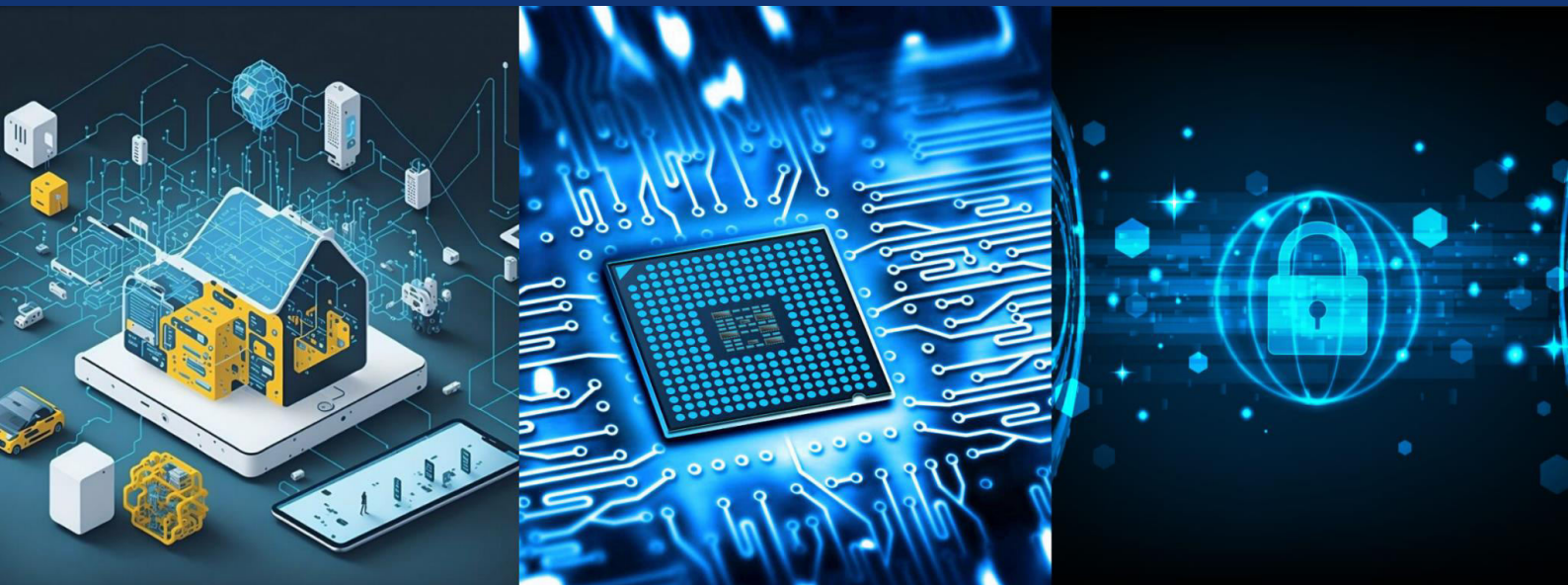
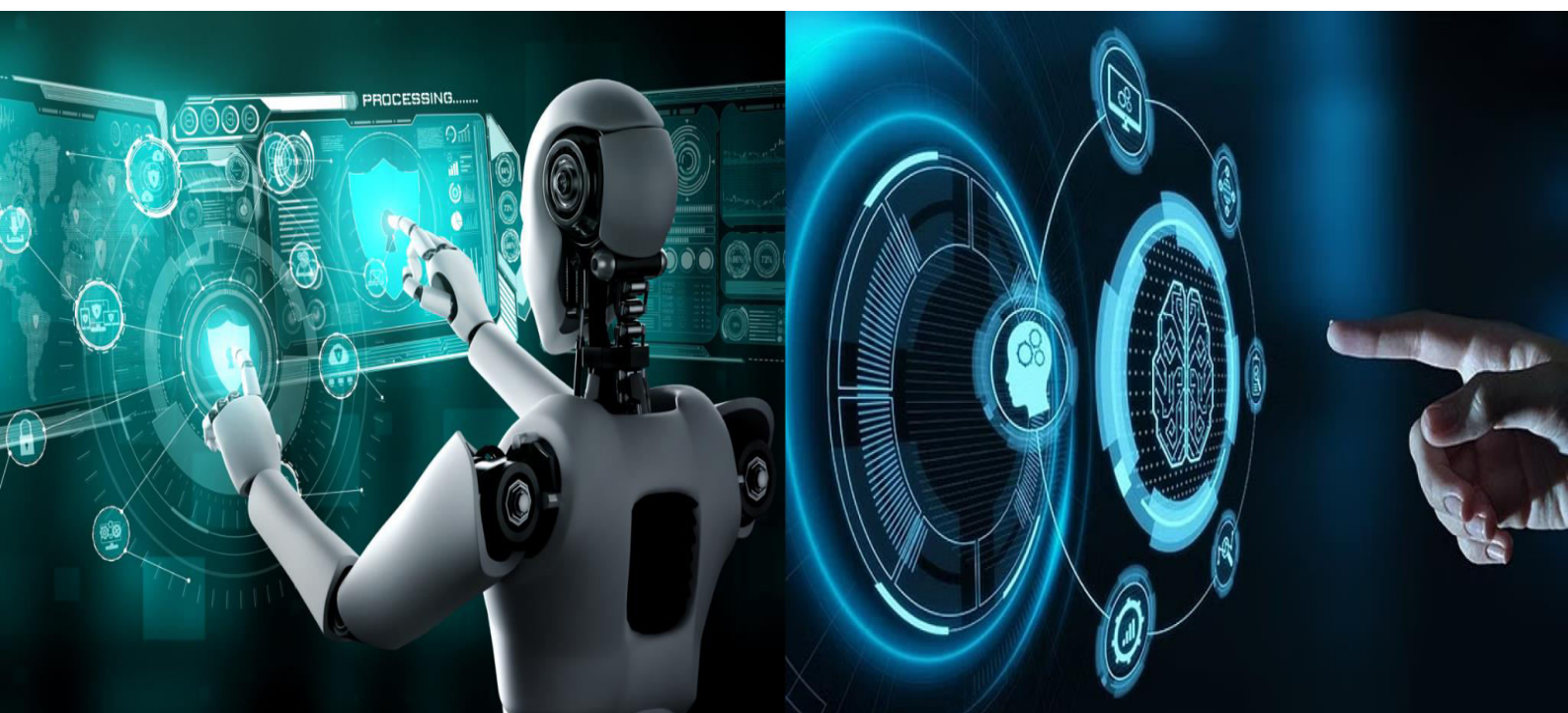




International Journal of Innovative Research in Computer and Communication Engineering

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)





Toward Trustworthy Edge Intelligence for Predictive Maintenance in Smart Manufacturing: A Critical Review of 30 Published Studies

Aditya Zade¹, Sheetal Devi², Radha Tadas³, Sumedh Meshram⁴, Prof. Dhiraj Kumar Gupta⁵,
Mr. Harsh Kalbande⁶

Department of Computer Science and Engineering, St. Vincent Pallotti College of Engineering and Technology,
Nagpur, Maharashtra, India¹⁻⁴

Guide, St. Vincent Pallotti College of Engineering and Technology, Nagpur, Maharashtra, India⁵

Alumni Mentor, St. Vincent Pallotti College of Engineering and Technology, Nagpur, Maharashtra, India⁶

ABSTRACT: Predictive maintenance is a key component of smart manufacturing, linking machine-health monitoring with reduced downtime, improved safety, quality, and maintenance cost. This review critically examines **30 published studies** covering machine learning, deep learning, bearing diagnosis, vibration analysis, Edge AI, TinyML, Industrial IoT, explainable AI, and benchmark datasets. The literature shows a shift from conventional feature-based methods toward deep learning, transfer learning, transformers, and resource-aware edge deployment. However, industrial adoption remains limited by laboratory-based datasets, inconsistent validation, domain shifts, model drift, energy constraints, and limited human-centered evaluation of explanations. The review highlights the need for lightweight and uncertainty-aware models, multi-sensor fusion, secure edge-cloud collaboration, realistic datasets, explainable decision support, and standardized reporting.

KEYWORDS: Predictive maintenance, smart manufacturing, Industry 4.0, bearing fault diagnosis, condition monitoring, vibration analysis, Edge AI, TinyML, Industrial IoT, explainable artificial intelligence, benchmark datasets.

I. INTRODUCTION

Predictive maintenance (PdM) has evolved from isolated machine-learning models into an end-to-end smart manufacturing process involving sensing, condition assessment, prognosis, and maintenance decisions. Its integration with Industry 4.0, Industrial IoT, and cloud-edge architectures has expanded its role in improving reliability and reducing downtime.

Despite advances in deep learning and other AI methods, industrial deployment remains challenging. High results on curated datasets may be affected by laboratory conditions, overlapping windows, and data leakage, limiting their representation of real industrial performance.

Edge computing and TinyML enable low-latency local processing of vibration and acoustic data while reducing bandwidth requirements, but introduce constraints related to memory, power, model size, secure updates, and explainability. XAI is also important because maintenance predictions can directly influence inspection, downtime, and safety decisions.

This review analyzes **30 published studies** across PdM, deep learning, Edge AI, TinyML, IIoT, XAI, and benchmark datasets to identify major trends, limitations, and research gaps toward **trustworthy edge-based predictive maintenance**.



International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

II. REVIEW METHODOLOGY

Scope and Selection Logic

The review prioritizes peer-reviewed work published from 2020 to 2025, while retaining foundational studies from 2008 to 2019 when they define benchmark datasets or remain essential to the field. Sources were selected from reputable publishers and scholarly venues, including IEEE, Elsevier, Springer Nature, MDPI, PHM Society, and the DCASE workshop. The final corpus was deliberately limited to exactly 30 papers to meet the manuscript constraint while preserving coverage of all ten requested themes.

Inclusion and Exclusion Criteria

A paper was included when its publication metadata and DOI or publisher link could be verified and when it contributed directly to predictive maintenance, machine fault diagnosis, edge deployment, explainability, IIoT, condition monitoring, vibration analysis, or benchmark datasets. Unverifiable titles, duplicate records, non-scholarly reposts, and papers whose relevance was only superficial were excluded. Citation counts are reported only as approximate visible counts at retrieval time in July 2026; they are platform-dependent and should not be interpreted as fixed bibliometric values.

Data Extraction and Critical Coding

For each paper, the review records bibliographic metadata, objective, methodology, dataset, algorithm, hardware, explainability method, principal findings, limitations, and the research gap reported or directly implied by the study. Papers were assigned one primary theme for table organization, although many span multiple themes. The synthesis emphasizes comparison across papers rather than sequential abstract-style summaries.

III. CROSS-THEME LITERATURE SYNTHESIS

Predictive Maintenance as an Industry 4.0 Lifecycle

Predictive maintenance is increasingly viewed as an end-to-end Industry 4.0 lifecycle involving sensing, data processing, diagnosis, prognosis, and maintenance decisions. Studies [1]–[6] highlight the importance of intelligent sensing, connected architectures, organizational integration, and trustworthy AI, showing that model selection is only one part of the overall maintenance system.

From Classical Machine Learning to Deep Learning

Classical ML remains useful under limited resource budgets, while deep learning dominates signal-rich applications. Reviews [7]–[9] show a shift from handcrafted features toward CNNs, LSTMs, transfer learning, attention, and transformers, reducing manual feature engineering but increasing dependence on data quality, computation, and validation.

Bearing Diagnosis and Vibration Analytics

Bearing diagnosis is a major application area, with CNNs used for vibration analysis [10] and transformer-based methods for long-range feature learning [11], [12]. Transfer learning, data augmentation, and reinforcement learning address cross-condition variation, limited data, and class imbalance [13]–[15], although field generalization remains a challenge.

Edge AI, TinyML, and Industrial IoT

Edge and TinyML approaches reduce latency and bandwidth requirements while enabling local inference [16], [17]. Studies demonstrate deployment on Raspberry Pi, ESP32, and ESP32-S3 with transfer learning and resource-aware inference [18], [19]. IIoT further supports low-cost sensing and real-time maintenance within connected smart-factory environments [23], [24]. Key challenges remain in secure updating, drift monitoring, energy efficiency, and lifecycle management.

Explainable and Trustworthy Maintenance Intelligence

XAI is moving toward maintenance-oriented explanations rather than generic visualizations. Shapley-based diagnosis, manufacturing case studies, and broader XAI surveys [20]–[22] show growing interest, but limited evidence exists on whether explanations improve troubleshooting, fault localization, or operator trust.



International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Benchmark Datasets and the External-Validity Problem

Benchmark datasets support reproducible evaluation but differ substantially in purpose and realism. CWRU and Paderborn support fault diagnosis, PRONOSTIA/FEMTO-ST and XJTU-SY provide run-to-failure data, C-MAPSS supports RUL prediction, and MIMII addresses acoustic anomaly detection [25]–[30]. However, many datasets remain laboratory, accelerated, simulated, or domain-specific, limiting direct industrial generalization.

IV. COMPARATIVE LITERATURE MATRIX

TABLE I. COMPARATIVE SYNTHESIS OF THE 30 SELECTED PAPERS

Sr. No.	Author(s)	Key Findings	Research Gap
1	Carvalho et al. [1]	Feature quality, failure-data scarcity, imbalance, and inconsistent evaluation strongly influence reported performance.	More realistic validation, online learning, domain adaptation, and deployment-oriented reporting are required.
2	Zonta et al. [2]	Successful PdM depends on integration across sensing, connectivity, analytics, and maintenance planning rather than model accuracy alone.	Common reference architectures, reproducible evaluation, and end-to-end industrial evidence remain insufficient.
3	Pech et al. [3]	Sensor selection, sensitivity, cost, placement, and data quality shape every downstream diagnostic stage.	Interoperable multi-sensor systems and standardized sensor-to-decision pipelines are needed.
4	Achouch et al. [4]	PdM should be treated as a lifecycle from acquisition and preprocessing through diagnosis, prognosis, and maintenance action.	Stronger integration among AI prediction, maintenance policy, human decisions, and business objectives is required.
5	Nunes et al. [5]	Data readiness, imbalance, changing operating conditions, and organizational integration are persistent barriers.	Practical implementation frameworks and robust evaluation under changing industrial conditions remain open.
6	Ucar et al. [6]	The research agenda is shifting from accuracy-only systems toward scalable, secure, explainable, and dependable maintenance intelligence.	Validated trustworthy-AI frameworks and human-centered evaluation in real maintenance workflows are limited.
7	Wang et al. [7]	Deep models learn degradation patterns effectively but remain sensitive to data scarcity, preprocessing, and domain shift.	Uncertainty estimation, consistent scoring, cross-domain transfer, and industrial validation need further work.
8	Hakim et al. [8]	Transfer learning is increasingly important for variable operating conditions and limited target labels.	Realistic datasets, cross-machine robustness, interpretable diagnosis, and deployment-aware evaluation remain underdeveloped.
9	Tama et al. [9]	CNN-based methods dominate, but preprocessing, dataset choice, and validation protocols vary substantially.	Reproducible benchmarks, physics-informed learning, graph models, and realistic operating-condition studies are needed.
10	Souza et al. [10]	High classification accuracy was achieved while reducing dependence on multiple sensor channels.	External validation, uncertainty handling, and interpretable edge deployment remain open.
11	Ding et al. [11]	Self-attention captured long-range feature relationships and delivered strong benchmark performance.	Compressed transformers, cross-domain tests, and maintenance-oriented interpretation are needed.



International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Sr. No.	Author(s)	Key Findings	Research Gap
12	Hou et al. [12]	Very high average classification accuracy was reported on both evaluation datasets.	Model compression, real edge latency tests, concept drift, and interpretable attention require further study.
13	Huo et al. [13]	The method improved target-domain diagnosis beyond approaches relying only on domain-invariant features.	Explainable, lightweight, and continually adaptable transfer learning is still needed.
14	Wang et al. [14]	Synthetic augmentation improved diagnosis accuracy and model robustness under limited data.	Validation on real industrial degradation, uncertainty quantification, and cross-bearing generalization are needed.
15	Li et al. [15]	The framework improved diagnosis under minority-fault conditions relative to conventional baselines.	Lightweight, stable, and explainable reinforcement learning for industrial maintenance remains open.
16	Kubiak et al. [16]	Edge computing reduces transfer delay and enables immediate feedback for manufacturing decisions.	Industrial reference implementations, workforce readiness, ROI evidence, and domain-specific edge standards are needed.
17	Huang et al. [17]	Local inference shortened decision time and reduced the need to transmit raw vibration streams.	Broader industrial datasets, edge-cloud coordination, cybersecurity, and trustworthy outputs require study.
18	Asutkar et al. [18]	Useful target performance was achieved with very little target-domain data and on-device inference was feasible.	Trustworthy TinyML, drift monitoring, federated updates, and multimodal embedded diagnosis remain open.
19	Gao et al. [19]	The study reported 88.28% accuracy, about 45 ms inference, and about 17.7 mJ per inference.	Open edge benchmarks, field trials, explainability, and safe model updating are needed.
20	Hasan et al. [20]	The pipeline retained diagnostic performance while showing which selected features supported each decision.	Deep yet equally interpretable models and stronger cross-domain validation are needed.
21	Hrnjica and Softic [21]	The work showed that explanations can bridge predictive models and shop-floor reasoning.	Human-centered evaluation, explanation fidelity, and operational decision impact require validation.
22	Alexander et al. [22]	Manufacturing-XAI research is fragmented across users, tasks, explanation forms, and evaluation approaches.	Unified manufacturing-XAI benchmarks, role-aware explanation design, and evidence of decision benefit are needed.
23	Cakir et al. [23]	Decision trees provided strong accuracy and fast inference; low-cost hardware supported real-time maintenance alerts.	Scalable secure architectures, standard edge benchmarks, and cross-asset validation are required.
24	Soori et al. [24]	IoT supplies the data and connectivity foundation for monitoring, optimization, safety, and predictive maintenance.	Interoperability, cybersecurity, latency guarantees, and maintenance-oriented reference architectures remain priorities.
25	Smith and Randall [25]	Some CWRU cases are unusually easy; careless splitting and record selection can exaggerate performance.	More demanding datasets and stricter benchmarking protocols are required.
26	Lessmeier et al. [26]	The dataset provides richer damage types and operating conditions than simpler seeded-fault benchmarks.	Field degradation datasets with richer process context and standardized split protocols are needed.



International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Sr. No.	Author(s)	Key Findings	Research Gap
27	Nectoux et al. [27]	The platform became a central resource for health indicators and RUL prediction.	More realistic, multimodal, and intervention-aware run-to-failure datasets are needed.
28	Lei et al. [28]	The dataset supplies complete degradation trajectories across multiple operating conditions.	More operating diversity, process metadata, and standardized RUL protocols are needed.
29	Saxena et al. [29]	C-MAPSS became the dominant multivariate benchmark for data-driven RUL prediction.	Field data, uncertainty-aware scoring, and explainable prognostics under richer operational variability are needed.
30	Purohit et al. [30]	MIMII established a key benchmark for acoustic machine anomaly detection beyond vibration-only monitoring.	Synchronized multimodal datasets, stronger domain-shift protocols, and interpretable acoustic diagnostics are needed.

V. HIGH-QUALITY REVIEW PAPERS

The review papers below are especially useful for framing a publication-level manuscript because they organize the field, define recurring vocabulary, and expose gaps that cannot be inferred reliably from one experimental study. Their value is complementary: some are maintenance-lifecycle reviews, others are method-centered, vibration-centered, edge-centered, or explainability-centered.

TABLE.II. HIGH-QUALITY REVIEW PAPERS

Ref.	Title	Coverage	Principal strengths	Principal limitations
[1]	A Systematic Literature Review of Machine Learning Methods Applied to Predictive Maintenance (2019)	Theme 2 - Machine Learning for Predictive Maintenance	Foundational ML taxonomy; strong baseline for method evolution.	Predates transformers, TinyML, and current XAI practice.
[2]	Predictive Maintenance in the Industry 4.0: A Systematic Literature Review (2020)	Theme 1 - Predictive Maintenance in Industry 4.0	End-to-end Industry 4.0 architecture framing.	Limited evidence on modern trustworthy edge deployment.
[3]	Predictive Maintenance and Intelligent Sensors in Smart Factory: Review (2021)	Theme 8 - Condition Monitoring	Connects sensing choices with smart-factory maintenance.	Broad scope; modest algorithm-level depth.
[4]	On Predictive Maintenance in Industry 4.0: Overview, Models, and Challenges (2022)	Theme 1 - Predictive Maintenance in Industry 4.0	Clear lifecycle treatment of CBM, PHM, and RUL.	No common quantitative benchmark.
[5]	Challenges in Predictive Maintenance - A Review (2023)	Theme 1 - Predictive Maintenance in Industry 4.0	Strong emphasis on implementation barriers.	Review findings are not converted into a standard implementation protocol.
[6]	Artificial Intelligence for Predictive Maintenance Applications: Key Components, Trustworthiness, and Future Trends (2024)	Theme 1 - Predictive Maintenance in Industry 4.0	Current trustworthiness and future-technology perspective.	Breadth limits method-by-method empirical comparison.
[7]	Remaining Useful Life Prediction Using Deep Learning Approaches: A Review (2020)	Theme 2 - Machine Learning for Predictive	Focused entry point for deep-learning RUL research.	Predates recent attention, transformers, and edge RUL.



International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Ref.	Title	Coverage	Principal strengths	Principal limitations
		Maintenance		
[8]	A Systematic Review of Rolling Bearing Fault Diagnoses Based on Deep Learning and Transfer Learning: Taxonomy, Overview, Application, Open Challenges, Weaknesses and Recommendations (2023)	Theme 3 - Deep Learning for Bearing Fault Diagnosis	Detailed bearing DL and transfer-learning taxonomy.	Most surveyed evidence is laboratory-based.
[9]	Recent Advances in the Application of Deep Learning for Fault Diagnosis of Rotating Machinery Using Vibration Signals (2023)	Theme 9 - Vibration-Based Fault Diagnosis	Strong synthesis of vibration-based deep diagnosis.	Inconsistent source-study reporting constrains comparison.
[16]	Possible Applications of Edge Computing in the Manufacturing Industry - Systematic Literature Review (2022)	Theme 4 - Edge AI and Edge Computing	Manufacturing-specific edge-computing map.	Broad manufacturing scope; maintenance architectures are not benchmarked directly.
[22]	An Interrogative Survey of Explainable AI in Manufacturing (2024)	Theme 5 - Explainable Artificial Intelligence	Role- and question-oriented manufacturing-XAI taxonomy.	Industrial human studies and explanation metrics remain scarce.
[24]	Internet of Things for Smart Factories in Industry 4.0: A Review (2023)	Theme 6 - Industrial Internet of Things	Broad IIoT infrastructure context for smart factories.	Not specific enough to serve as the sole PdM review.

VI. BENCHMARK DATASETS USED IN PREDICTIVE MAINTENANCE

Benchmark datasets make model comparison possible, but each dataset encodes a specific experimental reality. The following table should therefore be read as a guide to appropriate use rather than a ranking. IMS and SEU are included because they are widely used in the literature; their source papers are not added to the reference list so that the corpus remains exactly 30 papers.

TABLE.III. BENCHMARK DATASETS USED IN PREDICTIVE MAINTENANCE

Dataset	Purpose	Sensor modalities	Common tasks	Advantages	Limitations	Primary source
CWRU	Seeded bearing-fault classification	Drive-end/fan-end accelerometer vibration; motor load/speed metadata	Fault classification, feature learning, transfer learning	Extremely common; easy access; supports reproducible comparisons	Seeded laboratory faults; often too easy; leakage-prone windowing; weak field realism	[25]
Paderborn	Bearing-damage classification across artificial and real damage	Vibration, motor current, speed/load and damage metadata	Classification, transfer learning, domain adaptation	Richer damage types and operating conditions than CWRU	Laboratory system; preprocessing and split conventions differ across studies	[26]



International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Dataset	Purpose	Sensor modalities	Common tasks	Advantages	Limitations	Primary source
IMS	Run-to-failure bearing degradation	Accelerometer vibration from endurance tests	Early fault detection, health indicators, prognosis, RUL	Natural degradation sequences; useful for progression analysis	Small number of runs; noisy signals; labels and failure stages are less standardized	Official NASA/IMS repository; source paper not included
FEMTO-ST / PRONOSTIA	Accelerated bearing degradation and prognostics	Horizontal/vertical vibration, temperature, speed, and force	RUL prediction, degradation modeling, PHM challenge evaluation	Canonical run-to-failure benchmark with controlled conditions	Accelerated degradation and challenge-specific scoring may not reflect field maintenance	[27]
XJTU-SY	Bearing accelerated life tests under multiple conditions	Vibration signals and operating-condition metadata	RUL, anomaly detection, degradation-stage recognition	Complete trajectories under three conditions; widely used in modern RUL studies	Accelerated tests, limited sensing, and relatively small asset diversity	[28]
SEU	Bearing and gearbox diagnosis under two operating conditions	Multi-sensor vibration from drivetrain dynamic simulator	Fault classification, domain adaptation, multi-sensor fusion	Includes bearings and gears; supports cross-condition evaluation	Laboratory faults; no universally adopted split protocol; metadata usage varies	Source paper not included in the 30-paper corpus
NASA C-MAPSS	Simulated turbofan degradation and RUL	Multivariate engine sensor channels and operating settings	RUL regression, sequence modeling, uncertainty estimation	Mature benchmark; multivariate; supports standardized train/test subsets	Simulation assumptions; benchmark overuse; limited direct field equivalence	[29]
MIMII	Industrial acoustic anomaly detection	Microphone recordings for valves, pumps, fans, and slide rails	Unsupervised/semi-supervised anomaly detection and domain shift	Real factory acoustics; broadens monitoring beyond vibration	Audio-only; environment and machine-specific domain shifts	[30]
PHM Challenge Datasets	Community benchmarking for prognostics and health management	Varies by challenge; includes vibration, temperature, current, and multivariate process data	RUL, diagnosis, anomaly detection, prognostic scoring	Shared tasks encourage comparable evaluation and reusable baselines	Challenge metrics may encourage overfitting; datasets often remain controlled or accelerated	[27], [29]



International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

VII. COMPARATIVE INTERPRETATION OF BENCHMARK SUITABILITY

Dataset choice determines what a model's results can demonstrate. CWRU and Paderborn mainly support fault diagnosis, while IMS, PRONOSTIA, XJTU-SY, and C-MAPSS are suited to degradation and RUL studies; MIMII focuses on acoustic anomaly detection. Results from seeded-fault or simulated datasets should therefore not be assumed to represent direct industrial performance.

Industrial conditions introduce variability from sensor placement, maintenance, component changes, environmental noise, and operating schedules. Hence, future studies should use machine-, campaign-, or time-based splits and cross-dataset evaluation to assess generalization.

Benchmark studies should report both predictive and computational metrics, including preprocessing, split strategy, model size, memory, latency, energy, and explanation stability.

VIII. MOST INFLUENTIAL PAPERS

The following 20 papers are the strongest anchor citations within the selected corpus. They were chosen for foundational influence, benchmark importance, methodological direction, or their ability to frame a critical review rather than merely for recency.

- [1] A Systematic Literature Review of Machine Learning Methods Applied to Predictive Maintenance - Foundational machine-learning taxonomy for PdM.
- [2] Predictive Maintenance in the Industry 4.0: A Systematic Literature Review - Core Industry 4.0 PdM architecture review.
- [4] On Predictive Maintenance in Industry 4.0: Overview, Models, and Challenges - Strong lifecycle framing across CBM, PHM, and RUL.
- [5] Challenges in Predictive Maintenance - A Review - Clear synthesis of implementation challenges.
- [6] Artificial Intelligence for Predictive Maintenance Applications: Key Components, Trustworthiness, and Future Trends - Trustworthiness and future-technology agenda.
- [8] A Systematic Review of Rolling Bearing Fault Diagnoses Based on Deep Learning and Transfer Learning: Taxonomy, Overview, Application, Open Challenges, Weaknesses and Recommendations - Key DL and transfer-learning bearing review.
- [9] Recent Advances in the Application of Deep Learning for Fault Diagnosis of Rotating Machinery Using Vibration Signals
- Major review of vibration-based deep diagnosis.
- [11] A Novel Time-Frequency Transformer Based on Self-Attention Mechanism and Its Application in Fault Diagnosis of Rolling Bearings - Representative transformer shift in vibration diagnosis.
- [12] Diagnosisformer: An Efficient Rolling Bearing Fault Diagnosis Method Based on Improved Transformer - Influential efficient transformer-based bearing model.
- [13] Enhanced Transfer Learning Method for Rolling Bearing Fault Diagnosis Based on Linear Superposition Network - Important cross-domain transfer-learning contribution.
- [16] Possible Applications of Edge Computing in the Manufacturing Industry - Systematic Literature Review - Manufacturing-specific edge-computing synthesis.
- [18] TinyML-Enabled Edge Implementation of Transfer Learning Framework for Domain Generalization in Machine Fault Diagnosis - Strong evidence of TinyML and real constrained deployment.
- [20] An Explainable AI-Based Fault Diagnosis Model for Bearings - Early interpretable bearing diagnosis with Shapley analysis.
- [21] Explainable AI in Manufacturing: A Predictive Maintenance Case Study - Early manufacturing predictive-maintenance XAI case.
- [22] An Interrogative Survey of Explainable AI in Manufacturing - Comprehensive question-driven manufacturing-XAI survey.
- [24] Internet of Things for Smart Factories in Industry 4.0: A Review - High-impact IIoT smart-factory infrastructure review.
- [25] Rolling Element Bearing Diagnostics Using the Case Western Reserve University Data: A Benchmark Study - Essential critique of CWRU benchmark practice.



International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

- [26] Condition Monitoring of Bearing Damage in Electromechanical Drive Systems by Using Motor Current Signals of Electric Motors: A Benchmark Data Set for Data-Driven Classification - Canonical Paderborn dataset publication.
- [27] PRONOSTIA: An Experimental Platform for Bearings Accelerated Degradation Tests - Foundational PRONOSTIA/FEMTO-ST run-to-failure platform.
- [29] Damage Propagation Modeling for Aircraft Engine Run-to-Failure Simulation - Defining C-MAPSS run-to-failure simulation paper.

IX. CRITICAL TRENDS AND RESEARCH GAPS

9.1. Edge Deployment and Resource Efficiency: Existing studies demonstrate low-latency local inference, but practical deployment still requires reliable sensing, secure model updates, drift monitoring, resource-aware optimization, and energy evaluation. Lightweight models should be assessed using accuracy, memory, latency, and energy together.

9.2. Explainability and Human-Centered AI: XAI methods can connect predictions with diagnostic features, but limited evidence exists on whether explanations improve fault localization, troubleshooting, or operator trust. Future work should provide stable, role-specific, and actionable explanations.

9.3. Generalization and Multi-Sensor Learning: Many studies rely on controlled datasets and fixed sensor conditions. Robust systems should address domain shift, sensor failures, asynchronous data, continual learning, and federated learning across heterogeneous machines.

9.4. Digital Twins, Data and Benchmarking: Digital twins and physics-informed methods remain promising but require stronger validation. Benchmark datasets also need machine-level splits, cross-condition testing, richer metadata, and realistic field data to improve external validity.

X. FUTURE RESEARCH AND EVALUATION FRAMEWORK

Future research should focus on lightweight models, cross-machine and cross-dataset validation, robust multi-sensor fusion, secure edge updates, drift detection, uncertainty estimation, and reproducible reporting. Evaluation should progress from leakage-resistant testing within a dataset to unseen operating conditions, cross-machine/dataset transfer, and continuous edge-hardware testing.

A common reporting protocol should include data-split strategy, preprocessing, class balance, model configuration, confidence intervals, memory, latency, energy, and quantization effects. XAI studies should additionally report explanation fidelity, stability, computational cost, and intended user role.

10.1. Human-Centered and Deployment Validation

Maintenance intelligence should be evaluated through its **impact on maintenance decisions**, including diagnosis time, fault localization, inspection effort, operator confidence, false alerts, and overrides. Reporting deployment maturity—from offline benchmarks to continuous plant operation—can distinguish experimental prototypes from industrial systems.

10.2. Reproducibility, Safety and Cost-Aware Evaluation

Reproducible studies should provide dataset partitions, preprocessing code, model settings, random seeds, and deployment configurations. Systems should explicitly handle uncertainty, missing data, sensor failures, and model disagreement through safe fallback or human/cloud review. Evaluation should also consider maintenance cost, lead time, inspection burden, and the consequences of false and missed alarms, while reporting failure and boundary conditions rather than only successful benchmark results.



International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

TABLE.IV. CONSOLIDATED RESEARCH AGENDA

Gap category	Publication-oriented research direction
Edge AI deployment	End-to-end edge pipeline with acquisition reliability, model versioning, drift monitoring, and CMMS integration.
Explainability	Role-specific, faithful, stable explanations evaluated through maintenance decisions and human studies.
Lightweight models	Joint optimization of accuracy, memory, latency, energy, and minority-fault performance.
Real-time inference	Latency distributions and worst-case timing under realistic sensor and communication load.
TinyML	Open microcontroller benchmarks, safe updating, and explainability under kilobyte-scale constraints.
Energy efficiency	Energy per sample/window/inference and adaptive duty-cycling strategies.
Edge-cloud collaboration	Dynamic partitioning, resilient offline operation, secure synchronization, and fallback logic.
Multi-sensor fusion	Asynchronous and missing-modality fusion across vibration, current, temperature, sound, and process data.
Federated learning	Non-IID industrial assets, privacy, communication cost, poisoning resilience, and safety assurance.
Digital twins	Validated closed-loop links among sensing, simulation, diagnosis, prognosis, and maintenance action.
Dataset limitations	Field-scale run-to-failure and post-maintenance datasets with rich metadata and realistic imbalance.
Standard benchmarking	Machine-level splits, cross-condition tests, reproducible preprocessing, and resource reporting.

The agenda is intentionally systems-oriented. It treats data acquisition, model design, embedded execution, explanation, and maintenance integration as a connected research problem. A publishable contribution should therefore state which layer it advances and should report the trade-offs introduced at adjacent layers.

10.3. Suggested Design of Future Experimental Studies

A strong experimental study should begin with a clearly defined maintenance decision, not only a classification target. Data should be separated at machine or campaign level, and at least one out-of-domain condition should be reserved for final evaluation. The selected model should be compared with a lightweight classical baseline, a compact deep baseline, and an accuracy-oriented upper bound. This design exposes whether additional complexity creates a meaningful operational benefit.

10.4. Edge and Explainability Co-Evaluation

Edge efficiency and explainability should be evaluated together because post-hoc explanation can consume more time and memory than inference itself. Studies should report whether explanations are computed for every window, only for confirmed anomalies, or on a nearby gateway. A useful design may combine a compact always-on detector with a more detailed diagnostic and explanation stage that is activated only when persistence logic confirms an abnormal condition.

10.5. Evidence Expected for Publication

Publication-level evidence should include reproducible code or pseudocode, exact dataset partitions, confidence intervals across repeated runs, and clear limitations. For hardware studies, continuous-run logs and failed-read statistics are as important as a single latency value. For human-centered studies, participant expertise and task design should be



International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

reported. These details allow reviewers to distinguish a technically promising prototype from a dependable maintenance system.

XI. CONCLUSION

The review of 30 published studies shows a progression from machine learning to deep learning, transfer learning, attention, transformers, Edge AI, TinyML, XAI. CNNs remain effective for bearing and vibration diagnosis; transfer learning improves cross-condition adaptation; transformers capture long-range signal relationships; augmentation helps with limited data; reinforcement learning can address imbalanced faults. Edge and TinyML demonstrate low-latency local inference, while XAI links predictions to diagnostic features and maintenance reasoning. However, important limitations remain across the literature. Many studies rely on laboratory, simulated, accelerated, or domain-specific datasets, sensors, overlapping windows, and inconsistent validation. Cross-machine generalization, uncertainty estimation, drift handling, secure model updates, energy reporting and long-term field validation remain limited. XAI studies rarely measure operator trust or maintenance. Future systems should combine uncertainty-aware models, multi-sensor data, leakage-resistant validation, secure edge deployment, faithful explanations, standardized reporting of accuracy, latency, memory, and energy. These findings position trustworthy edge intelligence as a practical path toward dependable industrial predictive maintenance.

REFERENCES

- [1] T. P. Carvalho *et al.*, "A systematic literature review of machine learning methods applied to predictive maintenance," *Computers & Industrial Engineering*, 2019, Art. no. 106024.
- [2] T. Zonta *et al.*, "Predictive maintenance in the Industry 4.0: A systematic literature review," *Computers & Industrial Engineering*, 2020, Art. no. 106889.
- [3] M. Pech *et al.*, "Predictive maintenance and intelligent sensors in smart factory: Review," *Sensors*, vol. 21, 2021, Art. no. 1470.
- [4] M. Achouch *et al.*, "On predictive maintenance in Industry 4.0: Overview, models, and challenges," *Applied Sciences*, vol. 12, 2022, Art. no. 8081.
- [5] P. Nunes *et al.*, "Challenges in predictive maintenance—A review," *CIRP Journal of Manufacturing Science and Technology*, vol. 40, 2023, pp. 53–67.
- [6] A. Ucar *et al.*, "Artificial intelligence for predictive maintenance applications: Key components, trustworthiness, and future trends," *Applied Sciences*, vol. 14, 2024, Art. no. 898.
- [7] "Remaining useful life prediction using deep learning approaches: A review," *Procedia Manufacturing*, vol. 49, 2020, pp. 81–88.
- [8] M. Hakim *et al.*, "A systematic review of rolling bearing fault diagnoses based on deep learning and transfer learning," *Ain Shams Engineering Journal*, vol. 14, 2023, Art. no. 101945.
- [9] B. A. Tama *et al.*, "Recent advances in the application of deep learning for fault diagnosis of rotating machinery using vibration signals," *Artificial Intelligence Review*, vol. 56, 2023, pp. 4667–4709.
- [10] R. M. Souza *et al.*, "Deep learning for diagnosis and classification of faults in industrial rotating machinery," *Computers & Industrial Engineering*, vol. 153, 2021, Art. no. 107060.
- [11] Y. Ding *et al.*, "A novel time-frequency Transformer based on self-attention mechanism and its application in fault diagnosis of rolling bearings," *Mechanical Systems and Signal Processing*, 2022, Art. no. 108616.
- [12] Y. Hou *et al.*, "Diagnosisformer: An efficient rolling bearing fault diagnosis method based on improved Transformer," *Engineering Applications of Artificial Intelligence*, vol. 124, 2023, Art. no. 106507.
- [13] C. Huo *et al.*, "Enhanced transfer learning method for rolling bearing fault diagnosis based on linear superposition network," *Engineering Applications of Artificial Intelligence*, vol. 121, 2023, Art. no. 105970.
- [14] H. Wang *et al.*, "FTGAN: A novel GAN-based data augmentation method coupled time-frequency domain for imbalanced bearing fault diagnosis," *IEEE Transactions on Instrumentation and Measurement*, vol. 72, 2023.
- [15] Y. Li *et al.*, "A deep reinforcement learning-based intelligent fault diagnosis framework for rolling bearings under imbalanced datasets," *Control Engineering Practice*, vol. 145, 2024, Art. no. 105845.
- [16] K. Kubiak *et al.*, "Possible applications of edge computing in the manufacturing industry—Systematic literature review," *Sensors*, vol. 22, 2022, Art. no. 2445.
- [17] Y. Huang *et al.*, "Edge computing and fault diagnosis of rotating machinery based on MobileNet in wireless sensor networks for mechanical vibration," *Sensors*, vol. 24, 2024, Art. no. 5156.



International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

- [18] S. Asutkar *et al.*, “TinyML-enabled edge implementation of transfer learning framework for domain generalization in machine fault diagnosis,” *Expert Systems with Applications*, vol. 213, 2023, Art. no. 119016.
- [19] Z. Gao *et al.*, “An edge-deployable TinyML approach enhanced by transfer learning for efficient bearing fault diagnosis,” 2025.
- [20] M. J. Hasan *et al.*, “An explainable AI-based fault diagnosis model for bearings,” *Sensors*, vol. 21, 2021, Art. no. 4070.
- [21] B. Hrnjica and S. Softic, “Explainable AI in manufacturing: A predictive maintenance case study,” 2020.
- [22] Z. Alexander *et al.*, “An interrogative survey of explainable AI in manufacturing,” *IEEE Transactions on Industrial Informatics*, vol. 20, 2024, pp. 7069–7081.
- [23] M. Cakir *et al.*, “The experimental application of popular machine learning algorithms on predictive maintenance and the design of IIoT based condition monitoring system,” 2021.
- [24] M. Soori *et al.*, “Internet of things for smart factories in Industry 4.0: A review,” *Internet of Things and Cyber-Physical Systems*, vol. 3, 2023, pp. 192–204.
- [25] W. A. Smith and R. B. Randall, “Rolling element bearing diagnostics using the Case Western Reserve University data: A benchmark study,” *Mechanical Systems and Signal Processing*, vols. 64–65, 2015, pp. 100–131.
- [26] C. Lessmeier *et al.*, “Condition monitoring of bearing damage in electromechanical drive systems by using motor current signals of electric motors: A benchmark data set for data-driven classification,” *PHM Society Conference*, 2016.
- [27] P. Nectoux *et al.*, “PRONOSTIA: An experimental platform for bearings accelerated degradation tests,” *IEEE PHM*, 2012.
- [28] Y. Lei *et al.*, “XJTU-SY rolling element bearing accelerated life test datasets: A tutorial,” *Journal of Mechanical Engineering*, vol. 55, 2019, pp. 1–6.
- [29] A. Saxena *et al.*, “Damage propagation modeling for aircraft engine run-to-failure simulation,” *IEEE International Conference on Prognostics and Health Management*, 2008, pp. 1–9.
- [30] H. Purohit *et al.*, “MIMII dataset: Sound dataset for malfunctioning industrial machine investigation and inspection,” *DCASE Workshop*, 2019.



INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

 9940 572 462  6381 907 438  ijircce@gmail.com



www.ijircce.com

Scan to save the contact details