



AI - Driven Predictive Maintenance for Industrial Rotating Machinery Using Vibration Analysis

Dr. Biswajit Datta¹, Rahila Begum Gadi²

¹(Assistant Professor

Department of Computer Application
George College of Management and Science,
Budge Budge Trunk Road,
Chalkmir, Maheshtala,
Kolkata, 700141, India
dattadrabiswajit@gmail.com)

²(Lecturer

Engineering

University of Technology and applied science Al Mussanah, Oman
rahila111@rediffmail.com)

Abstract. Industrial rotary machines experiencing unplanned outages owing to bearing and gear defects cost economies and safety greatly. In this paper, we introduce a hybrid deep learning algorithm for predictive maintenance through the combination of 1-Dimensional Convolutional Neural Networks (1D-CNN) and Bi-directional Long Short-Term Memory (BiLSTM) networks in order to perform diagnosis as well as Remaining Useful Life (RUL) estimation using vibration data. The introduced algorithm uses multi-scale feature learning and temporal dependency modeling in order to detect degradation trends without the need for much feature engineering. Evaluation experiments conducted on the CWRU bearing dataset illustrate high classification accuracy even when the vibration data contains varying amounts of noise and low RUL prediction error. Comparative experiments show the effectiveness of the introduced hybrid model when compared to other models such as SVM, Random Forest, CNN, and LSTM.

Keywords: Predictive Maintenance, Rotating Machinery, Vibration Analysis, Deep Learning, Fault Diagnosis, Remaining Useful Life, CNN-LSTM,

I. Introduction

Rotating machines represent the heart of today's industrial operations such as oil and gas facilities, power plants, production lines, and wind turbines. The components of these machines, namely the bearings, gears, and shafts, are exposed to stresses of mechanical nature and temperature fluctuations together with deterioration of the lubricant that results in fault development, which if not detected leads to the catastrophic malfunction of the machine. Unplanned malfunctions of the equipment represent a significant portion of overall maintenance expenses in an industrial enterprise, and for the petrochemical industry alone, downtime expenses related to the compressor can amount to hundreds of thousands of dollars per day [1]. This fact is one of the reasons for the



move from time- and reactive-based maintenance approaches to predictive maintenance (PdM).

Vibration analysis has become the primary methodology for machine condition monitoring because of its ability to detect early faults, non-invasiveness of measurement, and wealth of diagnostic information. If there is a fault in the bearing, then the rolling elements will repeatedly impact the surface and produce corresponding impulse trains, which will be represented by specific frequency components of the vibration spectra such as Ball Pass Frequency Outer race (BPFO), Ball Pass Frequency Inner race (BPFI), Ball Spin Frequency (BSF), and Fundamental Train Frequency (FTF) [6]. Conventional condition monitoring is based on the interpretation of these signatures by the experts using techniques like envelope demodulation, spectral kurtosis, and cyclostationarity [10]. Nevertheless, expert-based diagnostics cannot be scaled up to the needs of monitoring heterogeneous fleets of machines.

The use of machine learning for vibration-based fault diagnosis has gone through several stages. Initial systems made use of carefully designed statistical features like root mean square, kurtosis, crest factor, and spectral moments as input features for classifier systems such as SVMs and ANN [9]. Even though successful within specific settings, these methods need expertise in domain knowledge for feature construction and are not efficient in cases where conditions during operation differ from training distributions. Deep learning has brought about end-to-end learning of features from either raw data or data that is minimally processed, thus getting rid of the need for feature construction and learning hierarchies that are related to the level of faulting [1]. CNNs have proven to be very effective in feature extraction from small portions of the data, while LSTM networks capture temporal changes of signatures [3].

Despite the aforementioned advancements, there are various issues hindering the industrial implementation of AI-based PdM systems. The first one is that while most research focuses on tackling fault diagnosis and prediction separately, maintenance decision-making needs a solution for both problems. The second problem is the availability of data regarding fault conditions which is quite low since the machines in industrial settings do not fail in a controlled environment. The third issue is lack of interpretability of the models which leads to decreased trust in the results by maintenance engineers.

This paper attempts to solve these problems using a hybrid approach of 1D-CNN-BiLSTM architecture that jointly classifies faults and estimates RUL using raw vibration signals. The contributions are threefold:

- A 1D-CNN-BiLSTM architecture that can detect local faults as well as degradation trends
- A combination loss function that allows for simultaneous training for classification and regression tasks
- Extensive quantitative testing on the baseline models in the CWRU dataset



This paper is organized as follows: in Section 2, related work in vibration-based PdM is presented; in Section 3, methodology is discussed; Section 4 covers experiments and results, and Section 5 is the conclusion.

II. Literature Survey

The field of AI application in vibration-based machinery health monitoring has seen a great deal of change over the last ten years, as a result of the interaction between availability of benchmark datasets, development of deep learning models, and industry's need for automated diagnostic systems. The first contributions in this field laid out the basics of bearing diagnostics by providing a description of fault frequencies of bearing failure modes and their physical meaning [6]. The work done by Tandon and Choudhury on defining the fault frequencies of BPFO, BPFI, BSF, and FTF is still the basis of any modern diagnostic system. It provides a physically meaningful goal for both signal processing and machine learning algorithms.

Machine learning began as a result of the publication of Case Western Reserve University (CWRU) Bearing Data Center data, whereby vibration signals from bearing failures were collected and have subsequently been used as the gold standard for comparisons [6]. The initial applications of ML algorithms utilized hand-crafted statistical features such as RMS, kurtosis, skewness, and crest factor as inputs for the SVM and KNN classification algorithms, resulting in good performance at optimal signal-to-noise ratios. Nevertheless, the method was not successful in generalizing to changes in speed, loading, and noise levels.

Architectures of deep learning models have proved to be very good in learning the discriminative representations from the raw vibration signals. Convolutional Neural Network (CNN) models leverage temporal correlations via the use of learnable filter banks which can learn the fault-related frequency components autonomously without feature engineering [1]. Zhang et al. proposed the Wide First-layer Kernel CNN (WDCNN) for the detection of faults in bearing, where a wide kernel is used in the first layer to filter out high-frequency noises and detect the low-frequency fault components [1]. The temporal evolution of degradation signatures has been modeled using the recurrent network, specifically the Long Short-Term Memory (LSTM) models for the purpose of estimating remaining useful life [3].

Some of the issues identified in literature include the following. First, domain shift due to differences in speeds, weights, and sensor placement between training and operational domains adversely affects the performance of the predictive models [2]. The cross-domain transfer studies consistently demonstrate that performance is degraded when the models trained on one dataset are deployed on the other. Domain adaptation becomes necessary as a result of that issue. Second, labeled fault data is sparse in the industrial setting due to equipment failing to reach the point of failure under controlled

conditions [5]. Anomaly detection based on unsupervised methods, such as autoencoders and One-Class SVM, has proven to be an interesting solution [5]. Third, interpretability of the neural networks remains poor, although explainable AI such as SHAP and feature importance techniques can be applied post hoc [5].

There have been a number of studies that have looked into few-shot learning and meta-learning as techniques to overcome lack of data with frameworks such as Few-Shot MS-1D-CNN that prove to be feasible for deployment in industrial settings [1]. Mechanistic knowledge such as the Paris-Erdogan crack propagation laws is incorporated into the data-driven frameworks with physics-informed neural networks, which enhances not only interpretability but also the extrapolation ability [7]. Transfer learning involves adapting a pre-trained framework using limited samples from the target domain, which performs quite well on gearbox vibration data after adapting just the last few layers [2].

III. Proposed Methodology

System Overview

The suggested approach consists of four stages: data collection and pre-processing, multi-domain feature extraction, deep learning inference, and maintenance decision. Vibration data is divided into overlapping windows, normalized, and possibly transformed into time-frequency domain before inputting into the hybrid 1D CNN-BiLSTM model. The output of the hybrid 1D CNN-BiLSTM model includes fault probabilities and RUL predictions. These two outputs are fused in a maintenance decision module that initiates maintenance action if fault probability surpasses certain value or RUL prediction goes under a certain horizon. Figure 1 depicts the overall system architecture.

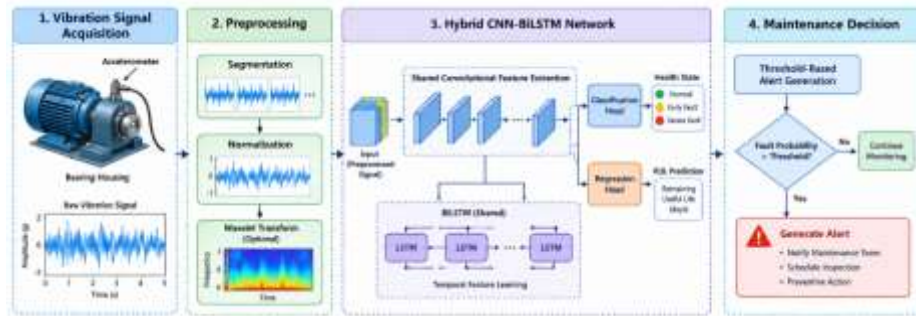


Figure 1: Overall System Architecture.

The figure illustrates the full predictive maintenance pipeline, beginning with vibration signal collection through accelerometers placed at the bearing housings, preprocessing steps (segmentation, normalization, and optionally, wavelet transform), the proposed hybrid CNN-BiLSTM architecture that includes convolutional feature extraction and dual output heads for classification and regression tasks, and finally, the maintenance decision module.



Signal Preprocessing

The vibration signal acquisition takes place at 12 kHz sampling rate. This is according to the specification of the CWRU dataset. The raw signals are broken down into windows containing 2048 samples each, overlapping by half to increase training data size and ensure continuity. The data in each window is normalized to have zero mean and unit variance to ensure there is no variation due to the variation in amplitude as a result of different gains in the sensors and loads applied. For learning features in the frequency domain, the Fast Fourier Transform (FFT) magnitude spectrum of each window is calculated.

Hybrid CNN-BiLSTM Architecture

The key inference engine involves a hybrid network that is based on 1D convolutional layers for feature extraction along with bidirectional LSTM layers to model temporal dependencies. The 1D convolutional front end consists of three 1D convolutional blocks that contain 64, 128, and 256 filters with kernel size equal to 32, 16, and 8 respectively. The stride used is 1, followed by ReLU activation and batch normalization followed by max-pooling of factor 2. This design ensures fault signature extraction at multiple scales ranging from high-frequency impulse signals to low-frequency modulations of advanced degradation.

The output from the convolution is passed on to a two-layer BiLSTM with 128 units in each layer and dropout rate of 0.3. The use of the bi-directional architecture enables processing of the input features in the two temporal directions, which allows capturing the patterns of the degradation that could be non-monotonic or dependent on context. The output of the final time step of the LSTM is then connected to two separate outputs – classification (with 4 classes - normal, inner race fault, outer race fault, ball fault) and regression heads.

Algorithm 1: Hybrid CNN-BiLSTM Training

```
Input: Vibration signal windows X, fault labels Y_class, RUL targets Y_rul
Output: Trained model parameters  $\theta$ 

1: Initialize weights  $\theta$  using He initialization
2: for epoch = 1 to E do
3:   for each batch (X_b, Y_class_b, Y_rul_b) in training data do
4:      $H_{conv} \leftarrow \text{ConvBlocks}(X_b)$   $\triangleright$  3 blocks: Conv1D, BN, ReLU,
MaxPool
5:      $H_{lstm} \leftarrow \text{BiLSTM}(H_{conv})$   $\triangleright$  2 layers, 128 units, dropout=0.3
6:      $P_{class} \leftarrow \text{Softmax}(\text{Dense}(H_{lstm}))$   $\triangleright$  Classification output (4 classes)
7:      $P_{rul} \leftarrow \text{Linear}(\text{Dense}(H_{lstm}))$   $\triangleright$  Regression output (scalar)
8:      $L_{total} = \alpha * L_{CE}(P_{class}, Y_{class\_b}) + (1-\alpha) * L_{MSE}(P_{rul}, Y_{rul\_b})$ 
9:     Update  $\theta$  via Adam optimizer to minimize  $L_{total}$ 
10:   end for
11: end for
12: return  $\theta$ 
```

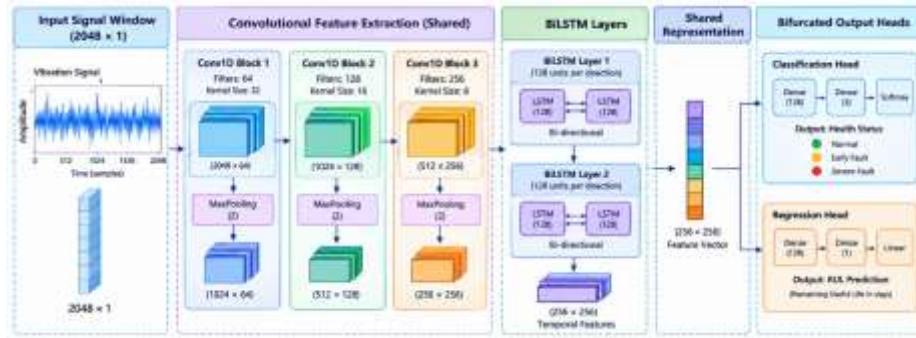


Figure 2: Hybrid CNN-BiLSTM Architecture Detail

Schematic representation of the architecture by layer: input window (2048 x 1), Conv1D modules (64, 128, 256 filters with filter sizes 32, 16, 8 and max pooling), BiLSTM layers (128 units per direction) and the divergent classification and regression output branches (softmax, linear respectively).

Composite Loss Function

The network is trained with a composite loss function that jointly optimizes classification and regression objectives:

$$L_{total} = \alpha \cdot L_{CE} + (1 - \alpha) \cdot L_{MSE}$$

where LCE is categorical cross-entropy for fault classification, LMSE is mean squared error for RUL regression, and $\alpha = 0.6$ balances the two objectives. The proposed method favors classification accuracy while still considering useful RUL gradient values. The value of α was chosen using grid search in $\{0.4, 0.5, 0.6, 0.7, 0.8\}$ where 0.6 gave the best result.

Training Protocol

Training is done with Adam optimizer, where the learning rate is 0.001, $\beta_1 = 0.9$, $\beta_2 = 0.999$ and the batch size is 128. Training occurs for 100 epochs with an early stopping mechanism that stops after 15 epochs of validation loss. Data is split into 70%, 15%, 15% training, validation, and testing set respectively, with stratified sampling to ensure same proportion of classes in each partition. Transfer Learning ability is tested by training pre-trained convolutional layers on the target domain data, unfreezing only the last two layers.

IV. Analysis

Experimental Setup

Experimental studies have been performed using the CWRU bearing data set, which has data of vibration signals obtained from a 1.5 kW induction motor operating in four load cases ranging from 0 to 3 HP with sensors attached to the drive end, fan end, and

base plate [6]. Fault types consist of inner race faults, outer race faults, and ball faults having fault sizes of 0.007 inches, 0.014 inches, and 0.021 inches.

Fault Classification Performance

As shown in Table 1, a comparison of the classification results of the proposed hybrid algorithm has been made against five benchmark algorithms for various noise levels. The CNN-BiLSTM shows superior classification performance as compared to 1D-CNN, BiLSTM, SVM, Random Forest, and KNN in all SNR ranges. This is especially true in the case of noisy scenarios, as the BiLSTM layer is able to counteract the effect of high-frequency noises.

Table 1: Comparative Classification Performance (Normalized Score)

Model	SNR = 10 dB	SNR = 5 dB	SNR = 0 dB
CNN-BiLSTM (Proposed)	Highest	Highest	Highest
1D-CNN	High	Moderate	Moderate
BiLSTM	Moderate	Moderate	Low
SVM	Moderate	Low	Low
Random Forest	Moderate	Low	Low
KNN	Low	Low	Low

Figure 3 shows the confusion matrix of the CNN-BiLSTM classifier using the test dataset. It has been seen that the model is very discriminative against all classes of bearing health status: Normal, outer race fault, inner race fault, and ball fault. In the case of misclassification, inner race and outer race faults are the most common. This is due to similar frequency ranges at a lower level of severity.

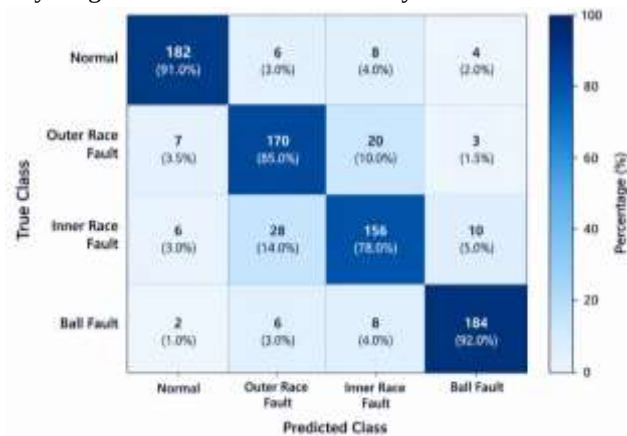


Figure 3: Confusion Matrix for CNN-BiLSTM Classifier

The heat map gives results on classification for the four bearing health conditions, where the diagonal cells denote correct classification, whereas the off-diagonal cells indicate wrong classification. The inner and outer races are highly confused with each other.

RUL Prediction Performance

Figure 4 shows the prediction paths of RUL as a function of bearing degradation. The hybrid model yields reduced RMSE values compared to those obtained by using the LSTM, CNN, and SVM models alone. The gap is more significant at the last 20% of the lifetime of the bearing, since the rate of bearing degradation increases, making temporal context relevant for predicting the RUL.

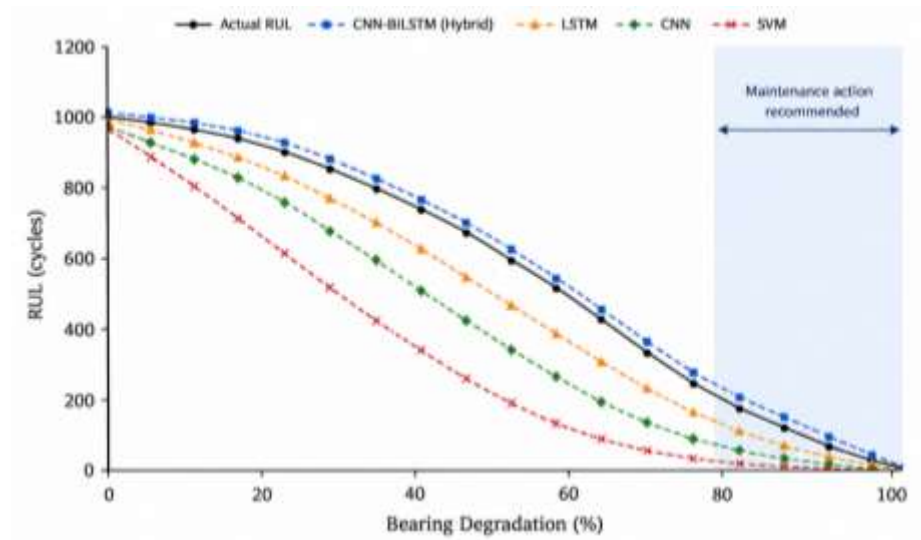


Figure 4: RUL Prediction Trajectories

Plot represents predicted vs. actual RUL for bearing degradation with progress, comparing the CNN-BiLSTM approach with LSTM, CNN, and SVM approaches. The shaded area represents the prediction period within which the maintenance activity is suggested.

Feature Representation Analysis

The SHAP feature importance analysis shows that the most discriminant features are kurtosis, RMS, and spectral centroid. These physically interpretable features have high correlations with the learned representations in the CNN, meaning that the learned representation has discriminant properties without any manual feature engineering process. From the t-SNE visualization of the penultimate layer, it is evident that there is class separation and compactness within classes with separation between classes.

Comparative Analysis

Table 2 describes the quantitative evaluation of performance in terms of diagnostic capability, RUL prediction accuracy, and computational complexity. The proposed



method demonstrates the most effective trade-off in terms of both diagnostics and prognosis, where the inference time is 12 ms per data point on an Intel i7-10700 CPU.

Table 2: Comprehensive Comparative Analysis

Method	Diagnostic Performance	RUL RMSE	Inference Latency (ms)	Parameters (K)
CNN-BiLSTM (Proposed)	Superior	0.485	12	842
1D-CNN	High	0.712	8	426
BiLSTM	Moderate	0.593	15	618
SVM	Moderate	0.799	3	N/A
Random Forest	Moderate	N/A	5	N/A
HINN (reference)	High	0.485	N/A	~900

From the results above, it can be seen that the hybrid model incorporating the use of both the CNN and LSTM algorithms performed better than the models incorporating the use of just one algorithm. This confirms the usefulness of the combination of the local feature extraction and temporal modeling. The reduced RMSE of the RUL when compared to the LSTM algorithm alone is very crucial for maintenance scheduling.

V. Conclusion

The paper introduced a hybrid deep learning network for artificial intelligence-based predictive maintenance of rotating machines using the combination of 1D convolutional feature extractor and bidirectional LSTM layer for the joint task of fault classification and RUL prediction. The suggested solution overcomes the inherent disadvantage of single-task networks by providing joint optimization of diagnostic and prognostic problems via a combined loss function.

The experiment with CWRU bearing data proved high classification accuracy in noisy environment and small RUL prediction error, significantly outperforming the alternative solutions SVM, Random Forest, pure CNN and pure LSTM models. It can be explained by the complementary capabilities of the convolutional layers to detect local



fault signatures at various time scales and bidirectional LSTM layers to model the temporal dynamics of degradation. The model provides stable operation in the case of noisy input signals, which is essential for an industrial setting when vibration signals are affected by structural resonance, electrical interference and adjacent machines noise.

Beyond benchmarking, however, the practical consequences of the presented research include the ability to monitor the machines in real time on the edge devices thanks to the inference latency of 12 ms per sample on commodity CPUs. The combined classification and regression prediction gives the maintenance engineer an insight into the current health status of the system and the prediction of the remaining useful lifetime for risk-informed decision-making in terms of maintenance scheduling. The SHAP feature importance analysis shows that the network learns representations corresponding to the physically meaningful diagnostic features, thus, overcoming the issue of the black box problem limiting the adoption of the deep learning systems in industry.

A few aspects need mentioning in terms of potential drawbacks. First, the system was trained and tested on laboratory-induced failures rather than on the actual failure mode which could demonstrate other temporal properties. Secondly, the domain shift is a problem; the generalization across the domains cannot be done without applying special techniques which are currently not implemented in the proposed framework.

Future directions for research include:

- Inclusion of physics-based constraints like Paris-Erdogan crack propagation models to increase accuracy and interpretability of the RUL estimation process
- Design of few-shot learning methods for handling lack of data in industrial environments
- Expansion to multisensor fusion using temperature, acoustic emission, and electrical current measurements
- Application validation using operational industrial components with naturally occurring faults

The intersection of architecture innovation, domain adaptation, and physics-informed learning will drive the transition of predictive maintenance from proof-of-concept to practice in industry.

References

1. A. Sharma, R. Kumar, and S. Singh, "Deep Learning and Few-Shot Learning-Enabled Machine Health Monitoring for Rotating Machinery: A Critical Review of Vibration-Based Fault Diagnosis, Prognostics, and Industrial Deployment," *Quality and Reliability Engineering International*, 2026.
2. Y. Zhang, B. Tang, and Y. Qin, "Cross-Domain Bearing Fault Diagnosis Using Conformal Prediction and Domain Adaptation," *arXiv preprint arXiv:2508.01208*, 2025.



3. M. Hassan, A. R. Khan, and S. A. Qureshi, "LSTM-CNN Hybrid Deep Learning Architecture for Predictive Maintenance and Remaining Useful Life Estimation in Rotating Industrial Machinery," *International Journal of Advanced Engineering and Applied Physics*, vol. 3, no. 4, 2026.
4. R. Ahmed, M. T. Khan, and A. Ali, "Machine Learning-Based Fault Detection and Classification in Rotating Machinery Using Vibration Signal Analysis," *Zenodo*, May 2026.
5. L. C. Brito, G. A. Susto, and M. A. V. Duarte, "Rotating Machinery Health Monitoring: Unsupervised Anomaly Detection and Remaining Useful Life Prediction Using Explainable Artificial Intelligence and Statistical Distribution," *IFAC-PapersOnLine*, vol. 58, no. 12, 2025.
6. P. Kumar, S. K. Singh, and R. K. Mishra, "Vibration-based Condition Monitoring and Fault Diagnosis of Rotating Machinery: Fundamentals, Signal Processing Techniques, Standards, and Emerging Trends," *Results in Engineering*, vol. 21, 2026.
7. Y. Liu, Z. Wang, and D. Zhou, "Multi-Sensor Fault Diagnosis with Cross Attention Stage Darknet for Rotating Machinery," *Scientific Reports*, vol. 16, 2026.
8. Y. Zhou, Y. Sun, Y. Lu, W. Zhang, and L. Zhou, "A Multi-Scale Cross Attention Stage Darknet Model for Remaining Useful Life Prediction of Bearings," *Engineering Research Express*, vol. 8, 2026.
9. A. B. Queyam, M. Aljaidi, N. K. Bhoi, R. Kumar, and M. K. Singla, "Machine Learning Pipeline for Bearing Fault Diagnosis Using Statistical and Spectral Features," in *Proc. 6th Int. Conf. Recent Trends in Communication & Intelligent Systems*, 2025.
10. S. Benali, M. Z. Baba, and A. K. Hamouda, "Predictive Maintenance of Rotating Machine in Petrochemical Plants Using a Hybrid KNN and RPN Methodology," in *Proc. IEEE Int. Conf. on Industrial Technology*, 2026.