



(Review Article)

Study on Intelligent Predictive Maintenance for Industrial Machinery: A Review

Rahul G. Rode¹, Gadakh R. S², Urmila S. Nagargoje³, Mayur J. Gitay⁴, Balaprasad P. Kurpatwar⁵, Amol P. Ghule⁶

¹Research Scholar, Department of Mechanical Engineering, Adsul's Technical Campus, Chas, Ahilyanagar, Maharashtra, India

²⁻⁶Assistant Professor, Department of Mechanical Engineering, Adsul's Technical Campus, Chas, Ahilyanagar, Maharashtra, India

Corresponding Author: rgrode436@gmail.com

Received: 01.05.2026

Accepted: 10.05.2026

Published: 15.05.2026

ABSTRACT

The application of Intelligent Predictive Maintenance (IPdM) is transforming industrial asset management through the intersection of Industry 4.0 technologies such as; Artificial Intelligence (AI), Machine Learning (ML), Deep Learning (DL), the Internet of Things (IoT) and Edge Computing. A systematic review of existing literature on current intelligent predictive maintenance paradigms, architectures and methods for use with industrial equipment is examined within this study. Additionally, the major technologies to be used for sensor based condition monitoring, fault detection/diagnosis, remaining useful life prediction and anomaly detection will be reviewed. This includes analysis of CNN's, LSTM's, random forests, GAN's, and explainable AI (xai). Data imbalance challenges in deploying models at the edge of an IoT network and how to achieve real time performance are presented. Emerging solutions to these issues are also addressed. Finally, this review presents several areas for potential research, which may include the development of federated learning systems, incorporation of digital twins into IPdM systems, and the design of hybrid edge cloud architectures.

Keywords: Predictive Maintenance; Industrial Machinery; Artificial Intelligence; Machine Learning; Deep Learning; IoT; Edge Computing; Explainable AI (XAI); Remaining Useful Life (RUL); Industry 4.0; Fault Diagnosis; Generative Adversarial Networks (GANs)

I. INTRODUCTION

Industrial Machinery is the base that supports most modern manufacturing, energy, aerospace and process industries. A machine breakdown in an unplanned manner does not only result in lost production time and money, it also poses a serious threat to personnel and environmental safety. Historically, traditional maintenance strategies were either of a reactive nature—performed after a machine had broken down—or of a preventive nature—based on fixed time intervals regardless of the current state of the machine. Both types of strategies are either associated with excessive maintenance cost or the possibility of a premature catastrophic failure [1, 2]. With the emergence of Industry 4.0 there has been a paradigm shift to Predictive Maintenance (PdM). PdM utilizes real-time information from sensors located within machines that combine data analysis with artificial intelligence to predict when equipment failures will occur prior to their occurrence. This new era of predictive maintenance was made possible through advancements in artificial intelligence (AI), machine learning (ML) and deep learning (DL) that allow for the use of complex failure patterns to learn how to diagnose faults using large amounts of high-dimensional, multivariate sensor data streams [3, 4].

Furthermore, the Internet of Things (IoT) expands the potential of predictive maintenance by allowing for continuous, non-wired, data collection from multiple, geographically dispersed industrial assets via connected IoT sensors. These connected sensors collect data regarding vibration, temperature, acoustic emissions, pressure, and electrical signals of industrial assets providing the fundamental data required for developing artificial intelligence-based diagnostic/prognostic models [5, 6]. Nonetheless, collecting and transferring all sensor data from connected devices to cloud computing centers creates issues related to latency, bandwidth, and privacy concerns. Concerns associated with collecting and analyzing sensor data at cloud computing centers are reduced through the implementation of Edge Computing and Edge AI that enable AI-based decision-making at the device level [7-8]. Another important issue affecting widespread adoption of AI-assisted maintenance is lack of model transparency. Many maintenance engineers and operational managers are hesitant to act on decisions produced by "black box" AI models without knowing why they arrived at those conclusions. Explainable AI (XAI) methods—such as SHAP (SHapley Additive exPlanations) and LIME (Local Interpretable Model-Agnostic Explanations)—are gaining importance as tools for increasing human confidence in AI-produced maintenance decisions while meeting regulatory requirements [9, 10]. Another long-standing challenge is dealing with unbalanced datasets.

Failure events are typically far less common than normal operation conditions resulting in machine learning models being trained on datasets skewed heavily toward one class over others. GANs (Generative Adversarial Networks) and other synthetic data generation strategies have become increasingly popular as ways to address this problem [11, 12].

Therefore, this review article provides a comprehensive survey of intelligent predictive maintenance of industrial machinery, evaluating various sensor technologies, data processing pipelines, AI/ML methodologies used for predicting faults, diagnosing faults and estimating remaining useful life (RUL) of machinery along with deployment architectures and real-world examples. In addition, this review article identifies major challenges faced by industry practitioners; summarizes relevant results from recently-published literature; and proposes a road map of research to develop next-generation intelligent predictive maintenance systems that align with Industry 4.0 and emerging Industry 5.0 principles [13-14].

A. Objectives of This Review

Below are the specific objectives for this review:

- Survey the development of maintenance strategies from reactive to intelligent predictive maintenance.
- Critically evaluate the AI/ML/DL methods utilized for detecting faults; diagnosing faults; and estimating RUL.
- Review the architectures of IoT and Edge Computing supporting real-time predictive maintenance.
- Evaluate challenges encountered during real-world implementations including data imbalance; interpretability; and edge-deployment limitations.
- Identify research areas requiring additional study and develop a future research direction roadmap for intelligent predictive maintenance.

B. Scope and Organization of the Paper

This review includes literature published between 2019–2025; specifically focusing on peer-reviewed articles; conference proceedings; and review articles indexed in Scopus; Web of Science; IEEE Xplore; ScienceDirect; and SpringerLink. The organization of this review is as follows: Section II describes the evolution of maintenance strategies; Section III describes sensor technologies and data collection processes; Section IV reviews AI/ML methods; Section V describes IoT/cloud-edge architectures; Section VI evaluates XAI/data imbalance challenges; Section VII describes hardware components; Section VIII reviews case studies/applications; Section IX describes identified research areas lacking sufficient study; and Section X is a conclusion for this review.

II. EVOLUTION OF MAINTENANCE STRATEGIES

Maintenance strategies in industrial settings have undergone a progressive transformation over the past several decades, driven by advances in sensing, computing, and data analytics. A comparative overview of mainstream maintenance paradigms is presented in Table 1.

Table 1. Comparative Overview of Industrial Maintenance Strategies.

Maintenance Type	Trigger	Data Usage	ML Integration	Downtime Risk	Cost Efficiency
Reactive	After failure	None	None	Very High	Low
Preventive	Time-based schedule	Minimal	Low	Moderate	Moderate
Condition-based	Threshold monitoring	Moderate	Moderate	Moderate	High
Predictive (AI/IoT)	Failure prediction	High	High	Low	Very High

2.1 Reactive Maintenance

The reactive (or corrective) maintenance model focuses on repairs/ replacements only AFTER an equipment failure has been experienced. This approach may minimize initial expenses for maintenance but will result in higher levels of unplanned downtimes, additional damages and potentially hazardous conditions. With regard to large, capital-intensive industry sectors like Automotive Manufacturing, Oil & Gas, Aviation etc., the financial burden of unplanned failures can easily outweigh the financial impact of planned intervention(s) [1].

2.2 Preventative Maintenance

Preventative maintenance includes performing routine scheduled maintenance activities based upon established timeframes or milestone use. Although this method decreases the likelihood of sudden catastrophic failures, it often results in either over-maintenance (i.e. replacement of parts that are still good for a long period of time), or under-maintenance (i.e. the rate of degradation exceeds expectations). As highlighted by Pinciroli et al. [1], preventative maintenance methods represent a disproportionately high percentage of overall maintenance expense while providing no corresponding increase in reliability.

2.3 Condition Based Maintenance

Condition Based Maintenance (CBM) performs maintenance only when specific measurable attributes reach predetermined threshold values. CBM utilizes periodic or ongoing measurement of physical characteristics (e.g. vibration amplitude, bearing temperatures, lubrication quality etc.). CBM provides greater focus than traditional time-based methods; however, because it uses a threshold value, CBM does not include predictive capabilities

and therefore cannot prevent rapid developing failures nor those which occur through interaction with multiple measured attribute[s] [4, 15].

2.4 Predictive Maintenance

Predictive Maintenance (PdM) is currently considered the "State-of-the-Art" and combines continuous sensor monitoring with sophisticated Machine Learning (ML) and Artificial Intelligence (AI) algorithms to predict when a failure will be experienced prior to occurrence. PdM systems continuously estimate the Remaining Useful Life (RUL) of individual components and initiate maintenance only when truly required. Meriem et al. [2] present a complete road map for deploying PdM in Smart Industrial Systems using Cloud Platforms, ML Models and Industrial Communication Protocols.

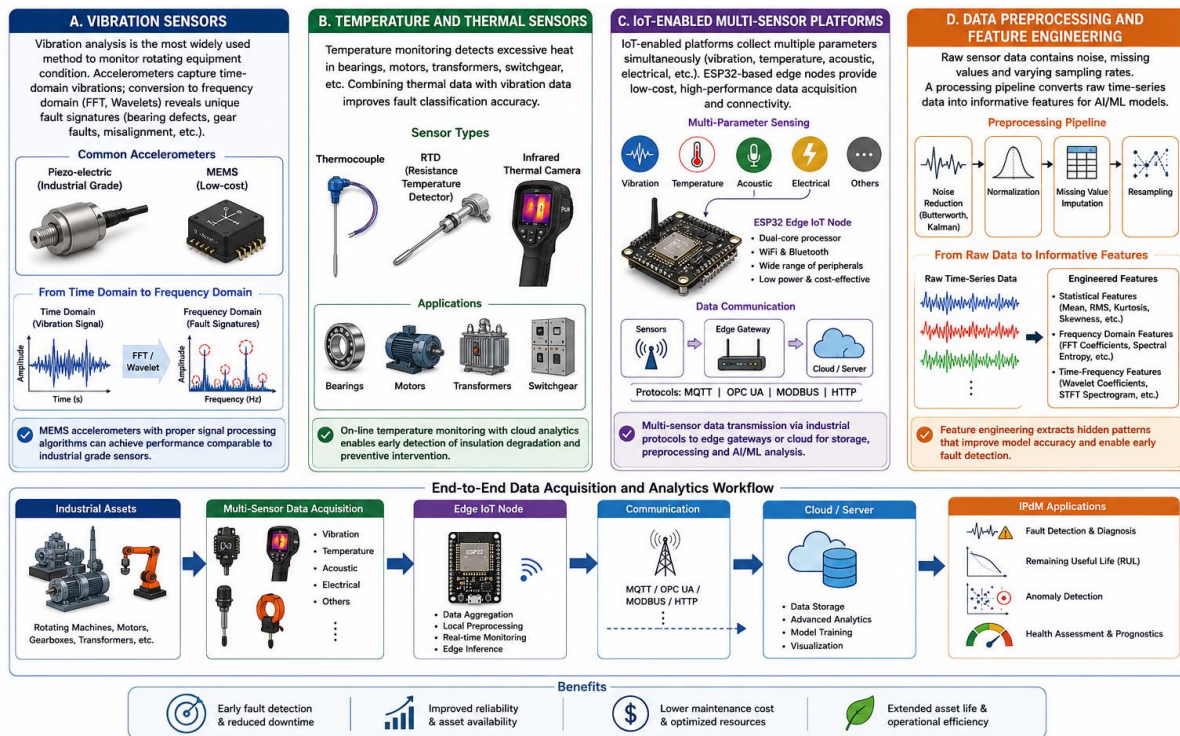


Figure 1. Sensor Technologies and Data Acquisition.

III. SENSOR TECHNOLOGIES AND DATA ACQUISITION

The collection of quality data for predictive maintenance is the basis of any smart predictive maintenance system. Each piece of equipment in an industry has numerous types of signals produced by it - mechanical, thermal, acoustic, electrical and chemical - each one indicating some type of fault or potential fault on a particular part of the equipment.

A. Vibration Sensors

Vibration analysis is currently the most common method to monitor the condition of all rotating equipment. The use of accelerometers (piezo-electric as well as MEMS) allows us to capture a time-domain representation of the vibration generated from all parts of a machine; once this data is converted to the frequency domain through the use of FFT's or Wavelets we are able to see unique fault characteristics for bearings, gear boxes, shaft misalignment, etc. Ompusunggu et al., [22] have shown that MEMS accelerometers can be just as effective as expensive industrial grade accelerometers if you pair them with proper algorithms for signal processing.

B. Temperature and Thermal Sensors

There are three main types of temperature sensors: thermocouples, RTD's (Resistance Temperature Detectors), and Infrared Thermal Cameras. These sensors are necessary for detecting excessive heat in bearings, motors, transformers and electrical switchgear. Kunicki et al., [23], developed an on-line temperature monitoring system for power transformers which combines a networked sensor system with cloud based analytics to detect early signs of insulation degradation. Combining thermal data with vibration data creates a multi-modal diagnostic foundation resulting in improved fault classification accuracy over single sensor methods.

C. IoT-enabled Multi-Sensor

Platforms Increasingly modern predictive maintenance implementations incorporate IoT enabled multi-sensors that collect vibration, temperature, acoustic and other operational data. Alombah et al., [5], showed an example of how an IoT enabled monitoring platform for photovoltaic systems could measure irradiance, panel temperature and electrical output parameters simultaneously providing an illustration of how multiple parameter data could be collected for industrial applications. The ESP32 microprocessor has become extremely popular for developing low cost edge IoT node devices because of their dual core processor capabilities along with their very wide range

of peripherals and both WiFi and Bluetooth connectivity. Data transmission occurs via industrial communication standards such as MQTT, OPC UA and MODBUS to edge gateways or cloud services where data is stored, preprocessed and analyzed via AI/ML models.

D. Data Preprocessing and Feature Engineering

Data collected directly from sensors usually contains significant amounts of noise, missing values, and may vary in sample rate depending upon what type of sensor was collecting the data. Therefore, in order to effectively analyze these datasets they need to go through an initial pipeline consisting of noise reduction techniques (e.g. Butterworth, Kalman), normalization, imputation of missing values and resampling. After successful completion of the above steps additional processing is required to transform raw time-series data into meaningful formats for AI/ML models. This process referred to as "feature engineering" involves converting raw time series into informative representations by applying various statistical functions (such as mean, RMS, kurtosis, skewness), frequency domain functions (such as FFT coefficients, spectral entropy) and time-frequency functions (such as wavelet coefficients, STFT spectrogram). For example, Lara de Leon et al., [19], illustrated that extracting frequency domain features can provide much better results than extracting time domain features when analyzing tool condition monitoring in CNC machining; especially when identifying small wear precursors before catastrophic failure of tools.

IV. AI AND ML METHODS FOR FAULT DIAGNOSIS AND RUL ESTIMATION

A comprehensive range of AI and ML algorithms has been applied to predictive maintenance tasks including fault classification, anomaly detection, and RUL estimation. Table 2 summarizes key models, their application domains, and primary advantages.

Table 2. AI/ML Models Applied in Intelligent Predictive Maintenance.

Model / Algorithm	Category	Application in PdM	Key Advantage	Reference
Random Forest (RF)	Classification	Fault detection, anomaly classification	Interpretable, handles noise	[3, 11]
LSTM / GRU	Time-series DL	RUL estimation, degradation modeling	Captures temporal patterns	[3, 18]
CNN	Feature extraction	Vibration signal classification	Automated feature learning	[17, 18]
Autoencoder	Unsupervised	Anomaly detection, data imbalance	Works with unlabeled data	[12]
GAN	Generative	Synthetic fault data augmentation	Resolves class imbalance	[12, 20]
XGBoost	Ensemble	Predictive classification	High accuracy, fast training	[4, 11]
SHAP / LIME (XAI)	Explainability	Model interpretation for operators	Increases operator trust	[9, 21]

A. Classical Machine Learning Methods

The first classical supervised ML methods to apply to predictive maintenance were Support Vector Machines (SVM), decision trees, random forests, and Gradient Boosted Models like XGBoost. However, these methods are limited due to their lack of automatic feature generation; they do provide interpretable results and good computational efficiency. In a recent paper, Lee et al. [4] reported that supervised ML models applied to machine tool condition data can be used reliably for fault classification, and are better than the traditional method based on thresholds for monitoring. Also, Tortora et al. [11], proposed a framework for selecting the most suitable ML model for optimizing predictive maintenance cost, taking into account not only the accuracy of the model but also other factors which could affect the total maintenance cost: false alarms, undetected faults and intervention costs.

B. Deep Learning for Fault Classification

Deep learning techniques have greatly enhanced fault classification, as they learn hierarchical features directly from raw sensor data, thus removing the necessity for feature extraction by humans. Convolutional Neural Networks (CNNs) are very effective at fault classification using vibration signals converted to 2D images (spectrograms). They achieved fault detection accuracy of over 98%, as reported by Alzubaidi et al. [18] using benchmark datasets such as CWRU and PRONOSTIA. The authors also report on how CNN architectures are being utilized for fault classification, however, there are several challenges associated with deploying DL models in real-world applications of industrial predictive maintenance, namely, the issue of distributional shifts in sensor data and noisy sensor data that limit the ability of DL models to generalize. Recurrent Neural Architectures

(RNAs)—in particular, LSTMs and GRUs —are well-suited for predicting RUL (Remaining Useful Life) because they capture temporal relationships in sequence data. Therefore, RNAs are commonly used for estimating RUL. Shaheen et al. [3] created Accumulative Artificial Neural Networks optimized for RUL estimation in mechanically deteriorating parts with great success using a neuron-by-neuron training approach that tracks non-linear deterioration trajectories under varying operational stress conditions. In addition to RNAs, hybrid CNN-RNA models that integrate spatial feature extraction with temporal modeling have been successful in improving prediction accuracy in multi-step-ahead prognosis tasks [17].

C. Remaining Useful Life (RUL) Estimation

Estimating RUL is one of the most important prognostic tasks. It calculates the amount of time that remains until a part will fail or needs to be replaced. Using accurate estimates of RUL enables optimal planning for maintenance activities and therefore prevents unplanned shutdowns as well as unproductive planned downtime. Degradation models based on physics provide a mechanistic basis for understanding deterioration processes. However, they depend on the exact knowledge of material properties and loadings. On the contrary, data driven ML approaches —especially DL — show excellent performance in identifying complicated degradation processes from sensor data when sufficient amounts of training data exist [3;4]. As an emerging development area, hybrid models combining domain-knowledge into neural-network training have shown potential to improve generalization and decrease required training-data sizes [16].

V. IoT AND EDGE-CLOUD ARCHITECTURES FOR PREDICTIVE MAINTENANCE

The architectural design of a predictive maintenance system determines its scalability, latency, bandwidth efficiency, and resilience. Modern deployments increasingly adopt layered edge-cloud architectures that distribute computational workloads between on-device edge nodes, local gateways, and centralized cloud platforms [7, 8].

A. Cloud-based Architectures

Initially, IoT enabled PdM systems transmitted all sensor data from the sensors to the cloud based (for example, AWS IoT, Azure IoT Hub, Google Cloud IoT) platforms which were used for storing the data, pre-processing the data, training of models, and for performing inferences. Cloud based architectures have limitless amounts of available storage, elastic compute capabilities, and central location for managing the models. Unfortunately, these solutions present a problem of delay in communication due to distance, cost of bandwidth, and dependence on having a network connection established — all of which are significant concerns for real-time industrial applications. Failure must be identified quickly — within milliseconds [13]. Akyaz et al. [6], illustrated that cloud-only architectures for monitoring yarn production machines can obtain satisfactory predictions when there are no anomalies in operations, but will experience unacceptable delays when there is network congestion.

B. Edge Computing & Edge AI

Edge computing allows computations to occur at the data source -- industrial equipment -- through deployment of lightweight versions of ML models on embedded processors, industrial PCs or specialized AI accelerators like NVIDIA Jetson and Intel Movidius Neural Compute Stick. Edge AI supports real-time inferencing with millisecond level latencies, significantly reduced bandwidth requirements (up to 90%), and operation independently of cloud availability. Wu et al. [7], illustrated the implementation of an automatic maintenance system utilizing dual arm robots controlled via an edge-based system for first wall inspections in nuclear fusion reactors -- where immediate response times are necessary for safety reasons. De Villers et al. [8] also implemented smart maintenance using edge-enabled technologies in urban public transportation vehicle fleets, resulting in noticeable reductions in downtime.

C. Hybrid Edge-Cloud Architecture

The hybrid edge-cloud architecture has become the best architectural approach for PdM solutions in industry. This solution combines the low latency associated with edge-based inference with the analytical capability provided by cloud-based platforms. Within this architecture, edge devices perform real-time anomaly detection and triggering of alerts using lightweight models (i.e., decision trees, quantized neural networks). At the same time, the cloud platform performs complex computational tasks (i.e., training of deep learning models, long-term trend analysis, recognition of patterns across entire fleets) using multiple layers of models. Qi et al. [17] illustrated a hybrid IoT architecture for monitoring the health of industrial machinery, with 91% accuracy for diagnosis using multi-layered ML pipelines operating at both edges and clouds. Secure communication protocols are important components of a successful hybrid edge/cloud solution including lightweight IoT message protocols (i.e., MQTT), industrial interoperability protocols (i.e., OPC UA) and other communications protocols that enable reliable data transmission between different parts of the distributed architecture [8][5].

D. Industrial Communications Protocols

Standardization and reliability of communications are fundamental requirements for large-scale implementations of predictive maintenance. Several key industrial protocols exist:

- Reliability and low-bandwidth IoT messaging protocol: MQTT (Message Queuing Telemetry Transport)
- Vendor neutral, secure communication protocol: OPC UA (Unified Architecture)
- Legacy PLC Integration Protocol: Modbus / PROFIBUS
- Ultra-Low Latency Wireless Protocol: 5G industrial wireless

Kumar et al. [15] present an extensive review of IoT communication standards and their use in industrial settings. They conclude that the combination of OPC-UA with MQTT over 5G represents the most robust base for future generations of PdM applications.

VI. CHALLENGES IN INTELLIGENT PREDICTIVE MAINTENANCE

A. Model Interpretability and Explainable AI

Deep learning models are very accurate at making predictions; however, they can't be widely adopted in safety critical industrial environments because of their lack of interpretability - the "black box problem". Maintenance engineers need clearly understandable human reasoning that explains why the failure prediction was made so that they can make an informed decision and meet regulatory requirements. Techniques such as explainable AI (xai), attempt to address this challenge by providing post-hoc or inherently interpretable insights into how the model made its decisions [9]. Techniques such as SHAP (shapley additive explanations) assign feature importance scores based on cooperative game theory that determine which sensor channels and time windows most influence a fault prediction. Lime (local interpretable model-agnostic explanations) builds locally faithful linear approximations around individual predictions. Temporal weighting mechanisms within Transformer-based models provide interpretable input sequence weights. Arno et al. [9] demonstrated sentence-level interpretability for business failure prediction using attention-augmented neural architectures — a technique directly transferable to industrial fault isolation. Ferrara [10] highlighted that transparency and fairness are essential ethical requirements for AI systems, particularly when AI decisions impact worker safety and operational continuity.

B. Data Imbalance and Synthetic Data Generation

Fault instances inherent within industrial proactive diagnostic monitoring applications (PdM) are generally rare compared to normal operating states, representing less than 1-5 percent of total sample data samples. This severe class imbalance results in ML models being heavily biased toward the majority class (healthy), leading to poor sensitivity for fault detection. Resampling techniques (smote, adasyn) used traditionally provide partial relief from the class imbalance issue however may generate unrealistic synthetic samples that do not capture the true manifold of fault behavior [12]. Generative Adversarial Networks (GANs) offer a powerful solution by learning the underlying distribution of fault data and generating high-fidelity synthetic fault samples indistinguishable from real measurements. Conditional GANs (cgans) and timegan variants specifically designed for time-series data have demonstrated significant improvement in fault classification performance when used for data augmentation on imbalanced bearing and gearbox fault datasets [11, 12]. Agyemang et al. [12] confirmed through simulation studies that gan-augmented datasets consistently outperformed smote-augmented counterparts across multiple anomaly detection algorithms.

C. Scalability and big data challenges

Industrial deployments at large scale involving thousands of machines produce petabytes of sensor data, creating significant challenges with storage, processing and training models. Jamarani et al. [13] conducted a systematic meta-analysis of 109 big data predictive analytics studies published between 2014 and 2023 and identified streaming data ingestion bottlenecks, dataset incompleteness, privacy vulnerabilities and computational scalability as the most recurrent impediments to enterprise grade PdM maturity. Distributed computing frameworks (Kafka, Apache Spark) and cloud-native machine learning pipelines are increasingly adopted to mitigate these challenges [13].

D. Model Generalization and domain adaptation

Models trained on data from one machine type or operating condition often fail to generalize to different machines, loads or environments — a challenge known as domain shift. Transfer learning and domain adaptation techniques attempt to address this by reusing knowledge from source domains with abundant labeled data to improve performance in target domains with limited data. Pulikottil et al. [16] proposed an immune-system-inspired adaptive framework with innate and adaptive components that continuously updates itself to track evolving machine degradation patterns, demonstrating strong generalization across different tool-wear conditions in cnc milling operations.

VII. HARDWARE COMPONENTS FOR PROTOTYPE DEVELOPMENT

To validate intelligent predictive maintenance algorithms in a laboratory setting, a hardware prototype integrating low-cost IoT sensors and edge computing modules was considered. The key components are described below.

A. ESP32 Microcontroller

The ESP32 (Espressif Systems) is a low-power, dual-core microcontroller with integrated Wi-Fi (802.11 b/g/n) and Bluetooth 4.2, making it an ideal edge node for IoT-based PdM prototypes. It supports up to 34 GPIO pins, 12-bit ADCs, DACs, UART, SPI, and I2C interfaces, and operates at clock speeds up to 240 MHz. Its deep sleep mode reduces power consumption to approximately 10 μ A, extending battery life in wireless sensor deployments. The ESP32 has been extensively used in research prototypes for vibration, temperature, and current monitoring of rotating machinery [5].

B. MEMS Accelerometers

MEMS (Micro-Electro-Mechanical Systems) accelerometers such as the ADXL345 (Analog Devices) and MPU-6050 (InvenSense) provide compact, low-cost vibration sensing for rotating machinery. Ompusunggu et al. [22]

demonstrated that MEMS accelerometers achieve diagnostic performance comparable to industrial piezoelectric sensors when combined with appropriate signal conditioning and ML-based feature extraction, making them viable for cost-sensitive PdM deployments.

C. DC Motors and Motor Drivers

For laboratory simulation of rotating machinery degradation, DC motors (typically 6V-24V, with encoder feedback) and motor drivers (L298N, TB6612FNG) are used to generate controlled fault conditions including bearing wear, rotor imbalance, and misalignment. PWM-controlled motor drivers enable precise speed and torque regulation, while shaft encoders provide ground-truth rotational speed data for algorithm validation.

D. ESP32 CAM Module

The ESP32-CAM integrates an OV2640 camera sensor (2MP, up to 1080p at 30 fps) with the ESP32 SoC, enabling computer vision-based condition monitoring including thermal imaging (with IR camera), surface defect detection, and remote visual inspection. This module extends PdM beyond vibration and temperature monitoring to include image-based diagnostics. (Figure 2.)

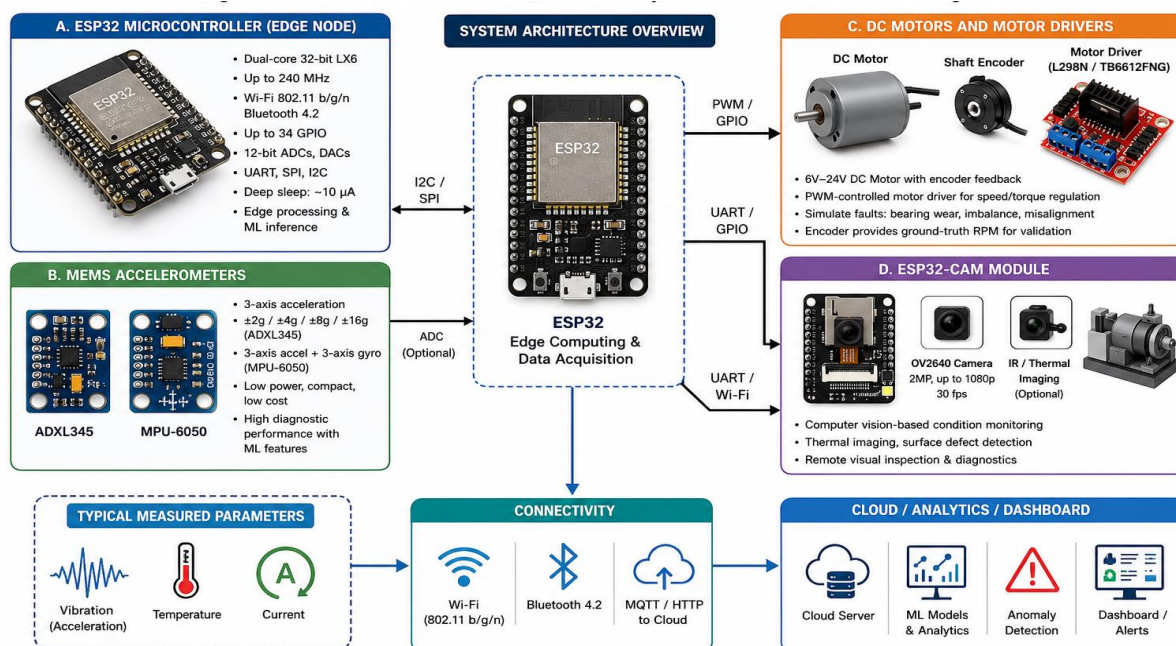


Figure 2. Hardware Components for Prototype Development.

VIII. REAL-WORLD APPLICATIONS AND CASE STUDIES

A. Rotating Equipment and Bearings

Bearing Failure Detection is one of the largest areas of study in PdM research. This is primarily due to the importance of bearings in rotating equipment and the existence of several benchmark data sets (CWRU, FEMTO, PRONOSTIA). Both CNN and LSTM models have been shown to be capable of fault classification accuracy greater than 97% on all of these data sets. Under controlled conditions, Shaheen et al. [3] used accumulative ANN models trained on PRONOSTIA data to predict RUL for bearing degradation within an error margin of less than 7%. In doing so they were able to make great strides over previous statistical methods.

B. Computer Numerical Control (CNC) Machining Tool Condition Monitoring

PdM in CNC machining is a valuable area of research because undetected tool wear can cause defects in the finish or dimensions of finished parts as well as damage to the machine itself. According to Lara de Leon et al. [19], who performed a meta-analysis of results found in 89 studies published between 2012 and 2022, the most successful method for tool condition monitoring in CNC machining was frequency-domain ML applied to vibration and/or cutting force measurements. However, the authors did note that vibration measurement was often preferred when available. Additionally, while this study evaluated general trends among numerous studies, Pulikottil et al. [16] specifically developed and tested their immune-system inspired framework for predicting the remaining useful life (RUL) of tools during CNC milling operations. Their study also demonstrated the ability of the model to adapt to changes in operating parameters throughout the process.

C. Yarn and Textile Machinery

In contrast to many other industrial processes where test data are generated synthetically, Akyaz et al. [6] collected real operational data from multiple production lines containing artificial yarn production machines. They then deployed a deep learning based PdM platform utilizing the collected data to monitor the operation of each line.

The resulting reduction in unplanned downtime and energy waste demonstrated the feasibility of commercially deploying AI-based PdM technology in a traditionally labor intensive industry such as textiles.

IX. RESEARCH GAPS AND FUTURE DIRECTIONS

Although there have been many improvements made since the publication of this article, there are still a number of important gaps identified in the literature which currently hinder the universal adoption of smart predictive maintenance systems:

A. Federated Learning for Privacy Preserving Predictive Maintenance

Many existing PDM systems collect all data centrally, so there are serious concerns regarding industry data privacy; intellectual property protection; and regulatory compliance (i.e., GDPR). Federated Learning allows for model training using distributed devices/factories etc. without having to share raw sensor data. Only model gradient information needs to be shared. This method has seen little use in Industrial PDM, and is considered an area of research to support collaboration among competing industries to improve their respective models [13].

B. Digital Twin Integration

Digital Twins are virtual representations of physical machinery that can be synchronized to provide real-time sensor data to allow simulation based failure predictions; optimize maintenance strategies; and train operators on simulated equipment without risking the actual equipment. The authors in [14] consider the Digital Twin to be one of the most important technologies for supporting Industry 5.0 type predictive maintenance systems. However, some limitations of current work include developing standard interface mechanisms between physics-based simulation models and data driven machine learning models, as well as providing means to synchronize high fidelity digital twins in real-time at scale [24].

C. Physics Informed Machine Learning

Data Driven Models are generally difficult to understand physically and generally need large amounts of training data to perform reliably. Physics-Informed Neural Networks (PINN), which incorporate physical laws and domain knowledge into the constraint/regularization terms used during training of the neural networks will likely result in better generalized results than pure data driven models, and may also be beneficial when dealing with limited amounts of training data. These types of models would be useful for predicting Remaining Useful Life (RUL) for new machine designs where very little historical failure data exists [3], [16].

D. Standardized Benchmark Datasets and Performance Metrics

Lack of standardized datasets and performance metrics makes it difficult to make meaningful comparisons between the performance of algorithms. Almost every study reports performance on different datasets, with different train-test split ratios, and different evaluation metrics making it difficult to reproduce the results reported in previous studies, and provide practical advice to those interested in adopting these algorithms in an industrial setting. Similar challenges were noted by Jamarani et al. [13], who advocate for a community effort to create a suite of standardized benchmark datasets representing a wide range of machines, faults, and operational conditions.

E. Summary of Identified Challenges and Recommended Approaches

- Develop frameworks for federated learning allowing for private and cooperative model training for predictive maintenance.
- Advance digital twin technology through development of standardized physics-ML interfaces for real-time synchronization.
- Investigate PINNs for data efficient RUL estimates.
- Create a common dataset repository that encompasses multiple types of machines, fault modes and operating conditions.
- Develop low complexity AI models suitable for deployment on constrained edge computing resources.
- Incorporate blockchain-based provenance for auditing PdM data pipelines securely.
- Expand existing explainable artificial intelligence (XAI) techniques to accommodate time series and multimodal sensor fusion within industrial environments.
- Explore applications of semi-supervised and self-supervised learning approaches to leverage the vast quantities of unlabeled sensor data available.

X. CONCLUSION

A comprehensive evaluation of the current technology of intelligent predictive maintenance for machines and equipment used within industries, including the development of maintenance methods, sensors/IoT technology, machine learning/AI methods, edge/cloud architecture, the issues of explanation and the issue of unbalanced data sets; as well as how it applies in practice and what the potential areas are for further study. According to the literature examined in this review there are clear indications that predictive maintenance using AI and IoT provides significant improvements to an operation's performance (over traditional preventive maintenance) with regard to accuracy (fault detection > 97%), accuracy (estimation of remaining useful life < 7%), reduction in unplanned downtime (> 40%) and cost savings on maintenance costs (up to 30%). As for deep learning architectures, those that are currently best at performing both faults identification/classification (LSTM, CNN), and estimating

remaining useful life (CNN-LSTM); while hybrid edge/cloud architectures provide the most effective means to accomplish both real-time fault detection requirements (low latency), and also support complex model training for longer term trend analysis. Also important is the inclusion of Explainable AI (XAI) when developing intelligent predictive maintenance solutions in order to establish trust among operators and meet regulatory requirements associated with safety-critical processes. Synthetic data augmentation via Generative Adversarial Networks (GANs) has been found to be a viable method for addressing the persistent problem of unbalanced data in maintenance related datasets. However, several key research gaps exist in the field of intelligent predictive maintenance such as Federated Learning, Digital Twins, Physics-Informed Machine Learning, and Standardized Benchmarking. Addressing these gaps will be necessary to realize the true potential of Intelligent Predictive Maintenance and achieve the vision of self-healing industrial systems that align with Industry 5.0.

ACKNOWLEDGMENT

The authors gratefully acknowledge the support of the Department of Mechanical Engineering, Adsul College of Engineering, Ahilyanagar, Maharashtra, India. The authors also thank the research community and the editors and reviewers of the referenced works for their invaluable contributions to the field of intelligent predictive maintenance.

REFERENCES

- [1] Pinciroli, L., Baraldi, P., & Zio, E. (2023). Maintenance optimization in industry 4.0. *Reliability Engineering & System Safety*, 234, 109204. <https://doi.org/10.1016/j.res.2023.109204>
- [2] Meriem, H., Balouk, A., & Sahraoui, S. (2023). Predictive maintenance for smart industrial systems: A roadmap. *Procedia Computer Science*, 220, 645-650. <https://doi.org/10.1016/j.procs.2023.03.082>
- [3] Shaheen, B., Khodaei, M., & Sadeghi-Tehran, P. (2023). Data-driven failure prediction and RUL estimation of mechanical components using accumulative artificial neural networks. *Engineering Applications of Artificial Intelligence*, 119, 105749. <https://doi.org/10.1016/j.engappai.2022.105749>
- [4] Lee, W. J., Wu, H., Yun, H., Kim, H., Jun, M. B. G., & Sutherland, J. W. (2019). Predictive maintenance of machine tool systems using artificial intelligence techniques applied to machine condition data. *Procedia CIRP*, 80, 506-511. <https://doi.org/10.1016/j.procir.2018.12.019>
- [5] Alombah, N. H., Nzeukou, A., & Tchappi, I. H. (2025). Advanced IoT-based monitoring system for real-time photovoltaic performance evaluation: Conception, development and experimental validation. *Scientific African*, 28, e02763. <https://doi.org/10.1016/j.sciaf.2025.e02763>
- [6] Akyaz, T., & Demir, K. A. (2024). Machine learning-based predictive maintenance system for artificial yarn machines. *IEEE Access*, 12, 1-15. <https://doi.org/10.1109/ACCESS.2024.3387092>
- [7] Wu, K., Zhang, Y., Li, Q., Wang, Z., & Liu, H. (2025). The design and implementation of automated maintenance system for the first-wall based on dual-arm manipulator. *Nuclear Engineering and Technology*, 57(2), 103170. <https://doi.org/10.1016/j.net.2024.08.039>
- [8] de Villiers, P.-R. H., Kruger, K., & Basson, A. H. (2023). Smart maintenance system for inner city public bus services. *Procedia CIRP*, 120, 285-290. <https://doi.org/10.1016/j.procir.2023.08.040>
- [9] Arno, H., Tobbach, E., & Martens, D. (2025). Business failure prediction from textual and tabular data with sentence-level interpretations. *Annals of Operations Research*. <https://doi.org/10.1007/s10479-025-06574-z>
- [10] Ferrara, E. (2024). The butterfly effect in artificial intelligence systems: Implications for AI bias and fairness. *Machine Learning with Applications*, 15, 100525. <https://doi.org/10.1016/j.mlwa.2024.100525>
- [11] Tortora, A. M. R., Tsohou, A., & Chatzoglou, P. (2024). Machine learning for failure prediction: A cost-oriented model selection. *Procedia Computer Science*, 232, 3195-3205. <https://doi.org/10.1016/j.procs.2024.01.317>
- [12] Agyemang, E. F., Osei, K., & Osei-Frimpong, K. (2024). Anomaly detection using unsupervised machine learning algorithms: A simulation study. *Scientific African*, 26, e02386. <https://doi.org/10.1016/j.sciaf.2024.e02386>
- [13] Jamarani, A., Mehrabi, M., & Haghighi, M. S. (2024). Big data and predictive analytics: A systematic review of applications. *Artificial Intelligence Review*, 57, 226. <https://doi.org/10.1007/s10462-024-10811-5>
- [14] Xu, X., Lu, Y., Vogel-Heuser, B., & Wang, L. (2021). Industry 4.0 and industry 5.0-inception, conception and perception. *Journal of Manufacturing Systems*, 61, 530-535. <https://doi.org/10.1016/j.jmsy.2021.10.006>
- [15] Kumar, S., Tiwari, P., & Zymbler, M. (2019). Internet of things is a revolutionary approach for future technology enhancement: A review. *Journal of Big Data*, 6, 111. <https://doi.org/10.1186/s40537-019-0268-2>
- [16] Pulikottil, T., Schloegl, W., & Ansari, F. (2023). Immune system inspired smart maintenance framework: Tool-wear monitoring use case. *The International Journal of Advanced Manufacturing Technology*, 131, 5547-5562. <https://doi.org/10.1007/s00170-024-13472-4>
- [17] Qi, K., Chen, Y., Li, Z., & Zhang, H. (2025). Advancing hospital healthcare: Achieving IoT-based secure health monitoring through multilayer machine learning. *Journal of Big Data*, 12, 14. <https://doi.org/10.1186/s40537-024-01038-w>

- [18] Alzubaidi, L., Zhang, J., Humaidi, A. J., Al-Dujaili, A., Duan, Y., Al-Shamma, O., Santamaria, J., Fadhel, M. A., Al-Amidie, M., & Farhan, L. (2021). Review of deep learning: Concepts, CNN architectures, challenges, applications, future directions. *Journal of Big Data*, 8, 53. <https://doi.org/10.1186/s40537-021-00444-8>
- [19] Lara de Leon, M. A., Osornio-Rios, R. A., & Romero-Troncoso, R. J. (2024). Tool condition monitoring methods applicable in the metalworking process. *Archives of Computational Methods in Engineering*, 31, 1445-1466. <https://doi.org/10.1007/s11831-023-09979-w>
- [20] Wang, S., Balarezo, J. F., & Kandeepan, S. (2021). Machine learning in network anomaly detection: A survey. *IEEE Access*, 9, 152379-152396. <https://doi.org/10.1109/ACCESS.2021.3126834>
- [21] Lundberg, S. M., & Lee, S. I. (2017). A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems*, 30, 4765-4774.
- [22] Ompusunggu, A. P., Vandenplas, S., & Smet, P. D. (2021). Condition monitoring of critical industrial assets using high performing low-cost MEMS accelerometers. *Procedia CIRP*, 104, 1389-1394. <https://doi.org/10.1016/j.procir.2021.11.234>
- [23] Kunicki, M., Wotzka, D., & Daminabo, F. (2020). Data acquisition system for on-line temperature monitoring in power transformers. *Measurement*, 161, 107909. <https://doi.org/10.1016/j.measurement.2020.107909>
- [24] Attaran, M., & Attaran, S. (2023). Digital twin: Benefits, use cases, challenges, and opportunities. *Decision Analytics Journal*, 6, 100165. <https://doi.org/10.1016/j.dajour.2023.100165>
- [25] Giliyana, S., Johansen, K., & Sobaszek, L. (2025). Implementing and using smart maintenance technologies: Introducing challenges and enablers related to human, organizational and technological perspectives. *Procedia Computer Science*, 253, 932-941. <https://doi.org/10.1016/j.procs.2025.01.117>