International Journal of Emerging Research in Science, Engineering, and Management
Vol. 1, Issue 1, pp.01-09, July 2025.
www.ijersem.com

AI-based Predictive Maintenance in Mechanical Systems

[1]Sayyed Nagulmeera, [2]G Rajesh, [3]Bandi Rajasekhar, [4]Shaik Hedayath Basha

[1]Assistant Professor, Department of CSE, DVR & Dr. HS MIC College of Technology, Vijayawada, India. syednagulmeera99@gmail.com
[2]Assistant Professor, Department of CSE, NBKRIST, Vidyanagar, India. rajeshg@nbkrist.org
[3]HOD, Department of CSE, Sree Venkateswara College of Engineering, Nellore, India. rsekharbandi@gmail.com
[4]Associate Professor, Department of ECE, R.M.K. College of Engineering and Technology, Chennai, India. shaikhedayathbasha@rmkcet.ac.in

Abstract: Predictive maintenance has emerged as a critical solution for minimizing unplanned downtime, extending equipment lifespan, and enhancing operational efficiency in mechanical systems. Recent advancements in artificial intelligence (AI) have enabled the development of intelligent maintenance strategies that leverage machine learning (ML), deep learning (DL), and hybrid algorithms to anticipate equipment failures with high accuracy. While numerous AI-driven predictive maintenance solutions have been proposed, most are application-specific, resulting in fragmented methodologies with limited transferability across domains. This paper proposes a unified conceptual framework for AI-based predictive maintenance tailored to mechanical systems. Drawing insights from diverse sectors—including HVAC, gas turbines, photovoltaic systems, manufacturing, and tunneling infrastructure—the framework integrates essential layers: data acquisition, preprocessing, modeling, decision-making, and action. The proposed model emphasizes modularity, scalability, and adaptability, and supports integration with real-time data sources, including edge computing platforms. This paper aims to consolidate current advancements, address cross-domain limitations, and offer a reusable framework that can guide the implementation of predictive maintenance across a variety of mechanical environments.

Keywords: AI-based prediction, HVAC, Gas Turbines, Mechanical Systems, Tunneling.

1 Introduction

Mechanical systems are at the core of critical infrastructure in industries ranging from energy and transportation to manufacturing and construction. Ensuring their optimal performance and reliability requires proactive strategies to detect faults and schedule maintenance before failures occur. Traditional maintenance approaches—reactive and preventive—are often insufficient, leading to costly downtime and reduced equipment lifespan. Predictive maintenance (PdM), powered by artificial intelligence (AI), offers a more intelligent alternative by leveraging operational data to anticipate failures and recommend timely interventions.

With the rise of Industry 4.0, AI-based PdM has gained momentum through the integration of sensors, the Internet of Things (IoT), and advanced analytics. Machine learning (ML) models such as Support Vector Machines (SVM), Artificial Neural Networks (ANN), and Extreme Gradient Boosting (XGBoost) have demonstrated their utility in predicting energy usage and estimating maintenance timelines in HVAC systems [1]. Similarly, deep learning architectures like Long Short-Term Memory (LSTM) networks have been employed in turbine health monitoring and photovoltaic system anomaly detection [2], [3].
Despite the proven benefits of AI-driven PdM, current research remains siloed within domain-specific solutions. For instance, gas turbines [2], manufacturing systems [4], and tunnel boring machines [5] have all seen unique implementations, often with distinct algorithms, input features, and deployment strategies. This fragmentation limits the scalability and interoperability of PdM solutions across industries.

The absence of a unifying framework hampers knowledge transfer increases deployment costs, and creates barriers for organizations attempting to adopt AI-based maintenance at scale. As such, there is a compelling need for a generalizable, modular, and scalable conceptual framework that can guide the development of AI-based PdM systems for diverse mechanical systems. Such a framework should accommodate varying sensor data streams, support both centralized and edge computing, and integrate seamlessly with existing maintenance workflows. This paper addresses this gap by proposing a conceptual framework for AI-based predictive maintenance in mechanical systems. The framework is synthesized from an extensive literature survey covering eleven state-of-the-art studies across multiple industrial domains. It is designed to standardize the key components of PdM—data acquisition, preprocessing, model training, decision-making, and action planning—while remaining adaptable to industry-specific constraints.

The remainder of the paper is organized as follows: Section 2 presents a categorized literature review; Section 3 identifies key gaps and limitations in current systems; Section 4 details the proposed conceptual framework; Section 5 maps real-world case studies to the framework; Section 6 discusses its benefits and challenges; Section 7 outlines future research directions; and Section 8 concludes the paper.

2 LITERATURE REVIEW

This section surveys recent advances in AI-based predictive maintenance, categorized by application domains, AI techniques employed, and notable implementation strategies. The review spans key industrial sectors including HVAC, gas turbines, photovoltaic systems, manufacturing, tunnel infrastructure, and autonomous vehicles. These case studies provide critical insights into the capabilities and limitations of existing models, forming the foundation for the proposed unified framework.

2.1 Application Domains

2.1.1 HVAC and Energy Systems

Amin et al. [1] employed a data-driven approach to predict energy consumption and plan maintenance activities for Active Chilled Beam (ACB) systems in office buildings. Using a dataset of 2,500 samples, machine learning models such as Gaussian Process Regression (GPR) and XGBoost were used to accurately predict both electricity consumption and maintenance time. Their work demonstrates the dual benefit of AI models in optimizing energy usage and preventive maintenance scheduling.

2.1.2 Gas Turbines

In [2], Brahimi et al. developed an intelligent monitoring system for MS5002C gas turbines using Adaptive Neuro-Fuzzy Inference Systems (ANFIS) and Long Short-Term Memory (LSTM) networks. Trained on decades of historical data (1985–2021), their model effectively predicted component degradation and enabled reliability-centered maintenance planning.

2.1.3 Photovoltaic Systems

Syamsuddin et al. [3] proposed a predictive maintenance approach for solar power plants based on anomaly detection using SCADA data. A Long Short-Term Memory Autoencoder (LSTM-AE) model was trained to reconstruct normal system behavior, with deviations flagged as potential faults. The model achieved high accuracy in identifying anomalies and demonstrated the effectiveness of unsupervised learning in real-world energy systems.

2.1.4 Manufacturing Systems

Marti-Puig et al. [4] explored predictive maintenance in a wooden piece manufacturing setting. They predicted motor temperature in the factory's extraction system using Extreme Learning Machines, leveraging IoT sensor data and novel data-cleaning techniques. Their approach emphasized fast training and implementation readiness, suitable for dynamic industrial environments. Artiushenko et al. [6] advanced this further by implementing a resource-efficient Edge AI system for tool condition monitoring (TCM) during milling processes. The system executed multiple ML and DL models on low-power edge devices, demonstrating the feasibility of real-time PdM deployment in constrained environments.

2.1.5 Tunnel and Infrastructure Maintenance

Zou et al. [5] addressed predictive maintenance of Tunnel Boring Machines (TBMs) using an attention-based Graph Convolutional Network (att-GCN) to predict machine wear and performance. Their model achieved superior predictive accuracy by integrating geotechnical and operational data, with an online learning variant improving performance during real-time operations. Lu et al. [7] evaluated the mechanical performance of shield tunnel structures post-fire using a Backpropagation (BP) neural network. The model correlated various structural damage indices to predict post-fire tunnel capacity, aiding in emergency maintenance decision-making.

2.1.6 Autonomous Systems

Aeddula et al. [8] focused on autonomous haulage vehicles and integrated AI-based predictive maintenance within Product-Service System (PSS) development. Using variational autoencoders, the system forecasted both expected and unexpected failures, accounting for complex interdependencies in vehicle components and enabling proactive planning during early PSS phases.

2.2 AI Techniques in Predictive Maintenance

A diverse range of AI techniques has been adopted across domains:

- **Supervised Learning**: SVM, GPR, and XGBoost models were used for predicting maintenance time and energy consumption [1], [3].
- **Deep Learning**: LSTM [2], LSTM-AE [3], BP Neural Networks [7], and attention-based GCNs [5] enabled dynamic sequence modeling and multi-output predictions.
- **Hybrid Approaches**: ANFIS-LSTM [2] and att-GCN [5] illustrate the growing trend of combining multiple architectures for improved accuracy and robustness.
- **Unsupervised Learning**: Autoencoders [3], clustering techniques [7], and online learning [5] allowed anomaly detection in systems lacking labeled failure data.
- **Edge AI**: As shown in [6], running models on embedded devices near the data source offers reduced latency, enhanced privacy, and better scalability.

2.3 Observations and Trends

From the reviewed studies, several key observations emerge:

- **Domain-Specific Implementations**: Most studies focus on highly specific use cases, leading to non-transferable models.
- **Data-Driven Emphasis**: SCADA, IoT, and sensor data are central to almost all implementations, with preprocessing playing a critical role.
- **Evaluation Metrics**: Common metrics include RMSE, MAE, R^2, and MAPE, enabling comparative performance evaluation.
- **Platform Diversity**: Implementation tools range from MATLAB [1] to open-source edge devices [6], indicating flexibility but also variability in standardization.

These trends highlight the innovative but fragmented nature of current research, justifying the need for a unified conceptual framework.

3 GAPS IN EXISTING SYSTEMS

Despite notable progress in AI-based predictive maintenance across various mechanical systems, current implementations often remain fragmented, domain-specific, and limited in scalability. The following key limitations are evident from the literature:

3.1 Lack of Generalizability Across Domains

Most existing predictive maintenance solutions are developed for specific industrial contexts, with limited consideration for cross-domain applicability. For instance, models tailored for gas turbines [2] or HVAC systems [1] are not easily transferable to manufacturing lines [4], photovoltaic systems [3], or tunnel infrastructure [5]. Each implementation typically employs customized datasets, feature sets, and learning strategies, hindering model reuse or adaptation. This siloed approach increases development costs and delays broader adoption. A generalizable framework would facilitate reusability of core components (e.g., preprocessing, modeling layers), reducing the time and expertise required for deploying AI-based PdM systems in new domains.

3.2 Challenges in Data Quality and Availability

AI models heavily depend on the availability and quality of operational data. However, many industrial environments suffer from issues such as:

- **Sensor blockages and noise** (e.g., addressed through preprocessing in [4])
- **Lack of labeled failure data**, especially in rare-event systems like photovoltaic farms [3]
- **Heterogeneous data formats** from SCADA systems, IoT devices, and legacy equipment

In [3], unsupervised learning via LSTM-AE was used as a workaround for unlabeled data, but the absence of standard data cleaning and labeling procedures remains a bottleneck. Moreover, high dependency on domain-specific data further reduces model interoperability.

International Journal of Emerging Research in Science, Engineering, and Management
Vol. 1, Issue 1, pp.01-09, July 2025.
www.ijersem.com

3.3 Limited Adoption of Real-Time and Edge Deployment

Only a few studies, such as [6], explored Edge AI—the deployment of models on lightweight, low-power hardware near the equipment. Most predictive models still rely on centralized computation, which may introduce latency, increase communication overhead, and raise data privacy concerns. Real-time decision-making is essential in mechanical systems like motors or autonomous vehicles [8], yet many current models are trained offline and lack online learning or inference capabilities. The adoption of edge-computing infrastructure and real-time adaptive learning models remains underdeveloped.

3.4 Fragmented Model Integration with Maintenance Workflow

Many AI models focus on predicting failures but fall short in actionable integration with existing maintenance management systems (e.g., CMMS). For instance, while [1] and [3] predicted maintenance time and anomalies respectively, there is little discussion on how these predictions feed into actual maintenance planning, scheduling, or resource allocation.
Additionally, user interface design, interpretability of model outputs, and collaboration with human decision-makers are largely overlooked. Without integration into broader decision support systems, AI outputs may be underutilized.

3.5 Limited Model Explainability and Trust

Industrial adoption of AI models requires transparency and explainability, especially in safety-critical systems. However, deep learning models such as LSTM [2], [3] or att-GCN [5] often function as “black boxes.” Few studies attempt to explain predictions, quantify uncertainty, or involve domain experts in the modeling loop.
A lack of explainability undermines trust, making stakeholders hesitant to rely on automated decisions. Future frameworks must incorporate interpretable AI methods and integrate visual analytics or explanation modules to bridge this gap.

3.6 Inconsistent Evaluation and Benchmarking

There is no consistent benchmarking framework across predictive maintenance studies. Models are evaluated using different metrics—RMSE, MAE, MAPE, R^2—on varying datasets. For example, [5] used MAPE for tunnel systems, [1] used RMSE and R^2 for HVAC predictions, and [3] used MAE for solar farm anomalies. This lack of standardization hinders fair comparisons and makes it difficult to identify universally effective approaches. A conceptual framework must promote evaluation consistency and define baseline metrics for diverse application areas.

4 PROPOSED CONCEPTUAL FRAMEWORK

The increasing complexity of mechanical systems, along with the vast amounts of data generated by sensors and control systems, necessitates a systematic and adaptable framework for AI-based predictive maintenance (PdM). Based on insights from the literature reviewed, we propose a multi-layered conceptual framework that integrates data collection, preprocessing, modeling, decision-making, and deployment into a cohesive structure. This modular framework is designed to be adaptable across domains, ensuring that AI-driven maintenance solutions can be scaled and optimized to meet the specific requirements of various mechanical systems.

The first layer of the framework is data acquisition, which serves as the foundational input layer. It encompasses both sensor-based real-time monitoring systems and historical data logs. This layer supports a wide variety of data types such as vibration, temperature, pressure, humidity, acoustic signals, system logs, maintenance history, and operating conditions. A robust data acquisition system ensures consistent, high-resolution inputs that are critical for detecting subtle anomalies and long-term degradation patterns. Integration with Supervisory Control and Data Acquisition (SCADA) systems and Internet of Things (IoT) platforms enables continuous streaming and archiving of mechanical performance metrics.

Once data is collected, it is processed through the data preprocessing and feature engineering layer. This component includes noise filtering, missing data imputation, normalization, and synchronization across multiple sensors or systems. Equally important is the extraction of meaningful features, such as statistical indicators (mean, RMS, kurtosis), spectral characteristics (FFT, wavelet transforms), and domain-specific parameters. In some applications, advanced methods like Principal Component Analysis (PCA) or autoencoders are used to reduce dimensionality while preserving critical information. Clean, well-structured data enhances model performance and interpretability, especially when used in real-time inference.

The modeling layer is at the core of the framework. It supports both supervised and unsupervised learning techniques, chosen based on the availability of labeled failure data. Supervised models such as Support Vector Machines (SVM), Random Forests (RF), Gradient Boosting Machines (GBM), and Deep Neural Networks (DNN) are suitable for applications with well-annotated datasets, while unsupervised methods such as clustering algorithms or autoencoders are preferred when failure labels are unavailable. Time-series models like LSTM and attention-based mechanisms can capture temporal dependencies in degradation.

In some cases, hybrid models—such as ANFIS or ensemble methods—are employed to combine predictive accuracy with interpretability. Model selection is guided by evaluation metrics such as Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), R-squared, or classification accuracy. Following the modeling stage, the decision and action layer interprets model outputs to generate actionable maintenance insights. This may include predicting the Remaining Useful Life (RUL), flagging potential faults, or classifying equipment condition into health states. Thresholds can be predefined or dynamically adjusted using anomaly detection techniques or risk-based prioritization. Decision rules are often integrated into maintenance management systems (CMMS), alert dashboards, or visual analytics interfaces to ensure timely and data-driven interventions. This layer also supports ranking of maintenance priorities, spare parts planning, and scheduling of interventions to minimize downtime.

At the top of the architecture is the deployment and integration layer, which addresses the system-level implementation of the predictive maintenance pipeline. Models can be deployed in centralized cloud environments, on local servers, or on edge devices depending on latency, privacy, and scalability needs. This layer includes API-based model serving, integration with ERP/SCADA platforms, real-time dashboards, and human-machine interfaces (HMI). In edge computing scenarios, lightweight models are optimized for performance-constrained devices to support in-situ decision-making. The deployment architecture must also support model version control, feedback loops, and retraining pipelines to ensure long-term adaptability.

A cross-cutting theme of the framework is explainability and human-in-the-loop interaction. For AI-based decisions to be actionable in industrial environments, users must understand the rationale behind predictions. Visualization tools, SHAP values, feature importance rankings, and interpretable model structures (e.g., decision trees or fuzzy systems) enable human operators to trust and validate the system. Moreover, incorporating domain expertise in model training and feedback loops improves accuracy and adoption.

The proposed framework is designed to be flexible, scalable, and explainable, offering a unified reference architecture that can guide the development and implementation of AI-based predictive maintenance systems in a variety of mechanical settings. Its structure supports both academic research and industrial deployment, bridging the gap between algorithmic innovation and real-world maintenance challenges.

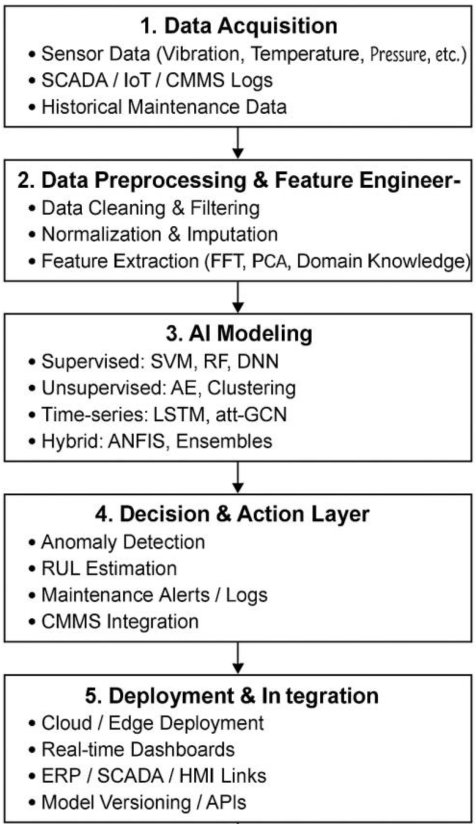

Fig. 1. Workflow of the proposed method

The proposed framework, as shown in Fig. 1, begins with data acquisition, the foundation for any predictive maintenance system. This includes real-time sensor data from equipment, historical logs from SCADA systems, and CMMS data such as maintenance history and operator reports. This layer ensures that both real-time and contextual data are available for intelligent decision-making. Next, data preprocessing and feature engineering is carried out to clean the data, handle missing values, and extract relevant features. This stage is crucial to improving the signal-to-noise ratio in datasets and enhancing model performance. Feature extraction methods such as Fourier transforms (FFT), wavelet analysis, and PCA are particularly useful in mechanical systems for detecting degradation trends.

The AI modeling layer supports various machine learning and deep learning algorithms. Supervised models like Random Forests and Deep Neural Networks are used when labeled data is available, while unsupervised models such as autoencoders or clustering techniques are deployed for anomaly detection in unlabeled datasets. LSTM and graph-based attention networks like att-GCN are leveraged for time-series modeling in complex systems with temporal dependencies. Hybrid approaches (e.g., ANFIS) combine the strengths of multiple algorithms for robust predictions.

The decision and action layer transforms model outputs into actionable insights. This could involve predicting the remaining useful life (RUL), detecting anomalies, or recommending specific maintenance actions. These decisions are logged, visualized, or sent to maintenance teams through integrated CMMS platforms. In the deployment and integration layer, the trained models are deployed to the cloud, local servers, or edge devices depending on system requirements. This layer also integrates model outputs into enterprise platforms such as ERP, SCADA, and HMI dashboards, enabling seamless flow of insights into operational workflows. Explainability and human-in-the-loop interaction ensures that predictions are interpretable and usable by field engineers and decision-makers. Visual tools (e.g., SHAP, LIME), feature importance scores, and intuitive dashboards help build trust and facilitate action. Operator feedback is incorporated to improve model learning and system performance over time.

5 Mapping the Framework to Real-World Cases

To validate the comprehensiveness and flexibility of the proposed conceptual framework, it is essential to examine how its components align with real-world implementations of AI-based predictive maintenance across different mechanical systems. The following discussion maps the key layers of the framework to selected studies from various domains, thereby demonstrating its applicability and generalizability.

In the context of HVAC systems, the study by Amin et al. [1] exemplifies the full realization of the data acquisition, modeling, and decision layers. Their work utilized a structured dataset comprising over 2,500 samples representing building and system characteristics, aligning directly with the framework's data acquisition layer. The preprocessing stage involved the careful selection of input features, including environmental and historical maintenance variables. Machine learning algorithms such as Gaussian Process Regression (GPR) and XGBoost were applied to predict both energy consumption and maintenance time, thereby populating the modeling layer of the framework. The outputs—accurate predictions of consumption and service intervals—enabled data-driven maintenance planning, feeding directly into the decision and action layer. Their use of MATLAB and EnergyPlus demonstrates platform-level flexibility, which fits well within the integration expectations of the proposed architecture.

Similarly, Brahimi et al. [2] offer a strong example from the energy sector, where ANFIS and LSTM models were deployed to monitor and anticipate faults in gas turbines. The modeling layer is enriched by the combination of neuro-fuzzy logic with deep learning, showcasing the hybrid approach supported by the framework. The models used a vast historical dataset, fulfilling the framework's requirement for robust input from the acquisition layer. Though less emphasis is placed on real-time feedback, the system effectively informs long-term maintenance schedules, supporting the decision-making component. Their reliability analysis further contributes to interpretability, aligning partially with the framework's vision of explainable and actionable AI.

In the manufacturing sector, Marti-Puig et al. [4] and Artiushenko et al. [6] reflect how predictive maintenance can be localized through edge deployment. Marti-Puig et al. focused on predicting the operational condition of motors used in wood residue extraction systems. Their preprocessing layer included an innovative algorithm to remove sensor-blockage artifacts, directly supporting the framework's emphasis on data quality. Meanwhile, Artiushenko et al. built an Edge AI prototype that implemented several machine learning and deep learning models on a low-power device to enable tool condition monitoring during milling. This strongly aligns with the proposed framework's provision for real-time deployment and edge integration, indicating the modular and adaptive nature of the architecture.

For photovoltaic systems, Syamsuddin et al. [3] employed LSTM autoencoders to detect anomalies in solar energy generation data. Their work represents a scenario where labeled failure data is unavailable, requiring the framework to support unsupervised learning strategies. The model reconstructs normal patterns and flags deviations, which is consistent with the modeling and decision layers of the framework. Although the action layer is not deeply integrated in their study, the detection of operational anomalies is itself a vital input for preventive strategies, thus validating the system's core functionality.

The case study presented by Zou et al. [5] on tunnel boring machines (TBMs) illustrates how complex mechanical systems operating in uncertain environments can benefit from advanced AI architectures such as attention-based graph convolutional networks (att-GCN). Their model processes both operational and geotechnical data to predict equipment degradation and adapts in real time through online learning. This implementation confirms the relevance of the framework's learning and decision layers in high-stakes, dynamic scenarios. The addition of an online variant further supports the framework's capability for continuous learning and performance improvement based on real-time data.

In autonomous systems, Aeddula et al. [8] demonstrated the use of variational autoencoders to predict maintenance needs for autonomous haulage vehicles. The study's strength lies in its focus on integrating predictive maintenance during the early phases of product-service system (PSS) development. This aligns closely with the framework's top-level integration layer, where AI insights feed into broader strategic planning tools. Their model's ability to identify hidden failure modes and enable early intervention supports the framework's goal of embedding intelligence across the lifecycle of mechanical assets.

Lu et al. [7] developed a backpropagation neural network to assess post-fire mechanical performance in shield tunnels. The study relied on clustering and numerical simulations to derive damage indices, which were then mapped to performance grades. This showcases the framework's capability to support multi-layered decision-making, where physical simulations and AI predictions work together to enable critical infrastructure maintenance.

Each of these real-world cases reflects elements of the proposed framework, whether in data acquisition, advanced modeling, real-time inference, or integration into operational workflows. While most studies focus on one or two components in depth, the proposed framework offers a unified structure that encapsulates all these elements, demonstrating its potential to serve as a reference architecture for scalable and intelligent predictive maintenance in diverse mechanical systems.

6 DISCUSSION

The proposed conceptual framework offers a unified, modular solution for implementing AI-based predictive maintenance (PdM) across diverse mechanical systems. By synthesizing insights from domain-specific applications, the framework promotes cross-sector applicability and encourages a structured approach to deploying intelligent maintenance strategies. This section discusses the broader significance of the framework, its potential advantages, implementation challenges, and opportunities for further enhancement.

One of the primary benefits of the proposed framework is its modularity, which allows different industrial sectors to adapt the architecture to their specific needs without reengineering the entire system. For example, a manufacturing plant and a photovoltaic power station may differ vastly in terms of operational dynamics and data formats, yet both can follow the same architectural blueprint, using interchangeable components such as data preprocessing pipelines, modeling blocks, or visualization interfaces. This modularity not only facilitates faster development and deployment but also supports the continuous evolution of each component based on technological advancements or domain requirements.

Another key strength of the framework is its support for real-time analytics and Edge AI deployment. Traditional predictive maintenance approaches often rely on offline models that analyze historical data, resulting in delayed insights and reactive decisions. The integration of edge computing capabilities, as seen in studies like [6], enables local inference with reduced latency, improved data privacy, and decreased network bandwidth consumption. This is especially valuable in remote or infrastructure-limited environments, where real-time fault detection and on-site response are critical [9]-[11].

The framework also emphasizes the importance of explainability and integration—two aspects frequently overlooked in existing systems. While highly accurate models such as deep neural networks and autoencoders offer powerful predictions, their opaque nature can limit trust and acceptance among field engineers and decision-makers. By including a dedicated layer for human interaction and explainable AI, the framework acknowledges the socio-technical reality of industrial environments, where AI must complement human expertise rather than replace it. Furthermore, seamless integration with enterprise systems such as Computerized Maintenance Management Systems (CMMS) or Enterprise Resource Planning (ERP) platforms ensures that model outputs directly influence operational decisions, thereby bridging the gap between insight generation and actionable maintenance [12].

Despite its strengths, the framework also presents several implementation challenges. One of the most prominent is the requirement for high-quality, continuous data streams, which are often lacking in traditional or legacy equipment setups. While modern industrial facilities increasingly adopt IoT-enabled sensors and SCADA systems, many mechanical systems still operate without sufficient instrumentation. Retrofitting such systems with appropriate data acquisition capabilities involves considerable investment and effort [13].

Another challenge is the model management and lifecycle maintenance. AI models require continuous monitoring, retraining, and validation to ensure accuracy over time. Changes in system behavior, equipment wear, or operational conditions may introduce concept drift, which can degrade model performance. Organizations must establish clear protocols for version control, periodic evaluation, and retraining strategies—preferably automated—to maintain the reliability of their predictive models.

Moreover, while the framework advocates for flexibility in algorithm selection and deployment platforms, this very flexibility can become a source of complexity. Organizations may face decision fatigue in choosing the right combination of algorithms, platforms, and tools, especially when expertise in AI is limited. To mitigate this, domain-specific implementation templates or model selection guidelines could be developed as supplementary resources to the framework.

Cybersecurity and data governance are critical concerns, particularly in environments that involve remote monitoring and cloud-based inference. As data becomes increasingly central to maintenance decision-making, ensuring its security, integrity, and compliance with regulatory standards is essential. The framework, therefore, must be implemented alongside robust data protection policies and secure communication protocols.

The proposed framework offers a promising foundation for generalizing AI-based predictive maintenance across mechanical systems. It brings structure to a currently fragmented field, promotes scalable and explainable deployment, and encourages integration with enterprise-level decision processes. At the same time, successful implementation demands careful consideration of data quality, model maintenance, user trust, and system security. Addressing these challenges proactively will be key to realizing the full potential of predictive maintenance in the era of Industry 4.0 and beyond.

7 Conclusions

This paper presented a conceptual framework for AI-based predictive maintenance (PdM) in mechanical systems, drawing on a comprehensive review of recent literature across diverse industrial domains. While numerous studies have demonstrated the potential of AI to enhance maintenance strategies in specific contexts—such as HVAC systems, gas turbines, manufacturing, photovoltaic farms, and tunnel boring machines—most existing solutions remain domain-specific, fragmented, and difficult to scale. Our proposed framework addresses these limitations by offering a modular, general-purpose architecture that integrates data acquisition, preprocessing, intelligent modeling, decision-making, and system-level integration.

The framework emphasizes flexibility in algorithm selection, supports real-time inference via edge computing, and incorporates explainable AI components to build trust among human operators. Its alignment with real-world applications validates its structure, while its generalizability allows organizations to adapt it to varied mechanical systems without extensive reengineering. Moreover, the framework highlights the importance of integrating predictive outputs with enterprise systems to ensure actionable maintenance planning.

Despite its potential, the deployment of such a framework poses challenges related to data quality, model maintenance, user trust, and cybersecurity. Overcoming these challenges requires a multidisciplinary effort involving AI experts, domain engineers, IT professionals, and organizational stakeholders. Future work may focus on implementing domain-specific instantiations of the framework, developing automated model lifecycle management tools, and creating standardized benchmarks for evaluating predictive maintenance systems.

By bridging the gap between advanced AI models and practical industrial needs, this conceptual framework provides a foundation for developing intelligent, scalable, and integrated predictive maintenance systems. As industries continue to embrace the principles of Industry 4.0 and beyond, such frameworks will play a critical role in enhancing asset reliability, reducing downtime, and optimizing operational efficiency.

Funding Information
This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Ethics Statement
This study did not involve human or animal subjects and, therefore, did not require ethical approval.

Statement Of Conflict Of Interests
The authors declare no conflicts of interest related to this study.

REFERENCES

[1] N. H. Amin, A. Etemad, and A. Abdalisousan, "Data-driven performance analysis of an active chilled beam air conditioning system: A machine learning approach for energy efficiency and predictive maintenance," *Results in Engineering*, vol. 23, p. 102747, Aug. 2024, doi: 10.1016/j.rineng.2024.102747.

[2] L. Brahimi, N. Hadroug, A. Iratni, A. Hafaifa, and I. Colak, "Advancing predictive maintenance for gas turbines: An intelligent monitoring approach with ANFIS, LSTM, and reliability analysis," *Computers & Industrial Engineering*, vol. 191, p. 110094, Mar. 2024, doi: 10.1016/j.cie.2024.110094.

[3] A. Syamsuddin, A. C. Adhi, A. Kusumawardhani, T. Prahasto, and A. Widodo, "Predictive maintenance based on anomaly detection in photovoltaic system using SCADA data and machine learning," *Results in Engineering*, p. 103589, Dec. 2024, doi: 10.1016/j.rineng.2024.103589.

[4] P. Marti-Puig, I. A. Touhami, R. C. Perarnau, and M. Serra-Serra, "Industrial AI in condition-based maintenance: A case study in wooden piece manufacturing," *Computers & Industrial Engineering*, vol. 188, p. 109907, Jan. 2024, doi: 10.1016/j.cie.2024.109907.

[5] X. Zou et al., "Advancing tunnel equipment maintenance through data-driven predictive strategies in underground infrastructure," *Computers and Geotechnics*, vol. 173, p. 106532, Jun. 2024, doi: 10.1016/j.compgeo.2024.106532.

[6] V. Artiushenko, S. Lang, C. Lerez, T. Reggelin, and M. Hackert-Oschätzchen, "Resource-efficient Edge AI solution for predictive maintenance," *Procedia Computer Science*, vol. 232, pp. 348–357, Jan. 2024, doi: 10.1016/j.procs.2024.01.034.

[7] Y. Lu, J. Jiang, B. Wang, W. Chen, and J. Ye, "AI-based evaluation method of mechanical performance of shield tunnel structures after fire," *Tunnelling and Underground Space Technology*, vol. 150, p. 105858, May 2024, doi: 10.1016/j.tust.2024.105858.

[8] O. Aeddula, M. Frank, R. Ruvald, C. J. Askling, J. Wall, and T. Larsson, "AI-Driven predictive maintenance for autonomous vehicles for Product-Service system development," *Procedia CIRP*, vol. 128, pp. 84–89, Jan. 2024, doi: 10.1016/j.procir.2024.06.008.

[9] C. Gnekpe, D. Tchuente, S. Nyawa, and P. K. Dey, "Energy performance of building refurbishments: Predictive and prescriptive AI-based machine learning approaches," *Journal of Business Research*, vol. 183, p. 114821, Jul. 2024, doi: 10.1016/j.jbusres.2024.114821.

[10] A. Stohr, P. Ollig, R. Keller, and A. Rieger, "Generative mechanisms of AI implementation: A critical realist perspective on predictive maintenance," *Information and Organization*, vol. 34, no. 2, p. 100503, Feb. 2024, doi: 10.1016/j.infoandorg.2024.100503.

[11] P. D. Paraschos and D. E. Koulouriotis, "Learning-based production, maintenance, and quality optimization in smart manufacturing systems: A literature review and trends," *Computers & Industrial Engineering*, p. 110656, Oct. 2024, doi: 10.1016/j.cie.2024.110656.

[12] Gusnul Mahesa, Ahmad Tajri, Eflan Ananda Pujito, Muhammad Rizal, Dimas Arya Bagaskara, Mohammad Givi Efgivia, "Integration of Artificial Intelligence in Network Technology: A Literature Review," *International Journal of Emerging Research in Engineering, Science, and Management*, vol. 4, no. 2, pp. 22-31, 2025. doi: 10.58482/ijeresm.v4i2.4.

[13] Allu Venkata Satyanarayana, Yerubandi Venkata Ramana Murthy, Chandra Mouli V.S.A, "Evaluation of Mechanical Properties of Al6061 Composite by Stir Casting Method," *International Journal of Emerging Research in Engineering, Science, and Management*, vol. 3, no. 3, pp. 51-58, 2024. doi: 10.58482/ijeresm.v3i3.9.