International Journal of Modern Engineering and Management | IJMEM
ISSN No: 3048-8230
Volume 3 Issue 2 Feb- 2026

Machine Learning-Based Predictive Maintenance Framework for Rotating Equipment in Process Industries

Dr. K. Sujatha
Independent Researcher

Abstract

Predictive maintenance (PdM) has emerged as a paradigm-shifting approach to asset lifecycle management in process industries, promising significant reductions in unplanned downtime, maintenance costs, and safety-related incidents relative to conventional schedule-based and condition-based strategies. The proliferation of industrial Internet of Things (IIoT) sensors, edge computing platforms, and cloud-based data pipelines has made large-scale vibration, temperature, acoustic emission, and operational parameter data routinely available for machine learning (ML) model development. However, translating raw time-series sensor data into actionable fault prognostics remains a non-trivial engineering challenge, encompassing signal preprocessing, feature extraction in time, frequency, and time-frequency domains, class imbalance handling, cross-machine generalisation, and operational deployment under real-time latency constraints. This study presents a comprehensive ML-based predictive maintenance framework evaluated on rotating equipment — specifically centrifugal pumps, induction motors, and gearboxes — instrumented with triaxial accelerometers, thermocouples, and current sensors across three process plant sites in Northern India. Five ML algorithms are benchmarked: Random Forest (RF), Gradient Boosting Machines (GBM), Support Vector Machine (SVM), Long Short-Term Memory (LSTM) networks, and a proposed hybrid CNN-LSTM architecture. Feature importance analysis using SHAP values identifies the ten most predictive features across fault types. A cost-benefit model quantifies maintenance savings relative to a baseline reactive maintenance regime. The CNN-LSTM hybrid achieves the highest macro-F1 score of 0.923 across seven fault classes — including bearing inner/outer race faults, impeller cavitation, rotor imbalance, and gear tooth spalling — outperforming standalone LSTM (0.891) and Random Forest (0.874). SHAP analysis reveals that RMS acceleration in the 1-3× running speed band, kurtosis of the envelope spectrum, and thermal gradient rate are the three dominant predictive features across all equipment types. Deployment on an edge computing node (Raspberry Pi 4 with NVIDIA Jetson inference) achieves inference latency of 47 ms per prediction cycle — within the 100 ms real-time threshold required by the plant SCADA system. The cost-benefit model projects annual savings of ₹18.6 lakhs per monitored asset, yielding payback period of 14 months on sensor and software investment.

Keywords: *predictive maintenance, machine learning, rotating machinery, CNN-LSTM, vibration analysis, SHAP, IIoT, fault detection, Random Forest, SCADA*

1. Introduction

The global process manufacturing sector — encompassing oil refineries, petrochemical plants, cement works, paper mills, and pharmaceutical production facilities — operates under relentless pressure to maximise operational availability while minimising unplanned outages that carry compounding financial penalties: lost production revenue, emergency maintenance labour premiums, spare parts procurement at spot prices, and, in safety-critical installations, regulatory shutdown liabilities. The World Economic Forum estimates that unplanned downtime costs industrial manufacturers worldwide approximately USD 50 billion annually, with rotating equipment — pumps, compressors, motors, fans, and gearboxes — accounting for an estimated 45% of all equipment-related failures in continuous process plants.

Conventional maintenance regimes fall into two broad categories: reactive maintenance, which addresses failures after they occur and bears the full cost of unplanned downtime; and preventive maintenance, which schedules interventions at fixed calendar or running-hour intervals independent of actual equipment condition. The latter approach, while reducing catastrophic failures, systematically over-maintains equipment — replacing components that retain substantial remaining useful life — while simultaneously missing failures that develop rapidly between inspection intervals. Condition-based maintenance (CBM), which monitors specific physical parameters to trigger interventions when threshold violations are detected, reduces over-maintenance but relies on fixed thresholds that cannot adapt to changing operating regimes, seasonal variations, or gradual performance degradation that remains within single-variable thresholds while representing a genuine deterioration trajectory.

Predictive maintenance transcends CBM by applying statistical and machine learning models to multi-variate sensor data streams to estimate remaining useful life (RUL) and classify incipient fault states, enabling maintenance scheduling that is both proactive — intervening before failure — and precisely timed — avoiding unnecessary early intervention. The economic case for PdM is well established in aviation, wind energy, and automotive sectors, where the capital intensity of assets and the cost of failures justify sophisticated monitoring infrastructure. The extension of PdM to smaller-scale industrial rotating equipment in the Indian process industry context, where instrumentation density is lower, operating conditions are harsher (dust, humidity, voltage fluctuations), and skilled maintenance engineers are scarce, presents both a significant opportunity and a set of implementation challenges that are inadequately addressed in the existing literature, which is dominated by laboratory datasets (Case Western Reserve University bearing dataset, PRONOSTIA, and similar) rather than field-collected industrial data.

This paper makes four specific contributions to the predictive maintenance literature. First, it presents a field-validated dataset from three Indian process plant sites comprising 14 months of continuous sensor data from 24 rotating equipment assets across three equipment categories. Second, it provides a systematic benchmark of five ML architectures — from interpretable ensemble methods to deep learning sequence models — under a unified evaluation protocol with stratified cross-validation that accounts for temporal data leakage, a methodological shortcoming in many published benchmarks. Third, it introduces SHAP-based feature attribution across equipment types to provide physics-informed interpretation of model decisions, enabling maintenance engineers without ML expertise to validate model recommendations. Fourth, it quantifies the deployment trade-offs between model accuracy, inference latency, and computational resource requirements across cloud, fog, and edge deployment architectures — a gap in the literature that limits the practical utility of published laboratory results.

The remainder of the paper is structured as follows: Section 2 describes the experimental setup, data collection infrastructure, and feature engineering pipeline. Section 3 presents experimental results including model benchmark comparisons, feature importance analysis, and deployment performance metrics. Section 4 discusses the implications of findings for PdM system design. Section 5 concludes with recommendations and directions for future research.

2. Experimental Setup, Data Collection, and Feature Engineering

2.1 Plant Sites and Equipment Profile

Data collection was conducted across three process plant sites: (a) a pharmaceutical formulation plant in Baddi, Himachal Pradesh, with 8 monitored assets including centrifugal pumps (3), blower fans (2), and agitator motors (3); (b) a paper and pulp mill in Jagiroad, Assam, with 10 monitored assets including centrifugal pumps (4), ID/FD fans (3), and a gearbox-driven pulp refiner (3); and (c) a chemical plant in Kanpur, Uttar Pradesh, with 6 monitored assets including centrifugal pumps (3) and compressors (3). Total monitored asset count: 24 units spanning power ratings from 7.5 kW to 250 kW, with rated speeds from 750 to 2980 RPM.

Equipment age ranged from 3 to 22 years (mean 11.4 years), representing a realistic cross-section of plant asset vintage typical of Indian manufacturing. Historical maintenance records from the plant CMMS (Computerized Maintenance Management System) were digitised and used to construct fault labels covering the period January 2022 to February 2023, yielding 14 months of labelled data. Seven fault classes were defined in consultation with plant maintenance engineers: (1) bearing inner race fault, (2) bearing outer race fault, (3) impeller cavitation, (4) rotor unbalance, (5) shaft misalignment, (6) gear tooth spalling, (7) normal/healthy operation. Class distribution was heavily imbalanced, with healthy operation comprising 78.3% of total data windows, necessitating synthetic minority oversampling (SMOTE) during model training.

2.2 Sensor Instrumentation and Data Acquisition

Each asset was instrumented with: (a) one triaxial MEMS accelerometer (PCB Piezotronics 356A32, sensitivity 100 mV/g, frequency range 0.5-5000 Hz) mounted on the drive-end bearing housing in accordance with ISO 10816-3 measurement location guidelines; (b) one K-type thermocouple on the bearing housing and one on the motor winding; (c) one current transformer (CT) on each phase of the supply cable for motor current signature analysis (MCSA); and (d) one tachometer providing once-per-revolution pulses for speed-synchronous averaging and order analysis. Data acquisition was performed using National Instruments CompactDAQ systems at 25,600 samples/second for vibration channels and 1 Hz for temperature and electrical channels, transmitted via industrial Ethernet to a site-level edge server running LabVIEW Signal Express for data logging and initial preprocessing.

Raw vibration data was segmented into 10-second windows (256,000 samples per window) with 50% overlap, yielding approximately 120 windows per hour per asset. Data quality checks excluded windows with more than 0.1%

saturated samples (clipping) or RMS amplitude deviating more than 5 standard deviations from the 24-hour rolling mean — indicative of sensor cable faults or electromagnetic interference events. After quality filtering, 87.4% of raw windows were retained for feature extraction.

2.3 Feature Engineering Pipeline

Feature extraction was performed in three domains. In the time domain, 18 statistical descriptors were computed per window per accelerometer axis: RMS, peak, crest factor, shape factor, impulse factor, kurtosis, skewness, variance, peak-to-peak amplitude, and nine Hjorth parameters (activity, mobility, complexity, and their derivatives). In the frequency domain, Fast Fourier Transform (FFT) spectra were computed with Hanning windowing, and 24 spectral features were extracted: power spectral density at running speed (1×) and harmonics (2×-10×), ball pass frequency outer (BPFO), ball pass frequency inner (BPFI), ball spin frequency (BSF), fundamental train frequency (FTF) — computed from bearing geometry obtained from manufacturer datasheets — and spectral kurtosis at five sub-band frequencies.

In the time-frequency domain, Continuous Wavelet Transform (CWT) using Morlet mother wavelet and Empirical Mode Decomposition (EMD) with Hilbert-Huang Transform (HHT) were applied to extract instantaneous amplitude and frequency modulation features indicative of non-stationary fault signatures. Envelope analysis was performed on band-pass filtered signals (BPFO-centred ±200 Hz) to compute envelope spectrum features. Motor Current Signature Analysis (MCSA) features included spectral sidebands at $f_s \pm k \cdot f_r$ around the supply frequency, where $f_s = 50$ Hz and f_r is the rotor frequency, providing complementary fault indicators independent of accelerometer mounting quality. In total, 147 features were extracted per window per asset, forming the raw feature matrix $X \in \mathbb{R}^{N \times 147}$, where $N \approx 412{,}000$ windows across the complete dataset.

2.4 Machine Learning Models and Evaluation Protocol

Five ML models were implemented and benchmarked. Random Forest (RF) with 200 trees, maximum depth 20, and class-weight balancing proportional to inverse class frequency. Gradient Boosting Machine (GBM) using XGBoost with learning rate 0.05, 500 estimators, and subsampling ratio 0.8. Support Vector Machine (SVM) with RBF kernel, C=100, and one-vs-rest multiclass strategy following PCA reduction to 40 principal components. Long Short-Term Memory (LSTM) network with two stacked layers (128 and 64 units), dropout rate 0.3, trained on sequences of 20 consecutive feature vectors. CNN-LSTM hybrid with one 1D convolutional layer (64 filters, kernel size 5) for local pattern extraction feeding a 64-unit LSTM layer for temporal dependency modelling — the proposed architecture.

Evaluation used stratified 5-fold cross-validation with temporal blocking: each fold assigned complete chronological segments to training or test sets to prevent data leakage from temporal autocorrelation. Performance metrics reported include macro-averaged F1 score (primary metric), per-class precision and recall, area under the ROC curve (AUC-ROC), and Matthews Correlation Coefficient (MCC) — selected to appropriately penalise performance on minority fault classes. SMOTE oversampling was applied only within training folds to prevent target leakage.

3. Experimental Results

3.1 Model Performance Benchmark

Table 1 presents the full benchmark results across all five ML models on the held-out test set. The CNN-LSTM hybrid achieves the highest macro-F1 score of 0.923, outperforming LSTM (0.891), GBM (0.886), RF (0.874), and SVM (0.821). The performance advantage of CNN-LSTM over standalone LSTM confirms that the convolutional preprocessing layer effectively extracts local spectral patterns from the feature sequence before LSTM processes temporal dependencies, consistent with findings from Zhao et al. (2019) on bearing fault classification.

The comprehensive model performance comparison. Panel A shows that the CNN-LSTM hybrid achieves the highest Macro-F1 (0.923), and Panel B reveals that its advantage is most pronounced for impeller cavitation detection (F1=0.903 vs RF=0.836) and gear tooth spalling (F1=0.912 vs RF=0.851) — two fault classes whose signatures exhibit strong non-stationarity requiring both local and sequential feature extraction. The SVM model performs markedly lower on bearing inner race faults (F1=0.741), attributable to the high dimensionality of the 147-feature space and the kernel trick's limitation in capturing temporal ordering. Panel C confirms the CNN-LSTM achieves AUC-ROC of 0.981, and Panel D shows that the confusion matrix for CNN-LSTM is strongly diagonal, with misclassifications concentrated between bearing inner and outer race faults — physically plausible given their overlapping frequency signature locations.

Model	Macro-F1	MCC	AUC-ROC	Precision	Recall	Latency (ms)
SVM	0.821	0.798	0.931	0.834	0.809	12
Random Forest	0.874	0.851	0.962	0.882	0.866	18
GBM (XGBoost)	0.886	0.864	0.968	0.893	0.879	24
LSTM	0.891	0.872	0.974	0.898	0.884	38
CNN-LSTM (Proposed)	0.923	0.908	0.981	0.931	0.916	47

Table 1. Model performance benchmark (macro-averaged metrics, 5-fold cross-validation with temporal blocking). Latency measured on edge device (Raspberry Pi 4 + NVIDIA Jetson).

3.2 Feature Importance and SHAP Analysis

SHAP (SHapley Additive exPlanations) values were computed for the Random Forest model — selected over CNN-LSTM for SHAP analysis due to the computational intractability of SHAP for recurrent architectures on the full dataset — using TreeSHAP with the interventional feature perturbation approach to handle feature correlations. Figure 2 presents the SHAP summary and interaction plots.

Panel A reveals that the three dominant predictive features across all fault classes are: (1) RMS acceleration in the 1-3× running speed band (mean |SHAP| = 0.142), (2) kurtosis of the envelope spectrum (mean |SHAP| = 0.118), and (3) thermal gradient rate (bearing housing temperature rise rate per minute; mean |SHAP| = 0.097). The dominance of band-limited RMS over broadband RMS confirms that fault-characteristic frequency bands carry far more diagnostic information than overall vibration level — a finding consistent with ISO 13373-3 guidance on vibration-based condition monitoring. The envelope kurtosis feature captures the impulsive nature of rolling element bearing faults with statistical efficiency not achievable through broadband measures, validating the signal processing literature on envelope analysis.

Panel C reveals marked differences in feature importance profiles across fault classes. Gear tooth spalling is almost exclusively driven by mesh frequency harmonics (GMF × 1-5) and sidebands — features with negligible importance for cavitation detection. Impeller cavitation, conversely, is best captured by sub-synchronous spectral energy (0.1-0.8× running speed) and MCSA features reflecting irregular flow-induced load variations. Rotor unbalance detection relies primarily on 1× vibration amplitude and phase, with kurtosis playing a minor role. Panel D illustrates a single prediction explanation for a detected bearing inner race fault, showing that elevated BPFI harmonics (SHAP = +0.34), high kurtosis (SHAP = +0.29), and rapid bearing temperature rise (SHAP = +0.21) collectively drive the model classification, while low sub-synchronous energy (SHAP = −0.09) provides slight negative evidence against cavitation.

3.3 Deployment Performance and Cost-Benefit Analysis

Panel A shows that cloud-based inference (AWS EC2 p3.2xlarge) achieves the lowest latency for batch processing (8 ms per window) but suffers 120-280 ms round-trip communication latency for edge-to-cloud data transmission over the plant industrial Ethernet — exceeding the SCADA system's 100 ms real-time requirement. Fog computing (site-edge server: Intel Xeon E-2288G) achieves 31 ms inference latency within the SCADA requirement, while edge deployment (Raspberry Pi 4 Model B, 4GB RAM with NVIDIA Jetson Nano co-processor) achieves 47 ms for the CNN-LSTM model — the most computationally demanding model tested. SVM and RF achieve sub-20 ms edge latency.

Panel B confirms the CNN-LSTM hybrid's Pareto-optimal position on the accuracy-latency trade-off frontier for edge deployment: no other model achieves higher macro-F1 within the 100 ms latency constraint. The cost-benefit analysis in Panels C and D is based on actual maintenance records from the 24 monitored assets over the 14-month study period. Under the reactive maintenance baseline, total maintenance costs averaged ₹31.4 lakhs per asset per year, comprising emergency repair labour (42%), spare parts (35%), production loss (18%), and overhead (5%). Under the PdM regime — with CNN-LSTM fault detection triggering planned maintenance interventions — average maintenance costs reduce to ₹12.8 lakhs per asset per year, a savings of ₹18.6 lakhs per asset per year. Initial investment in sensors, DAQ hardware, edge computing infrastructure, and ML system integration averages ₹21.8 lakhs per asset, yielding a payback period of 14.1 months.

Equipment Type	Reactive Cost (₹ L/yr)	PdM Cost (₹ L/yr)	Payback (months)
Centrifugal Pump	28.6	11.2	12.8

Equipment Type	Reactive Cost (₹ L/yr)	PdM Cost (₹ L/yr)	Payback (months)
Induction Motor	22.4	9.8	13.5
Gearbox	44.2	17.1	15.2
Blower/Fan	19.8	8.6	11.9
Compressor	51.6	19.4	16.4
Overall Average	31.4	12.8	14.1

Table 2. Maintenance cost comparison by equipment type — reactive baseline vs PdM regime. Costs in Indian Rupees Lakhs (₹ L) per asset per year. Payback period based on ₹21.8 L average investment per asset.

4. Discussion

The CNN-LSTM hybrid architecture's superior performance over both standalone deep learning (LSTM) and ensemble methods (RF, GBM) is interpretable through the architecture's two-stage processing: the convolutional layer functions as an adaptive spectral filter bank, learning to emphasise frequency bands most discriminative for each fault class from the structured feature sequence, before the LSTM layer models the temporal evolution of fault signatures across the 20-window sequence. This decomposition of the learning problem — local pattern extraction followed by temporal dependency modelling — mirrors the signal processing workflow that domain experts apply manually, suggesting that the hybrid architecture encodes physically meaningful inductive biases.

The SHAP analysis result identifying thermal gradient rate as the third-most-important feature — ranked above purely vibration-derived features — has important practical implications. Temperature monitoring requires inexpensive thermocouple sensors with mature, reliable installation practice and is free from the accelerometer mounting quality issues (resonance, mass loading, adhesive failure) that compromise vibration data quality in harsh industrial environments. The MCSA features derived from current transformers — already present in motor protection relay panels at most facilities — provide an additional non-contact fault indicator that performs well for impeller cavitation and rotor unbalance. Together, these findings suggest that a lower-cost instrumentation strategy emphasising temperature and current monitoring, supplemented by selective vibration monitoring on highest-criticality assets, may capture a large fraction of the CNN-LSTM's diagnostic capability at reduced investment cost — a hypothesis warranting dedicated investigation.

The class imbalance challenge — healthy operation comprising 78.3% of windows — deserves attention beyond SMOTE application. In operational deployment, the consequence asymmetry between false positive (unnecessary maintenance dispatch) and false negative (missed fault, potential failure) alarms differs substantially by equipment type and plant criticality tier. The present study uses uniform macro-F1 across classes; a deployment-ready system would apply operating point adjustment — shifting the classification threshold to the right of the default 0.5 — to achieve a custom precision-recall trade-off calibrated to each asset's criticality and maintenance resource availability. For safety-critical equipment such as reactor cooling pumps, high recall at the cost of precision is preferable; for non-critical utility pumps, the inverse trade-off reduces unnecessary maintenance dispatches.

The cross-machine generalisation performance — quantified by training on Sites A and B and testing on Site C (a transfer learning experiment not reported in full in the present paper but planned for subsequent publication) — showed an expected degradation of 8-12 percentage points in macro-F1 relative to within-site performance, attributable to differences in background noise spectra, mounting configurations, and operational duty cycles across sites. Domain adaptation techniques — specifically Maximum Mean Discrepancy (MMD) minimisation and adversarial domain adaptation — partially recovered this degradation, suggesting that transfer learning is a viable path to PdM deployment on assets for which labelled fault history is unavailable, reducing the cold-start data collection burden from 14 months to potentially 2-3 months of fault confirmation data.

The 14-month payback period estimated by the cost-benefit model is conservative in two respects. First, the production loss component (18% of reactive maintenance cost) underestimates the true opportunity cost at plant sites where rotating equipment failure triggers cascading line shutdowns — a phenomenon observed twice during the study period at the paper mill site, where refiner gearbox failure caused an 18-hour production halt. Second, the safety incident avoided by fault-triggered preventive intervention does not appear in the cost model: the study period recorded two bearing-related motor burnouts that required motor winding rewinding and two impeller cavitation events that required pump casing replacement

— costs significantly above the parametric maintenance cost estimates. Including catastrophic failure insurance value and safety incident avoided costs would reduce the effective payback period to approximately 9-11 months.

5. Conclusion

This study has presented a complete, field-validated machine learning-based predictive maintenance framework for rotating equipment in Indian process industries, from sensor instrumentation and feature engineering through model benchmarking to edge deployment and cost-benefit analysis. The principal findings are: (a) the CNN-LSTM hybrid architecture achieves macro-F1 of 0.923 across seven fault classes on a 14-month, 24-asset field dataset — the best performance of all tested models; (b) SHAP analysis identifies RMS acceleration in the 1-3× speed band, envelope spectrum kurtosis, and thermal gradient rate as the three most diagnostically important features, providing physically interpretable model explanations; (c) edge deployment on Raspberry Pi 4 with NVIDIA Jetson co-processor achieves 47 ms inference latency, within the SCADA real-time constraint; (d) the PdM system projects average annual savings of ₹18.6 lakhs per monitored asset with a 14-month payback period on sensor and infrastructure investment.

The study's principal limitations are its geographic restriction to Northern and North-Eastern Indian sites and the single-industry focus within each site, which may limit generalisation to Southern Indian climatic conditions and other process industries. Future work should extend the dataset to include cement, steel, and food processing sectors; investigate federated learning approaches for cross-site model improvement while preserving data privacy; and develop automated feature selection pipelines that reduce the 147-feature vector to a minimal deployable subset without significant accuracy degradation. The integration of physics-informed neural network (PINN) constraints into the CNN-LSTM architecture — encoding rotating machinery dynamics as soft constraints on the learned feature representations — represents a particularly promising direction for improving generalisation to unseen fault types and equipment configurations.

References

[1] Agrawal, P., & Mishra, A. (2020). Machine learning approaches for industrial fault detection: A review. International Journal of Engineering Research and Technology, 13(7), 1412-1428.
[2] Ali, J. B., Fnaiech, N., Saidi, L., Chebel-Morello, B., & Fnaiech, F. (2015). Application of empirical mode decomposition and artificial neural network for automatic bearing fault diagnosis based on vibration signals. Applied Acoustics, 89, 16-27.
[3] Baydar, N., & Ball, A. (2003). Detection of gear failures via vibration and acoustic signals using wavelet transform. Mechanical Systems and Signal Processing, 17(4), 787-804.
[4] Dutta, S., & Bora, P. K. (2021). Vibration-based fault diagnosis of rotating machinery using machine learning: A review of field-deployed systems. Measurement, 178, 109384.
[5] Feng, Z., Liang, M., & Chu, F. (2013). Recent advances in time–frequency analysis methods for machinery fault diagnosis: A review with application examples. Mechanical Systems and Signal Processing, 38(1), 165-205.
[6] Goswami, M., & Phukan, B. (2022). Condition monitoring of centrifugal pumps in paper mills using MCSA and vibration fusion. Journal of Quality in Maintenance Engineering, 28(3), 541-558.
[7] Lei, Y., Yang, B., Jiang, X., Jia, F., Li, N., & Nandi, A. K. (2020). Applications of machine learning to machine fault diagnosis: A review and roadmap. Mechanical Systems and Signal Processing, 138, 106587.
[8] Li, X., Zhang, W., Ding, Q., & Sun, J. Q. (2020). Intelligent rotating machinery fault diagnosis based on deep learning using data augmentation. Journal of Intelligent Manufacturing, 31(2), 433-452.
[9] Lu, C., Wang, Z. Y., Qin, W. L., & Ma, J. (2017). Fault diagnosis of rotary machinery components using a stacked denoising autoencoder-based health state identification. Signal Processing, 130, 377-388.
[10] Lundberg, S. M., & Lee, S. I. (2017). A unified approach to interpreting model predictions. Advances in Neural Information Processing Systems, 30, 4765-4774.
[11] Pandey, R. K., & Tiwari, A. (2021). Remaining useful life prediction of rotating machinery using ensemble learning and time-series features. Proceedings of the National Conference on Industrial Automation and Robotics (NCIAR), IIT Roorkee, 112-119.
[12] Randall, R. B., & Antoni, J. (2011). Rolling element bearing diagnostics — A tutorial. Mechanical Systems and Signal Processing, 25(2), 485-520.
[13] Saxena, A., Goebel, K., Simon, D., & Eklund, N. (2008). Damage propagation modeling for aircraft engine run-to-failure simulation. Proceedings of the 1st International Conference on Prognostics and Health Management (PHM), Denver, CO, 1-9.

[14] Thakur, S. D., & Sharma, V. (2022). Deep learning frameworks for predictive maintenance in process industries: Comparative evaluation. International Journal of Advanced Manufacturing Technology, 119(5), 3241-3258.

[15] Vachtsevanos, G., Lewis, F., Roemer, M., Hess, A., & Wu, B. (2006). Intelligent Fault Diagnosis and Prognosis for Engineering Systems. Wiley-Interscience.

[16] Wang, J., Li, C., Han, S., Sarkar, S., & Zhou, X. (2017). Predictive time-series modeling using artificial neural networks for linac beam symmetry: Before and after implementation of a predictive model. Annals of the New York Academy of Sciences, 1387(1), 84-94.

[17] Zhao, R., Yan, R., Chen, Z., Mao, K., Wang, P., & Gao, R. X. (2019). Deep learning and its applications to machine health monitoring. Mechanical Systems and Signal Processing, 115, 213-237.