Everest Advances in Science and Technology (EAST), Vol. 2, No. 1, 2026
ISSN: 3102-0410 (Print), 3102-0429 (Online)
DOI: 10.3126/east.v2i1.98622

Intelligent Fault Diagnosis of Rolling Element Bearings Using Statistical Feature Extraction and Random Forest Classifier

MANOJ ADHIKARI[1], BIKKI CHHANTYAL[2*], MAHESH CHANDRA LUINTEL[1]

Abstract. Accurate detection of faults within rotary machinery components is vital for assuring the structural reliability of equipment in manufacturing facilities and power generation systems. Automatic fault detection is an increasingly common approach whereby data acquired from sensors are analyzed by machine learning algorithms to distinguish between normal and faulty component states. Vibration signatures recorded from healthy and faulty ball bearings, taken from the Case Western Reserve University (CWRU) 12k Drive End database, were utilized in the present study. These vibration signatures were segmented into windows of 0.1 s duration. Standard statistical time-domain features (root mean square, kurtosis, peak value and crest factor) were extracted from each sample. A random forest (RF) classifier was trained on vibration data from healthy bearings and several types of faulty bearings, partitioned into training (70% of data) and testing (30%) sets. The RF classifier distinguished healthy from faulty bearings with 98.8% accuracy across ball, inner race, outer race and normal categories. Feature importance analysis indicated that RMS (50% importance) and peak value (28% importance) were the most influential contributors, together comprising about 78% of the decision making. The RF model further recorded a recall of 1.00 for the normal baseline category, indicating that healthy bearings were identified with zero false negatives. The results demonstrate that ensemble machine learning, combined with statistically significant time-domain features, provides a highly accurate and physically interpretable solution for industrial condition monitoring, serving as a foundation for more advanced rotor-dynamic fault detection.

Keywords: Bearing fault diagnosis, Vibration analysis, Random forest, Feature importance, CWRU dataset

AMS Subject Classification: 68T05 (Learning and adaptive systems), 62P30 (Applications of statistics to industrial engineering), 70Q05 (Control of mechanical systems)

[1] Department of Mechanical and Aerospace Engineering, IOE Pulchowk Campus, Tribhuvan University, Lalitpur, Nepal.
[2] Department of Automobile and Mechanical Engineering, IOE Thapathali Campus, Tribhuvan University, Kathmandu, Nepal. *E-mail: chhantyalbikki@tcioe.edu.np
* Corresponding author

Manuscript received: 24 March 2026; Revised: 15 June 2026; Accepted: 10 July 2026.

1. Introduction

Rotating machinery serves as the backbone of modern industrial infrastructure in Nepal and worldwide. Faults within bearings result in detrimental component damage and financial losses arising from equipment repair and replacement. Traditionally, experienced maintenance technicians manually used vibration equipment to test and analyze bearing components in order to assess component health. This process requires significant experience to diagnose accurately, especially given the increasingly complicated machines used in industry. Owing to advancements in sensor technology and rising machine complexity, machines are increasingly monitored in a fully automated fashion using machine learning (ML) classification programs that diagnose operating condition from real-time vibration signals.

The principal challenge in using machine learning for fault diagnosis is converting unstructured and dynamic vibration time series into discernible, meaningful features. Rolling element bearings are used to support the rotating shaft. Fault detection and characterization in rolling element bearings occur through periodic spikes in vibration that result from the impacts the rolling elements make upon the affected area of either the inner race, outer race or ball of the bearing. During each revolution, if a rolling element strikes any section of the affected area, a shock-like pulse is created that can be clearly seen in the time-domain signal, especially for larger faults. While an expert would examine this signal to diagnose bearing defects on a spectrogram plot, a machine learning algorithm, if trained effectively, can do the same through a variety of statistical measures derived from the same signal, such as RMS or kurtosis [1].

The phenomenon of detecting a bearing fault is closely intertwined with rotor dynamics and the structural condition of rotating shafts. In preceding analytical work, researchers analyzed the effect of slant cracks in rotor shafts, specifically the coupled interaction of the axial and bending components of vibration, upon rotor shaft vibration signatures [2]. A slant crack creates a variable "breathing" phenomenon, in turn altering the rotor shaft's stiffness matrix. The other major source of shaft vibration in rotating machinery is a defect in the bearing components. Localized defects create a non-linear stiffness characteristic around the bearing and therefore affect the shaft's response during operation [3]. Thus, to produce accurate ML models for bearing defect classification, it is vital that the engineers creating the models have a foundational understanding of the mechanical characteristics and vibration behaviors described above. In this regard, the present paper builds on established foundations within mechanical engineering by extending shaft crack analysis and demonstrating how it can be utilized in a statistical feature-based classification system to identify defects arising in rotating shafts as a result of bearing fatigue over time.

The automated fault diagnostic pipeline established in this study is derived from data gathered by the Case Western Reserve University (CWRU) Bearing Data Center, specifically the 12k drive-end ball bearing data, and shows the viability of time-domain statistical features in training a random forest classifier. Furthermore, this study uses feature importance calculation as a mechanism to address the "black box" concern surrounding modern artificial intelligence techniques, thus linking the output of the classifier to accepted principles of mechanical engineering that can be applied by any maintenance technician without requiring an understanding of artificial intelligence.

The specific contributions of this paper are as follows:

- A systematic evaluation of time-domain statistical features (RMS, kurtosis, peak value and crest factor) for multi-class bearing fault classification on the CWRU 12k drive-end dataset.
- A feature importance analysis within the random forest framework to produce a physically interpretable diagnostic model, bridging the "black-box" nature of artificial intelligence with established mechanical engineering principles so that the model's decisions are transparent to maintenance engineers.

- A demonstration of the practical viability of this approach for real-time condition monitoring in resource-constrained industrial environments such as Nepal's hydropower sector, where computationally efficient and interpretable diagnostics are prioritized over deep learning complexity.

2. Methodology

The research follows a methodological pathway comprising four key steps: acquisition of the experimental data, signal processing by partitioning the data into time segments (windowing), extraction of relevant statistical features from those segments, and classification into fault categories using an ensemble learning approach. The overall pipeline is summarized in Fig. 1.

Data Acquisition (CWRU 12k DE) → Windowing (1200 samples, 0.1 s) → Feature Extraction (RMS, Kurtosis, Peak, Crest Factor) → Random Forest Classification → Fault Label + Feature Importance

FIGURE 1. Proposed diagnostic pipeline for bearing faults.

2.1. **Data Acquisition.** The experimental data were obtained from the CWRU Bearing Data Center database [4]. The test setup includes a 2 HP Reliance Electric motor coupled to an encoder/torque transducer and connected to a power meter that logs voltage and current. Vibration data are obtained from accelerometers installed at the drive end (DE) of the motor.

Following common practice in the bearing fault detection literature, the 12,000 Hz (12 kHz) sampling rate dataset was chosen for this study so that the outcomes of the methodology can be directly compared against previous research [5]. The data collected included readings for the bearing in normal operation (normal baseline), inner race (IR) fault, ball (BL) fault and outer race (OR) fault. The IR, ball and OR faults are further classified as having a fault diameter of either 0.007 in or 0.014 in. Faulty bearings in these categories, all single-point faults, were induced via electrodischarge machining (EDM). For the outer race defects specifically, faults were introduced at the six o'clock position. For the normal baseline, multiple operating load levels were used, including 0, 1, 2 and 3 HP. The composition of the dataset is summarized in Table 1.

TABLE 1. Summary of the CWRU 12k Drive End Dataset and Sample Distribution

Category	Fault dia. (in)	Motor load (HP)	RPM (approx.)	Windows (samples)
Normal baseline	N/A	0, 1, 2, 3	1797–1730	∼400
Inner race fault	0.007, 0.014	0	1797	∼200
Ball fault	0.007, 0.014	0	1797	∼200
Outer race fault	0.007, 0.014	0	1797	∼200
Total dataset				∼1000

2.2. **Signal Preprocessing.** The raw data extracted from the CWRU dataset consist of continuous and complex vibration signals generated during motor rotation. Machine learning algorithms require this time-dependent information to be presented in a more tractable form. Accordingly, a segmentation (windowing) process was applied to transform the continuous vibration streams into smaller, discrete samples. The signals, acquired at sampling frequency f_s, were divided into separate samples of 1200 data points each. Given the sampling rate of 12 kHz, each sample covers a 0.1 s interval, as expressed in (1):

$$T_w = \frac{N_w}{f_s} = \frac{1200}{12000} = 0.1 \text{ s}, \tag{1}$$

where T_w is the window duration and N_w is the number of samples per window. This window length was chosen to strike a compromise between capturing sufficient data for statistical significance and keeping computational cost low. For a motor running at approximately 1797 RPM, a 0.1 s interval corresponds to nearly three shaft rotations, capturing rolling element impacts multiple times within a given sample. As a result, the original limited number of *.mat* files was transformed into a data bank of 10,553 labelled bearing vibration examples spanning the fault and normal classifications, sufficient for a machine learning model to learn meaningful patterns [6].

2.3. **Feature Extraction.** Four key statistical features were computed from each 1200-sample, 0.1 s window. These features represent reduced dimensions that nonetheless encode crucial diagnostic information about the health of the rolling element bearings. They were selected for their known sensitivity to the transient, impulsive and energetic characteristics associated with the vibration produced during component wear. The extracted features are as follows.

2.3.1. *Root Mean Square.* The RMS is computed as the square root of the average of the squares of the signal, as given in (2):

$$x_{\text{RMS}} = \sqrt{\frac{1}{N}\sum_{i=1}^{N} x_i^2}, \tag{2}$$

where x_i is the ith sample of the window and N is the number of samples. RMS provides a statistical measure of the magnitude, or overall vibrational energy, of the signal. Faults inherently create shock-like impacts that increase the energetic content of the signal, making RMS a primary indicator of defect magnitude. The increase in RMS with increasing defect size, observed here from 0.007 in to 0.014 in, results from higher overall kinetic impact energy during rolling element passage.

2.3.2. *Kurtosis.* This statistic measures the degree to which a distribution is prone to peaks, and is defined in (3):

$$K = \frac{\frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{x})^4}{\left[\frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{x})^2\right]^2}, \tag{3}$$

where $\bar{x}$ is the mean of the window. In signal processing, kurtosis is commonly used to detect "spikiness" indicative of shock events or impulses related to mechanical defects such as cracked elements or surface imperfections. A perfectly Gaussian distribution has a kurtosis of 3. A bearing in good working condition exhibits a kurtosis value close to 2.9, while defective bearings typically exhibit higher values as a consequence of the impulsive, non-Gaussian nature of the impacts.

2.3.3. *Peak Value.* This feature measures the largest absolute amplitude of the signal within a 1200-sample interval, as defined in (4):

$$x_{\text{peak}} = \max_{1 \le i \le N} |x_i|. \tag{4}$$

It directly captures the strongest shock pulse recorded within each observed 0.1 s portion of the rotating shaft's motion.

2.3.4. *Crest Factor.* The crest factor is calculated as the ratio of the peak value to the RMS of the signal, as given in (5):

$$\mathrm{CF} = \frac{x_{\mathrm{peak}}}{x_{\mathrm{RMS}}}. \tag{5}$$

It is a dimensionless parameter used to identify the impulsivity or sharpness of the signal. High crest factors can indicate localized defects such as cracks before the damage extends into a widespread surface imperfection.

Representative values of the four extracted features for each bearing condition are presented in Table 2.

Table 2. Extracted Statistical Features for Different Bearing Conditions

Bearing condition	Fault dia. (in)	RMS (g)	Kurtosis	Peak value	Crest factor
Normal baseline	N/A	∼0.07	∼2.9	∼0.2	∼3.5
Inner race fault	0.007	∼0.29	∼5.4	∼1.4	∼4.9
Inner race fault	0.014	∼0.19	∼21.7	∼1.6	∼8.1
Ball fault	0.007	∼0.13	∼2.9	∼0.5	∼3.4
Ball fault	0.014	∼0.14	∼7.3	∼0.8	∼5.4
Outer race fault	0.007	∼0.66	∼7.6	∼3.2	∼4.8
Outer race fault	0.014	∼0.10	∼3.0	∼0.35	∼3.4

2.4. **Classifier Configuration.** Following the extraction of statistical features, the samples were labelled according to health status: normal, inner race fault, outer race fault and ball fault. To emulate testing under real-world industrial conditions, the finalized feature matrix was divided into a 70% training set and a 30% test set. This separation prevents data contamination, which could otherwise lead to exaggerated estimates of classifier accuracy [7]. A random forest classifier was employed as the ensemble method of choice. It was selected for its resistance to overfitting, its ability to handle feature-rich data effectively, and its capacity to compute feature importances that aid interpretability. The forest was assembled from 100 decision trees to ensure reliable predictions. The implementation used the scikit-learn library [8].

3. Results and Discussion

This research produced significant outcomes regarding the automated diagnosis of bearing faults through the analysis of vibration signals and the assessment of random forest classification performance.

3.1. **Preliminary Data Visualization.** Initial inspection of the time-domain signals provided clear evidence of mechanical distress in faulty bearings, as shown in Fig. 2.

Two observations follow from Fig. 2. First, regarding the steady-state baseline, normal and healthy bearing vibration signals are predominantly smooth, low-amplitude waveforms with amplitude generally within $\pm 0.2g$. Second, regarding the impulsive fault signature, outer race faults (0.007 in) display distinctly patterned spikes exceeding the range of a healthy signal, reaching up to $3g$.

3.2. **Classification Performance.** The random forest classifier exhibits strong diagnostic capability, yielding high precision and recall figures across all fault types, as reported in Table 3.

Based on the evidence presented in Table 3, the method achieved 99% average classification accuracy on unseen data. Notably, the model reached perfect precision (1.00) for the ball (0.007 in), inner race (0.007 in), normal and outer race (0.007 in) categories. The corresponding confusion matrix is presented in Fig. 3, which indicates that almost all samples were categorized correctly,

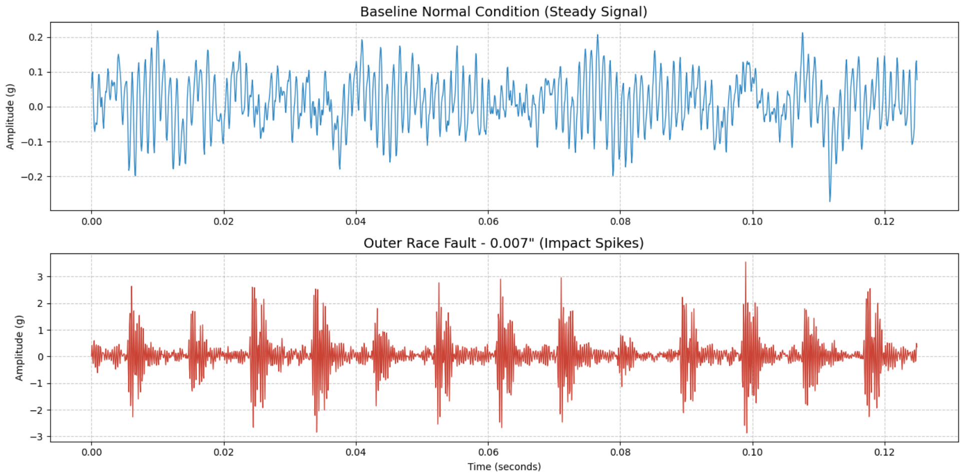

FIGURE 2. Time-domain vibration signals for healthy and faulty (0.007 in) bearing conditions.

TABLE 3. Classification Performance Metrics for the Random Forest Model

Fault category	**Precision**	**Recall**	**F1-score**	**Support**
Normal	1.00	1.00	1.00	416
Inner race (0.007 in)	1.00	1.00	1.00	31
Inner race (0.014 in)	0.97	0.95	0.96	39
Ball (0.007 in)	1.00	1.00	1.00	27
Ball (0.014 in)	0.89	0.86	0.87	28
Outer race (0.007 in)	1.00	1.00	1.00	28
Outer race (0.014 in)	0.97	1.00	0.99	37
Accuracy			**0.99**	**606**

with the limited misclassification that did occur arising predominantly between the ball (0.014 in) and inner race (0.014 in) categories.

Fig. 3 illustrates the nearly flawless classification performance of the random forest model for the normal baseline and the 0.007 in faults. Importantly, not a single false negative was recorded in which a faulty bearing was declared healthy, a critical characteristic for the industrial domain, where false negatives often have severe repercussions for the operation of the bearing and ultimately the machine. Confusion between the ball (0.014 in) and inner race (0.014 in) categories was observed. As the fault diameter becomes larger, the impacts generated by a ball rolling over the race path begin to share similar shapes in the time domain, creating confusion between these types. Even with such minor differences between fault characteristics, the diagnostic accuracy remained above that of typical thresholding schemes, owing to the feature selection adopted in this study.

3.3. **Feature Importance and Interpretability.** To understand how machine learning models "see" faults compared with traditional mechanical engineering practice, the importance of individual statistical features within the classification was examined. The result is presented in Fig. 4.

As shown in Fig. 4, the RMS feature ranks highest and is responsible for over 50% of the model's predictions. According to mechanical vibration principles, the presence of a defect causes an imbalance, which in turn leads to greater energy in the vibration signal. Following RMS, peak value accounts for around 28% of decisions, trailed by kurtosis at approximately 15%. Kurtosis and peak value are effective at capturing the impulsive and "spiky" vibration caused by bearing

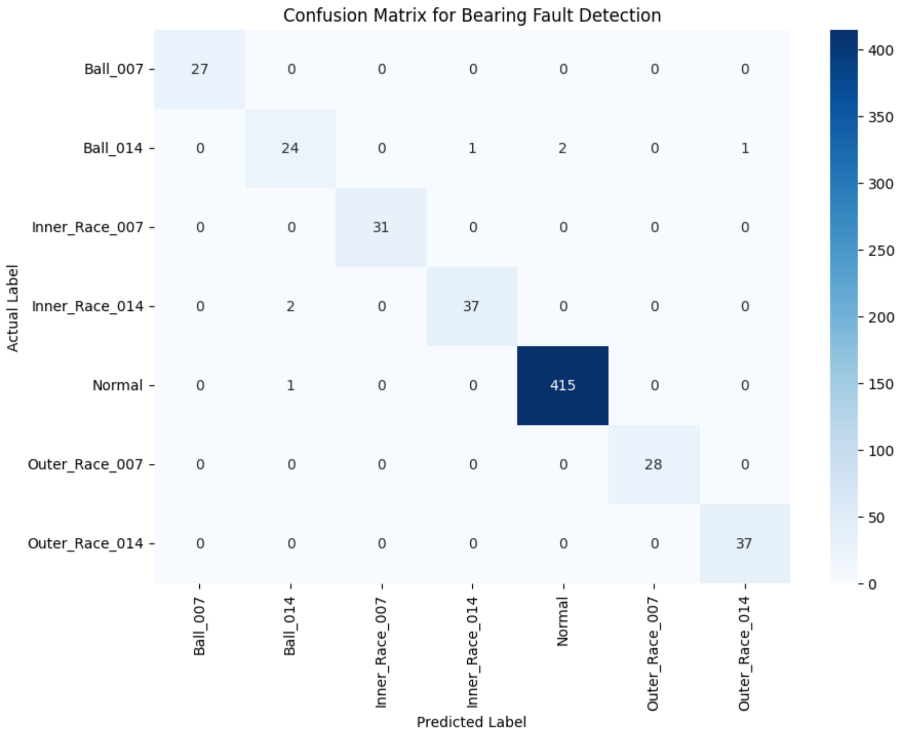

Figure 3. Confusion matrix of the random forest classifier on the testing set.

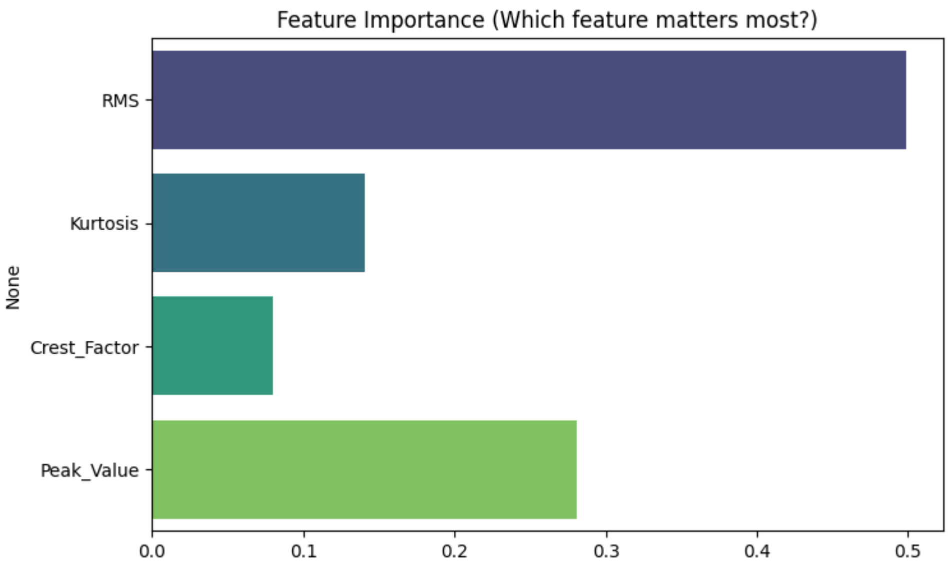

Figure 4. Relative importance of statistical features in the classification process.

defects. The crest factor, at approximately 7%, also indicates shock events during a fault but is less significant than the other selected features within this particular ensemble structure.

3.4. **Comparison to Traditional Rotor Dynamics.** The significance of the energy-based RMS feature and the impulse-based peak feature resembles findings in previous analytical research on slant cracks in shafts. Just as slant crack characteristics induce variations in rotor system stiffness due to repetitive contact, localized faults on a race or in a ball also result in recurring impacts, which are detected, analyzed and interpreted by the random forest model through multiple dimensionally reduced statistical indicators present in the vibration signal. Furthermore, the 99% classification accuracy achieved on the CWRU 12k drive-end dataset is consistent with published benchmarks for time-domain feature-based methods. For instance, classical spectral methods have been reported to achieve accuracies in the range of 90% to 97% on the same dataset [5], while unsupervised feature learning approaches have demonstrated accuracies reaching approximately 98% [6].

4. Conclusion

This study develops an intelligent fault diagnosis framework for rolling element bearings based on time-domain statistical features and the random forest classifier, leveraging data from the CWRU 12k drive-end bearing fault dataset to bridge raw vibration measurements and classified bearing health.

The following conclusions are derived from this study. Time-domain statistical features, in particular RMS and peak value, enable reliable detection of a wide array of fault categories in mechanical machinery, with RMS alone accounting for approximately 50% of the decisions made by the model. The random forest model produced an average accuracy of 99% on unseen data, and the classifier produced a recall of 1.00 for normal baseline bearings, meaning that zero false negatives were recorded in identifying a healthy bearing. Through feature importance analysis and confusion matrices, the internal workings of the machine learning algorithm are demystified; these parameters help translate artificial intelligence, which can appear opaque or "black box" in nature, into the familiar mechanics and engineering principles applied by maintenance personnel. Finally, windowing and feature selection offer a computationally efficient methodology suitable for real-time fault detection in rotating machinery, a framework that will be beneficial to industries such as the power sector in developing countries like Nepal.

In future work, the authors intend to employ frequency-domain features and deep learning architectures to detect and diagnose early-stage faults.

Acknowledgment

The authors would like to thank the Case Western Reserve University Bearing Data Center for providing public access to the bearing vibration dataset used in this study.

References

[1] S. S. Rao, *Mechanical Vibrations*, 6th ed. Hoboken, NJ, USA: Pearson, 2018.
[2] M. Adhikari and P. R. Shakya, "Slant crack detection in shaft using analytical vibration methods," in *Proc. IOE Graduate Conf.*, vol. 6, 2019, pp. 521–526. [Online]. Available: https://conference.ioe.edu.np/publications/ioegc2019-summer/IOEGC-2019-Summer-091.pdf
[3] Y. Jiang, W. Huang, J. Luo, and W. Wang, "An improved dynamic model of defective bearings considering the three-dimensional geometric relationship between the rolling element and defect area," *Mech. Syst. Signal Process.*, vol. 129, pp. 694–716, Aug. 2019.
[4] Case Western Reserve University Bearing Data Center, "12k drive end bearing fault data," [Online]. Available: https://engineering.case.edu/bearingdatacenter, accessed: Jul. 24, 2026.
[5] W. A. Smith and R. B. Randall, "Rolling element bearing diagnostics using the Case Western Reserve University data: A benchmark study," *Mech. Syst. Signal Process.*, vol. 64–65, pp. 100–131, Dec. 2015.
[6] Y. Lei, F. Jia, J. Lin, S. Xing, and S. X. Ding, "An intelligent fault diagnosis method using unsupervised feature learning towards mechanical big data," *IEEE Trans. Ind. Electron.*, vol. 63, no. 5, pp. 3137–3147, May 2016.
[7] A. Géron, *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems*, 2nd ed. Sebastopol, CA, USA: O'Reilly Media, 2019.

[8] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, "Scikit-learn: Machine learning in Python," *J. Mach. Learn. Res.*, vol. 12, pp. 2825–2830, Oct. 2011.