niezawodnosc - Maintenance and Reliability. 2027 29(1). (Online: 23 July 2026). https://doi.org/10.17531/ein/225078.
[21] Liao X, Zhou W, Xu Y, et al. Intelligent gearbox fault diagnosis under variable load conditions via Zero-Shot learning. Shock and Vibration. 2026,2026(1). https://doi.org/10.1155/vib/5563009.
[22] Vasudevan A, El Khider M, Yogeesh N, et al. A fuzzy-neuro approach to predictive maintenance using vibration signature classification. Sound & Vibration. 2026,60(2): 3950. https://doi.org/10.59400/sv3950.
[23] Nijalingappa Y, Vasudevan A, Hui SE, et al. Audio signal approximation and fuzzy logic-based acoustic noise-risk assessment using Fourier–Dirichlet analysis and hyperbolic series representations. Sound & Vibration. 2026,60(5):4180. https://doi.org/10.59400/sv4180.
[24] Muhammad YH, Oruabena B, Ul Nabi A. Inclusive review of machine learning in the automotive sector today. SciWaveBulletin. 2023,1(1):7–11. https://doi.org/10.61925/SWB.2023.1102.
[25] Vartika V, William P. Machine learning and cloud computing case study for online app usage tracking. SciWaveBulletin. 2023,1(4): 32–37. https://doi.org/10.61925/SWB.2023.1405.
[26] Nijalingappa Y, Vasudevan A, El Khider M, et al. Type-1 fuzzy inference system for epistemic uncertainty quantification in computational vibro-acoustic simulations. Sound & Vibration. 2026,60(3):3951. https://doi.org/10.59400/sv3951.