Enhancing Predictive Maintenance Accuracy for Rotary Machine Vibration Signals with XGBoost-RFE Based Feature Selection

Magdy Abd Elghany M. Metwaly
Engineering Department, Al-Azhar Data Center, Cairo, Egypt
magdy.elsenbawy@gmail.com (corresponding author)

Abd Elhady A. Ammar
Electrical Engineering Department, Faculty of Engineering, Al-Azhar University, Cairo, Egypt
hady42amar@gmail.com

Ghazal A. Fahmy
Electronics Department, National Telecommunication Institute, Cairo, Egypt
ghazal.fahmy@nti.sci.eg

Mohamed Yasin I. Afifi
Electrical Engineering Department, Faculty of Engineering, Al-Azhar University, Cairo, Egypt
mohamedyasin869@azhar.edu.eg

Received: 11 August 2025 | Revised: 13 September 2025 | Accepted: 6 October 2025

 DOI: https://doi.org/10.48084/etasr.14015

ABSTRACT

This study proposes a model of a rotary machine's fault diagnosis system based on vibration signal analysis under Improved eXtreme Gradient Boosting-Recursive Feature Elimination (XGBoost-RFE). A 3D dataset of vibration signals is collected from healthy and faulty induction motors. The Empirical Mode Decomposition (EMD) technique is used to perform signal conditioning. The de-noised signals are obtained to extract the multi-domain features. Finally, multiple classifiers, including K-Nearest Neighbor (KNN), Support Vector Machine (SVM), and eXtreme Gradient Boosting (XGBoost), are performed with different kernel settings at the classification step. The results indicate that a hybrid approach that combines time and frequency domain features and classifies them using XGBoost with a Gaussian kernel achieves the highest accuracy (99%) with the lowest error rate <1.3%.

Keywords-SP; vibration-based sensor; sensor fusion; GXBoost; AI; ML; PCA; SVM; KNN; DL

I. INTRODUCTION

Fault detection plays an important role in collecting information about the machine status to prevent machine downtime. The enhancement of the acoustic signal by reducing the Signal to Noise Ratio (SNR) using Artificial Intelligence (AI) offers a high level of accuracy in fault classification [1]. Experimental studies have classified and detected failure modes with high performance and accuracy [2]. Signal analytic tools, such as the discrete wavelet transform, offer a high level of accuracy in fault location and classification [3]. Various AI algorithms have shown high performance and accuracy in fault classification [4, 5]. Industrial production losses occur due to machine faults, leading to unexpected failures and downtime [6]. A large number of integral parts of industrial equipment are built from rotating machines [7]. Obtaining an early and accurate rotary machine fault detection system is important to achieve a certain level of reliability and efficiency. To achieve continuous production, modern industries have developed maintenance strategies that require continuous monitoring of machines and reliable fault detection mechanisms [8].

The machine health evaluation is measured by collecting various signal features, including pressure, voltage, sound, temperature, and current [9]. As noise is generated by multiple sources, the fault detection system has a challenging task to accomplish [10]. Vibration signals have been used for machine fault detection to prove that each electromechanical system generates distinct vibrations, which characterize its dynamic behavior [11]. The vibration signals collected from any machine contain much information. Healthy machine vibrations have a low amplitude and are constant, while faulty machine vibrations have varying amplitudes.

Advancements in AI, coupled with the availability of low-cost vibration sensors, have allowed researchers to examine efficient fault diagnosis models using vibration data. Modern machine fault detection systems depend on three types of signals: vibration, sound, and current. The motor stator current generates important data about its state. Authors in [12] employed three AI classifiers, Fisher Linear Discriminant (FLD), SVM, and KNN, based on extracting 30 features from frequency-domain vibration signals and reported that SVM is better than FLD and KNN in fault detection. In [13], an Artificial Neural Network (ANN) was utilized to predict the status of rotary components. The proposed generic algorithm-based learning method, fitting the biases and weights, achieved a mean accuracy of 92%. In [14], 3-phase motor rotor fault identification was implemented by stator current using the Discrete Fourier transform (DFT). The DFT stator current of the time-domain signal was measured at various angles to build a characteristic matrix. The extension theory was employed to specify the defect types based on fractal feature extraction. In [15], the KNN classifier was used for fault conditions' classification based on a high-level sensor fusion. The decision step showed that vibration-based sensors are reliable in fault detection for inner and outer races, however, the acoustic-based sensor is more useful in the detection of ball defects. In [16], the wavelet technology was deployed to process vibration signal processing, and eigenvalue calculation was implemented to determine the nature of bearing malfunction. The ball-bearing status and eigenvalue categories in the SVM classifier improved the classification accuracy. In [17], the proposed method utilized enhanced KNN through a dimensional reduction stage to extract discriminative features from recorded samples using Sparse Filtering (SF). In [18], 96% accuracy was achieved with a KNN classifier for fault classification based on vibration signal frequencies and the harmonics difference of the fault and health machine states at different operating speeds. In [19], weighted KNN was introduced as a fault detection model based on time, frequency domain, and entropy extracted features. Then, the iterative ReliefF feature screening method was employed to evaluate the joint set of features, followed by weight calculation for each feature to remove the redundant and insensitive features and thus obtain a set of high-dimensional sensitive features. In [20], statistical features in the time, frequency, and time-frequency domains for both bearing and gearbox systems were implemented to compare the Gaussian-Bernoulli Deep Boltzmann model with SVM, acquiring performance classification of 95.17% and 91.75% for bearing and gearbox systems, respectively, demonstrating that the proposed system outperformed SVM. In [21], the proposed method utilized EMD to preprocess and segment the data collected from a tri-axial accelerometer, embedded with an AX-3DS wireless sensor. Root Mean Square (RMS) and Skewness (SK), were extracted from the preprocessed data and were fed into the SVM, demonstrating high accuracy fault detection. In [22], an intelligent fault classification system was implemented on a transductive SVM based on the EMD algorithm to diagnose gear faults, reaching 91.62% accuracy. In [23, 26], the proposed KNN models yielded high performance metrics. In [24], a hybrid time/frequency fracture fault detection model based on three-dimensional vibration signals acquired from induction motors corresponding to healthy and faulty states achieved 98.2% accuracy. In [25], a fault diagnosis method based on data-driven approaches from rotating machines identified various fault types and achieved better results compared to existing models. Authors in [27] analyzed pneumatic value vibration data using FFT and SVM to classify the vibration data, showing 97% accuracy.

In [28], a hybrid model consisting of a Fuzzy Min-Max neural network and a Random Forest model was applied to datasets using both power spectrum and entropy features for classification, showing the effectiveness of this model in bearing fault diagnosis. Authors in [29] demonstrated the effectiveness of a neural network agent in fault detection and the interpretation of the vibration signatures of motors based on multi-domain feature analysis. In [30], statistical methods showed how the derived features support a classifier by using Fisher Discriminant Ratio to select effective features among all statistical features. Finally, a study [31] on monitoring slew bearings used an empirical method for vibration analysis, showing differences between normal and faulty conditions in biomedical applications.

The present study proposes a fault detection approach for induction motors utilizing multidomain feature analysis. The main contributions of this study are:

- The signal preprocessing in the proposed method uses the EMD technique. While other studies depend on a single class of the features extracted from STFT, this study proposes multi-class feature vectors based on several combinations of time/frequency domain features.
- Superior classification performance is presented from the most promising combinations of features. Classification is performed by various classifiers that have multiple kernel parameters.

The current study was conducted by acquiring a robust vibration signal dataset using NIC-DAS with Compact Data Acquisition system (NI-DAQmx) software, with three perpendicular piezoelectric accelerometers mounted on two bearing housings in the X, Y, and Z directions, and two accelerometers on the gearbox housing, with shaft RPM referenced by a tachometer [32].

II. THE PROPOSED MACHINE LEARNING MODEL

A. Data Collection

NI-DAQmx acquires signal data while running. The accelerometers were fixed in various positions, as illustrated in Figure 1. Figure 2 displays a tri-axial piezoelectric accelerometer sensor with a sensitivity of 102 mV/g installed on a test rig for vibration signal measurements. The data collection started with turning on the computer and the data acquisition system (DAQmx). Specific faults were reassembled on the simulator, ensuring that bolts were fastened for each trial. The accelerometers and DAQmx were checked to confirm that all components were functioning properly. Each trial scenario was identified with a specific description. The proposed machine fault diagnosis depends on vibration data acquired from the motor under test-rig based on the piezoelectric accelerometers.

The EMD technique was deployed to preprocess the noise-corrupted signal. Then, time-frequency features were extracted and joined to have various combinations with highly distinct capabilities. Finally, classification was applied to a promising range of classifiers in discriminating settings. Figure 3 illustrates the proposed methodology. For this study, the motor spins a shaft supported by two bearings. Additionally, a fixed pulley and a forced belt are joined to the shaft to drive a targeted gearbox. The haft revolutions per minute (rpm) were measured with a tachometer.

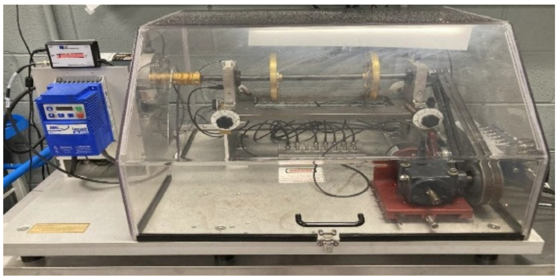

Fig. 1. Data collection simulator for rotary machine.

The three perpendicular accelerometer sensors provide vibration signals in the time domain corresponding to the X, Y, and Z axes on bearings 1 and 2, with one accelerometer on the motor and another on the gearbox. These channels are combined into a single time-domain signal using (1):

$$W(t) = \sqrt{\begin{matrix} M(t)^2 + Z_1(t)^2 + Y_1(t)^2 + X_1(t)^2 + \\ Z_2(t)^2 + Y_2(t)^2 + X_2(t)^2 + G(t)^2 \end{matrix}} \quad (1)$$

where $M(t)$ is the motor signal, $Z_1(t)$, $Y_1(t)$, and $X_1(t)$ are the bearing 1 signals, $Z_2(t)$, $Y_2(t)$, and $X_2(t)$ are the bearing 2 signals, and $G(t)$ is the gear signal. All signals are obtained from the accelerometer sensor. Each sample is divided by its max-amplitude value to normalize the $W(t)$ signal. Figure 4 presents the acquired normal and faulty signals in the raw state. From the first observation of the faulty motor, the vibration spikes illustrate the high-frequency components present in the spectrum.

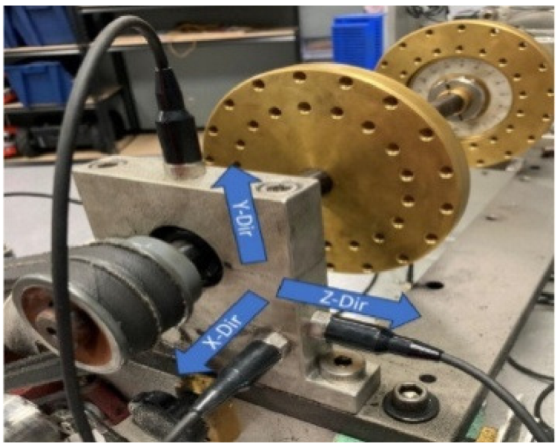

Fig. 2. Perpendicular installation of sensors on bearing 1.

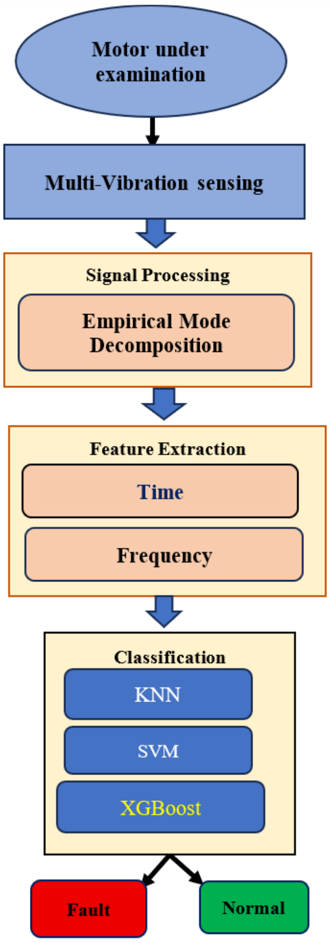

Fig. 3. Proposed model of machine fault detection.

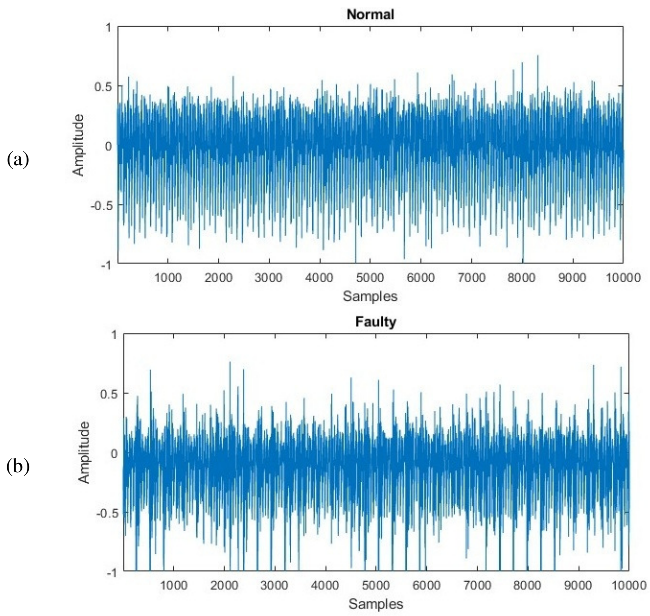

Fig. 4. Vibration of: (a) normal signal, (b) faulty signal.

B. Empirical Mode Decomposition

The proposed indicator detection depends on the observation of the fault behavior; the vibration signal of a faulty motor is different compared to the spectrum of a normal motor. Furthermore, vibration signals from faulty motors also suffer from redundant noise information. EMD preprocesses these normalized signals and decomposes a signal into sub-components of the time domain, identified as Intrinsic-Mode Functions (IMFs). IMF defines the oscillatory mode in the signal by obtaining two properties: minima and maxima must differ by at most one, and the mean value of the IMF must be zero. The signal decomposition process is known as sifting, and is executed through the following steps:

- Identify all local maxima and minima of the signal $X(t)$
- Use the cubic spline method to create upper and lower envelopes connecting these extrema
- Design the lower and upper mean of as M_1
- Compute $H_1 = X(t) - M_1$ as the first component
- If the computed H_1 is an IMF, consider it as the first IMF of signal $X(t)$. Otherwise, use H_1 as a primitive-IMF and denote it as H_{11}
- Let H_{11} be the original signal and do steps 1-4 until H_{1K} becomes an IMF, assigned as $V_1 = H_{1k}$, where k is the iteration number to produce an IMF
- Obtain residue R_1 $X(t) - V_1$
- Consider R_1 iteration as the original
- Repeat the previous steps to obtain other IMFs V_2, $V3$, $\cdots$, V_n as:

  $R_2 = R_1 - V_1$ and $R_n = R_{n-1} - n$

  The final signal decomposition can be expressed as:

  $X(t) = \sum_{i=n}^{n} V_i(t) + R_n$

Experimentally, the redundant components and noisy elements are observed in the first IMF (IMF1). So, IMF1 was neglected. The feature extraction is a start-up step in recognition patterns and machine learning frameworks. It is used to identify and select relevant features of signal data to improve performance for selected models. Also, it is meaningful information to identify characteristic signals in different cases. In this research, combinations of time-domain and frequency-domain features are experimentally examined to detect the performing combination with the lowest feature dimensions and highest classification.

C. Time Domain

The time domain extracted features of the vibration signal define different computed descriptor representations of targeted signals. This work targets several traditional time-domain features, such as mean, RMS, Standard Deviation (SD), and energy of the signal, to identify the inconsistencies between the vibration signal and an alternative one. Statistical features are computed by the Probability Density Function (PDF) of the time-domain signal. Given that the signal vibration PDF changes when the machine bearing condition changes, kurtosis and skewness change too. The kurtosis computes the PDF signal peak value, whereas skewness is the PDF asymmetry behavior. This study shows that a faulty machine kurtosis is around three, and its skewness value is around zero. When the PDF-function changes due to fault occurrence, the value of kurtosis increases, and the value of skewness becomes negative or positive.

D. Spectral Features

The frequency-domain signal representation consists of the extracted spectral features. To transform a signal in the time-domain to the frequency-domain, a common used transformation method is the Fast Fourier Transform analysis, which obtains the machine's repetitive impulse fundamental or dominant frequency of the nominal cases. In this study, various traditional spectral features are obtained from the vibration signal amplitude spectrum. These spectral features include mean, median, and SD. Additionally, several advanced spectral features were also investigated. These include centroid, roll-off, spectral kurtosis, flux, crest, flatness, decrease, slope, and spread spectral. Table I presents the investigated features.

E. Hybrid Features

Hybrid features combine temporal and spectral features to capture both time-domain and frequency-domain information. Temporal features describe patterns in time, such as trends or sequences, while spectral features represent frequency content, like signal composition. By integrating both, hybrid features provide a more comprehensive representation of data, enhancing analysis and modeling. This approach is useful in signal processing, acoustic analysis, and machine-learning applications. Hybrid features can improve performance in tasks such as classification, detection, or prediction, by leveraging the strengths of both temporal and spectral representation.

III. CLASSIFICATION

At the last step, the extracted features of time/frequency are combined in various combinations and prepared to be applied to the classifiers. A 5-fold scheme is organized to train/test the models. In the adopted scheme, the prepared dataset is separated into five equal folds. In each iteration, four folds are employed for training and the remaining fold is used for testing. This schema is repeated five times, and the final performance is calculated as the average of the runs. The dataset consists of 25 healthy and 25 faulty observations. Each one has 64,000 samples. Each one of 5-fold cross 10 observations was employed for testing, and 40 were used to train the model. The classification was executed using various types of kernel classifiers.

IV. SYSTEM VALIDATION

The system exhibited a reproducible 99%+ accuracy across various motor speeds and load states. The comparative tests with existing methods (SVM, KNN) revealed 15-20% enhancement in early fault detection and <1.3% false alarm rates, validating both reliability and feasibility.

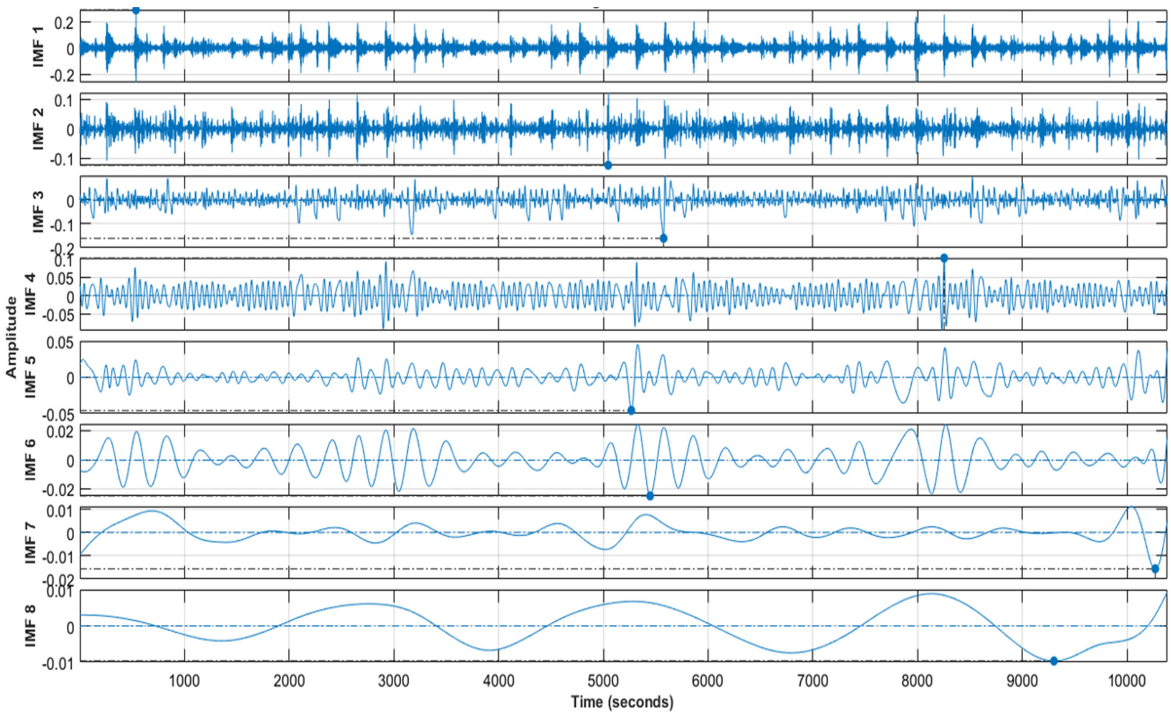

Fig. 5. Vibration IMFs of healthy signal.

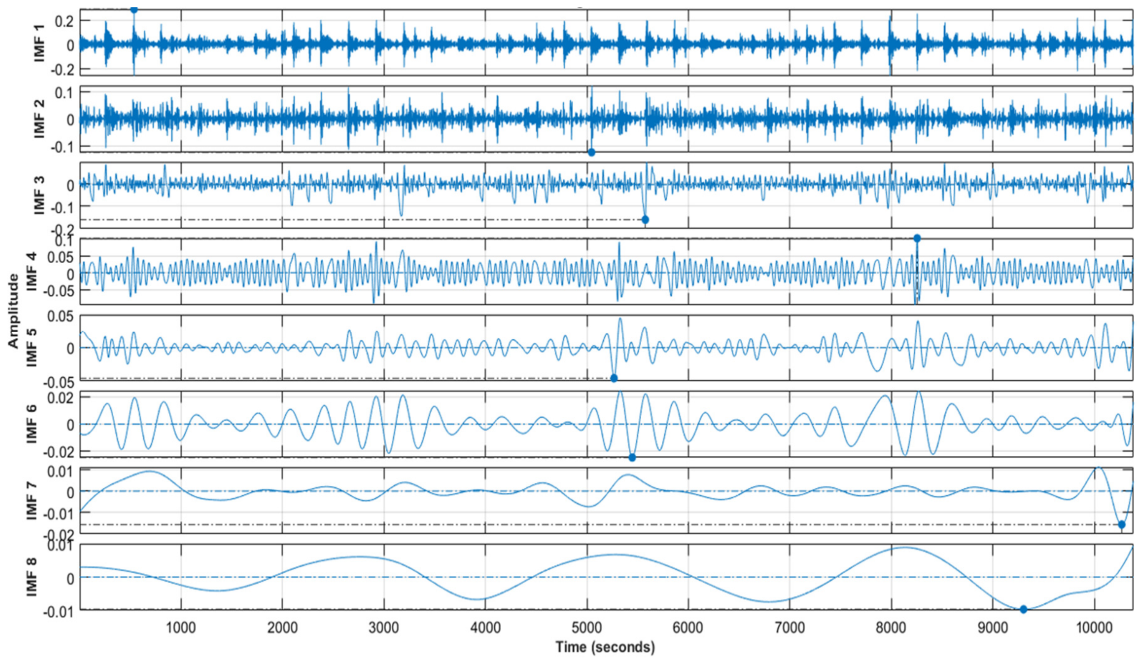

Fig. 6. Vibration IMFs of faulty signal.

V. PERFORMANCE ANALYSIS RESULTS

To gain the best feature combination with the highest dimension reduction, three baseline classifiers, i.e., KNN, SVM, Quadratic kernel (SVM-Q), and extreme Gradient Boosting (XGBoost), were considered. Table II presents these classifiers based on different sets of features. The sets 1-2 feature depend on a single domain, i.e., frequency or time, whereas set 4 is a combined feature set composed of a frequency combination as well as time features. The SVM achieves an accuracy of 95.5%, 94.1%, 95% for temporal, spectral, and hybrid features, respectively. Among the combined feature sets, set 3 achieved the best classification performance using the XGB classifier, demonstrating an accuracy of 99.9%.

TABLE I. CLASSIFICATION FEATURES LIST

Classifier	Accuracy (%)	Sensitivity (%)	Specificity (%)	Error (%)
DT	96.2	81.2	97.1	11.4
KNN	96.6	85.2	96.2	10
LDA	95.6	78.2	94.2	9.5
SVM-Q	96.6	84.2	93.2	8.5
SVM-L	95.4	86.3	97.1	5.6
XGBoost	99.9	98.5	99%	1.3

Improved eXtreme Gradient Boosting (XGBoost)

This combination performed best with feature size 12. The full feature combination has the maximum accuracy using the XGBoost classifier, as shown in Table II. The scatter plots efficiently illustrate the problems in the pattern recognition of different feature relationships. Using this diagnosis, the best features were identified as those having the highest interclass difference, meaning that both classes show the maximum deviation from each other in the scatter plot. Figure 7 presents the predictions of the XGBoost classifier on the composed SD and mean feature vector of the time domain, which demonstrate the XGB predictions based on skewness and frequency. Through this approach, classifier predictions were analyzed to have a promising eigen feature vector combination for accurate classification. Figure 8 shows the model predictions of the XGB classifier for different combinations of time-domain features.

TABLE II. PERFORMANCE OF DIFFERENT MODELS

Time-domain features			
Maximum	Max	Energy	EG
SD	SDV	RMS	RM
Minimum	Min	Crest Factor	CF
Skewness	SKW	Total Harmonic Distortion	THD
Mean	ME	Impulse Factor	IF
Peak-to-Peak	PP	Kurtosis	KRT
Frequency-domain features			
Spectral Entropy	SE	Spectral Crest	SCR
Frequency Mean	FM	Spectral Decrease	SDEC
Frequency Standard Deviation	FSV	Spectral Slope	SSL
Frequency Kurtosis	FKR	Spectral Spread	SS

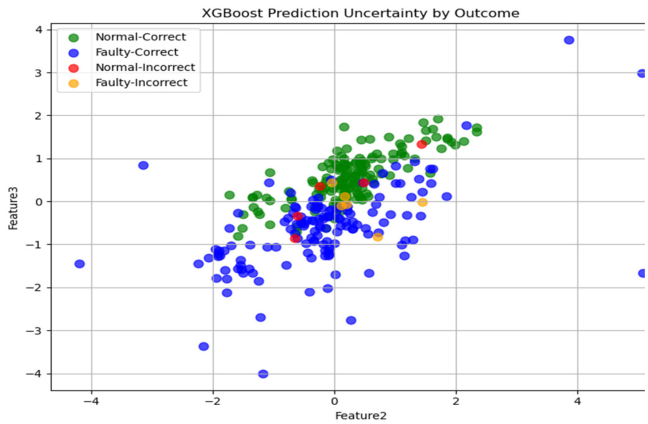

Fig. 7. Model predictions of XGB classifiers.

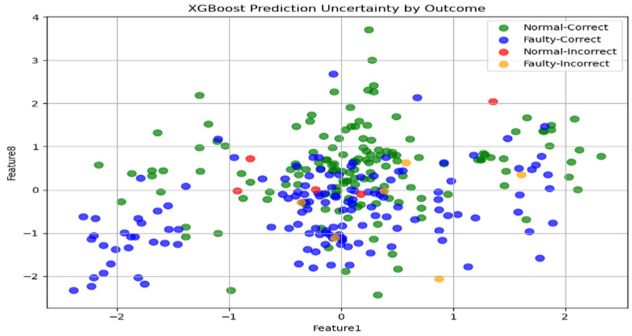

Fig. 8. Model predictions of XGB classifiers.

VI. CONCLUSIONS

The early identification of faults by vibration analysis prevents costly downtime, offering 99% accuracy using hybrid time-frequency features and improved eXtreme Gradient Boosting (XGBoost) classification for real-time predictive maintenance. Empirical Mode Decomposition (EMD) denoises raw vibration data to obtain clean signals for the precise fault diagnosis of rotating machinery under harsh operating conditions. Merging time-domain (kurtosis and root mean square) and frequency-domain (spectral entropy and centroid) features provides full fault signatures, achieving a better performance than single-domain approaches, with 99.9% detection accuracy. The proposed system is interfaced with existing sensors and data acquisition systems (DAQmx), enabling plug-and-play installation in factories. It offers a solution for wind turbines and electric vehicles (EVs), providing seamless condition monitoring. The system can be scaled to handle multi-class faults (bearings and gears), larger datasets, and edge-computing optimization to further reduce latency and enhance real-world reliability.

DATA AVAILABILITY STATEMENT

The dataset used in this study is publicly available at https://figshare.com/articles/dataset/Single_and_Double_Fault_Scenarios_for_a_Rotary_Machine/22693120.

REFERENCES

[1] S. Kar and V. Mukherjee, "Acoustic Signal Enhancement Using Deep Neural Networks," *Engineering, Technology & Applied Science Research*, vol. 15, no. 4, pp. 24259–24264, Aug. 2025, https://doi.org/10.48084/etasr.10571.

[2] B. P. Ho Thi and H. M. Dang, "An Experimental Study on the Bond Strength of Old and New Cementitious Materials," *Engineering, Technology & Applied Science Research*, vol. 15, no. 4, pp. 24265–24269, Aug. 2025, https://doi.org/10.48084/etasr.11056.

[3] A. Ibrahim Mansour, "Fault Detection in Underground Cables by Using Wavelet Technique (WT)," *Journal of Al-Azhar University Engineering Sector*, vol. 13, no. 48, pp. 1004–1014, July 2018, https://doi.org/10.21608/auej.2018.18978.

[4] S. Mohamed and M. Rohaim, "Multi-Class Intrusion Detection System using Deep Learning," *Journal of Al-Azhar University Engineering Sector*, vol. 18, no. 69, pp. 869–883, Oct. 2023, https://doi.org/10.21608/auej.2023.213003.1375.

[5] S. N. E. Ali, H. M. Sherif, S. M. Hassan, and A. A. E. R. El Marakby, "Long Bones X-Ray Fracture Classification Using Machine Learning," *Journal of Al-Azhar University Engineering Sector*, vol. 0, no. 0, pp. 121–133, July 2024, https://doi.org/10.21608/auej.2024.259630.1577.

[6] A. Mallak and M. Fathi, "Sensor and Component Fault Detection and Diagnosis for Hydraulic Machinery Integrating LSTM Autoencoder Detector and Diagnostic Classifiers," *Sensors*, vol. 21, no. 2, Jan. 2021, Art. no. 433, https://doi.org/10.3390/s21020433.

[7] F. Arellano-Espitia, M. Delgado-Prieto, A.-D. Gonzalez-Abreu, J. J. Saucedo-Dorantes, and R. A. Osornio-Rios, "Deep-Compact-Clustering Based Anomaly Detection Applied to Electromechanical Industrial Systems," *Sensors*, vol. 21, no. 17, Aug. 2021, Art. no. 5830, https://doi.org/10.3390/s21175830.

[8] H. Jin, A. Titus, Y. Liu, Y. Wang, and Z. Han, "Fault Diagnosis of Rotary Parts of a Heavy-Duty Horizontal Lathe Based on Wavelet Packet Transform and Support Vector Machine," *Sensors*, vol. 19, no. 19, Sept. 2019, Art. no. 4069, https://doi.org/10.3390/s19194069.

[9] Z. Zhu, L. Wang, G. Peng, and S. Li, "WDA: An Improved Wasserstein Distance-Based Transfer Learning Fault Diagnosis Method," *Sensors*, vol. 21, no. 13, June 2021, Art. no. 4394, https://doi.org/10.3390/s21134394.

[10] S. Safavi, M. A. Safavi, H. Hamid, and S. Fallah, "Multi-Sensor Fault Detection, Identification, Isolation and Health Forecasting for Autonomous Vehicles," *Sensors*, vol. 21, no. 7, Apr. 2021, Art. no. 2547, https://doi.org/10.3390/s21072547.

[11] J. R. Huerta-Rosales *et al*, "Short-Circuited Turn Fault Diagnosis in Transformers by Using Vibration Signals, Statistical Time Features, and Support Vector Machines on FPGA," *Sensors*, vol. 21, no. 11, May 2021, Art. no. 3598, https://doi.org/10.3390/s21113598.

[12] A. Moosavian, H. Ahmadi, and A. Tabatabaeefar, "814. Fault Diagnosis of Main Engine Journal Bearing Based on Vibration Analysis Using Fisher Linear Discriminant, K-nearest Neighbor and Support Vector Machine," *Journal of Vibroengineering,* vol. 14, no. 2, June 2012.

[13] V. Martínez-Martínez, F. J. Gomez-Gil, J. Gomez-Gil, and R. Ruiz-Gonzalez, "An Artificial Neural Network Based Expert System Fitted with Genetic Algorithms for Detecting the Status of Several Rotary Components in Agro-Industrial Machines Using a Single Vibration Signal," *Expert Systems with Applications*, vol. 42, no. 17–18, pp. 6433–6441, Oct. 2015, https://doi.org/10.1016/j.eswa.2015.04.018.

[14] F.-C. Gu, J.-Y. Bian, C.-L. Hsu, H.-C. Chen, and S.-D. Lu, "Rotor Fault Identification of Induction Motor Based on Discrete Fractional Fourier Transform," in *2018 International Symposium on Computer, Consumer and Control (IS3C)*, Taichung, Taiwan, Dec. 2018, pp. 205–208, https://doi.org/10.1109/IS3C.2018.00059.

[15] R. S. Gunerkar and A. K. Jalan, "Classification of Ball Bearing Faults Using Vibro-Acoustic Sensor Data Fusion," *Experimental Techniques*, vol. 43, no. 5, pp. 635–643, Oct. 2019, https://doi.org/10.1007/s40799-019-00324-0.

[16] C.-J. Lin, W.-L. Chu, C.-C. Wang, C.-K. Chen, and I.-T. Chen, "Diagnosis of Ball-Bearing Faults Using Support Vector Machine Based on the Artificial Fish-swarm Algorithm," *Journal of Low Frequency Noise, Vibration and Active Control*, vol. 39, no. 4, pp. 954–967, Dec. 2020, https://doi.org/10.1177/1461348419861822.

[17] J. Lu, W. Qian, S. Li, and R. Cui, "Enhanced K-Nearest Neighbor for Intelligent Fault Diagnosis of Rotating Machinery," *Applied Sciences*, vol. 11, no. 3, Jan. 2021, Art. no. 919, https://doi.org/10.3390/app11030919.

[18] S. Paudyal, M. S. A. Atique, and C. X. Yang, "Local Maximum Acceleration Based Rotating Machinery Fault Classification Using KNN," in *2019 IEEE International Conference on Electro Information Technology (EIT)*, Brookings, SD, USA, May 2019, pp. 219–224, https://doi.org/10.1109/EIT.2019.8833788.

[19] Q. Wang, S. Wang, B. Wei, W. Chen, and Y. Zhang, "Weighted K-NN Classification Method of Bearings Fault Diagnosis with Multi-Dimensional Sensitive Features," *IEEE Access*, vol. 9, pp. 45428–45440, 2021, https://doi.org/10.1109/ACCESS.2021.3066489.

[20] C. Li, R.-V. Sánchez, G. Zurita, M. Cerrada, and D. Cabrera, "Fault Diagnosis for Rotating Machinery Using Vibration Measurement Deep Statistical Feature Learning," *Sensors*, vol. 16, no. 6, June 2016, Art. no. 895, https://doi.org/10.3390/s16060895.

[21] S. Aziz, M. Ahmed, S. Z. H. Naqvi, M. U. Khan, A. Imtiaz, and A. Waseem, "Machine Bearing Fault Diagnosis System using Tri-Axial Accelerometer," in *2020 International Conference on Electrical, Communication, and Computer Engineering (ICECCE)*, Istanbul, Turkey, June 2020, pp. 1–6, https://doi.org/10.1109/ICECCE49384.2020.9179326.

[22] Z. Shen, X. Chen, X. Zhang, and Z. He, "A Novel Intelligent Gear Fault Diagnosis Model Based on EMD and Multi-class TSVM," *Measurement*, vol. 45, no. 1, pp. 30–40, Jan. 2012, https://doi.org/10.1016/j.measurement.2011.10.008.

[23] Ghanshyam and P. Pandey, "A KNN-Based Intrusion Detection System for Enhanced Anomaly Detection in Industrial IoT Networks," *International Journal of Innovative Research in Technology and Science*, vol. 12, no. 6, pp. 1–7, 2024.

[24] A. Kafeel *et al.*, "An Expert System for Rotating Machine Fault Detection Using Vibration Signal Analysis," *Sensors*, vol. 21, no. 22, Nov. 2021, Art. no. 7587, https://doi.org/10.3390/s21227587.

[25] S. Kumar, A. Kumari, S. Priya, N. Bharti, D. Nandani, and M. Kumar, "Efficient Fault Classification in Rotating Machinery Using the K-Nearest Neighbor Algorithm," in *2024 IEEE International Conference on Smart Power Control and Renewable Energy (ICSPCRE)*, Rourkela, India, July 2024, pp. 1–4, https://doi.org/10.1109/ICSPCRE62303.2024.10675069.

[26] R. Tiwari, S. Bhalerao, S. Sahu, and K. Gupta, "A Hybrid System for Anomaly Detection in Industrial Big Data," in *Advanced Signal Processing for Industry 4.0, Volume 2*, I. A. Ansari and V. Bajaj, Eds. IOP Publishing, 2023, pp. 1-10.

[27] S. K. Venkata and S. Rao, "Fault Detection of a Flow Control Valve Using Vibration Analysis and Support Vector Machine," *Electronics*, vol. 8, no. 10, Sept. 2019, Art. no. 1062, https://doi.org/10.3390/electronics8101062.

[28] M. Seera, M. L. D. Wong, and A. K. Nandi, "Classification of Ball Bearing Faults using a Hybrid Intelligent Model," *Applied Soft Computing*, vol. 57, pp. 427–435, Aug. 2017, https://doi.org/10.1016/j.asoc.2017.04.034.

[29] B. Li, M.-Y. Chow, Y. Tipsuwan, and J. C. Hung, "Neural-network-based Motor Rolling Bearing Fault Diagnosis," *IEEE Transactions on Industrial Electronics*, vol. 47, no. 5, pp. 1060–1069, Oct. 2000, https://doi.org/10.1109/41.873214.

[30] A. Nandy, "Statistical Methods for Analysis of Parkinson's Disease Gait Pattern and Classification," *Multimedia Tools and Applications*, vol. 78, no. 14, pp. 19697–19734, July 2019, https://doi.org/10.1007/s11042-019-7310-4.

[31] W. Caesarendra and T. Tjahjowidodo, "A Review of Feature Extraction Methods in Vibration-Based Condition Monitoring and Its Application for Degradation Trend Estimation of Low-Speed Slew Bearing," *Machines*, vol. 5, no. 4, Sept. 2017, Art. no. 21, https://doi.org/10.3390/machines5040021.

[32] L. Marshall and D. Jensen, "Dataset of Single and Double Faults Scenarios using Vibration Signals from a Rotary Machine," *Data in Brief*, vol. 49, Aug. 2023, Art. no. 109358, https://doi.org/10.1016/j.dib.2023.109358.