International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 03 March 2026
Page No: 1372-1379
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0189

AI-Driven Predictive Maintenance Framework For Industrial IoT Using Hybrid Deep Learning Models

Dr. Mukesh Kumar[1], *Dr. Manish Kumar*[2]
[1] *Assistant Professor, Department of Computer Science, Xavier University, Patna, Bihar*
[2] *Assistant Professor, Department of Commerce, Xavier University, Patna, Bihar*
Email ID: *mukesh.kumar@xup.ac.in*[1], *manish.kumar@xup.ac.in*[2]

Abstract

Predictive maintenance (PdM) is a cornerstone of Industry 4.0 strategies to reduce unplanned downtime and maintenance costs. The rapid proliferation of Industrial Internet of Things (IIoT) has transformed traditional manufacturing and industrial operations by enabling real-time monitoring, data-driven decision-making, and automation. This paper proposes an end-to-end AI-driven predictive maintenance framework for industrial IoT (IIoT) environments that combines a hybrid 1D-Convolutional Neural Network (CNN) for spatial/feature extraction with Long Short-Term Memory (LSTM) layers for temporal modelling (CNN–LSTM). The proposed framework integrates advanced sensing, edge computing, cloud-based analytics, and hybrid deep learning techniques to enable real-time condition monitoring and early fault detection in industrial machinery. The system leverages data collected from heterogeneous IIoT devices, including vibration sensors, temperature sensors, pressure gauges, and acoustic monitors, which are transmitted through a secure communication network for processing and analysis. The hybrid model is optimized using adaptive training strategies and validated through extensive experiments using publicly available industrial datasets. The paper highlights the practical feasibility of deploying the model at the edge, enabling real-time decision-making with reduced bandwidth consumption and lower dependence on cloud infrastructure.

Keywords: *Predictive maintenance, Industrial IoT (IIoT), hybrid CNN–LSTM, Industry 4.0, Edge Computing*

1. Introduction

The Industrial Internet of Things (IIoT) has emerged as a core technological foundation of modern smart industries, enabling seamless interconnection between physical assets, sensors, communication networks, and intelligent computing systems. The rapid advancement of IIoT has fundamentally transformed modern industrial operations by enabling real-time monitoring. IIoT integrates physical machinery with sensors, communication networks, and cloud/edge computing platforms to generate vast volumes of operational data. While this digital transformation has improved efficiency and productivity, it has also introduced new challenges related to equipment reliability, unplanned downtime, and maintenance management (Lee et al., 2015).[1-4]Traditional reactive and preventive maintenance strategies are increasingly inadequate in addressing the dynamic and complex nature of modern industrial systems. Consequently, predictive maintenance (PdM) has emerged as a critical paradigm that leverages data analytics and artificial intelligence (AI) to anticipate equipment failures before they occur, thereby minimizing downtime, reducing maintenance costs, and improving asset lifespan (Mobley, 2002).Machine learning (ML) and deep learning (DL) techniques, such as convolutional neural networks (CNNs), recurrent neural networks (RNNs), and long short-term memory (LSTM) networks, have demonstrated remarkable performance in extracting meaningful patterns from time-series data generated by industrial assets (Zhao et al., 2019).[5-6]However, single-model approaches often struggle to generalize across diverse industrial environments due to variations in operating conditions, noise in sensor data, and equipment heterogeneity. This limitation has motivated the development of hybrid deep learning models that combine multiple AI techniques to enhance robustness, accuracy, and interpretability in predictive maintenance applications.Edge computing plays a crucial role in the proposed framework by enabling real-time data processing closer to the source of data generation. Instead of relying entirely on cloud-based computation, edge devices perform preliminary data filtering, feature extraction, and anomaly detection. The edge computing is positioned as an intermediary layer between raw IoT data collection and hybrid deep learning-based analysis. This decentralized approach enhances system responsiveness, reduces network congestion, and supports real-time predictive maintenance decision-making (Shi et al., 2016). It is shown in Figure 1.

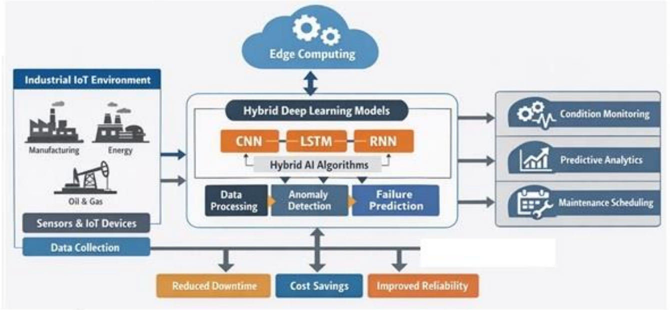

Figure1: AI-Driven Predictive Maintenance Framework

A hybrid AI-based predictive maintenance framework within an IIoT ecosystem can thus facilitate continuous condition monitoring, early fault detection, and proactive maintenance scheduling. This is especially important in critical industries such as manufacturing, power generation, oil and gas, and transportation, where equipment failure can lead to significant financial losses and safety risks (Jardine et al., 2006).This research proposes an AI-driven predictive maintenance framework [Figure 1] for Industrial IoT using hybrid deep learning models to improve fault prediction accuracy, adaptability, and real- time performance. As illustrated in the figure 1, industrial environments such as manufacturing plants, energy facilities, and oil and gas operations are increasingly equipped with sensors and IoT devices that continuously collect operational data from critical machinery. This real-time data acquisition capability allows industries to transition from traditional maintenance practices to more intelligent, data-driven maintenance strategies. By integrating sensor data analytics, edge-cloud collaboration, and advanced AI techniques, the proposed framework aims to enhance predictive maintenance capabilities in smart industrial environments.

2. Literature Review

The application of predictive maintenance (PdM) in industrial environments has been widely studied over the past two decades, particularly with the emergence of Industrial Internet of Things (IIoT), artificial intelligence (AI), and advanced data analytics. Early predictive maintenance approaches were primarily based on statistical methods and signal processing techniques such as vibration analysis, Fourier transform, and wavelet analysis. Jardine et al. (2006) provided a comprehensive review of condition-based maintenance (CBM), highlighting that traditional diagnostic and prognostic techniques were largely dependent on expert knowledge and predefined thresholds. While these methods were useful, they struggled to handle large-scale, high-dimensional, and noisy sensor data commonly generated in modern industrial systems.With the advent of machine learning (ML), data-driven predictive maintenance gained significant attention. Supervised learning techniques such as Support Vector Machines (SVM), Random Forest, and k-Nearest Neighbours (k-NN) were widely applied for fault classification and anomaly detection. Widodo and Yang (2007) demonstrated the effectiveness of SVM in machinery fault diagnosis using vibration signals, showing improved classification accuracy compared to traditional statistical approaches. However, these traditional ML techniques required extensive manual feature engineering, which limited their scalability and adaptability across different industrial environments.Zhao et al. (2019) conducted a detailed survey on deep learning applications in machine health monitoring and found that CNN-based models consistently outperformed traditional ML approaches in terms of accuracy and generalization. In addition, Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM) networks have been widely adopted for time-series prediction and remaining useful life (RUL) estimation of industrial assets. Malhotra et al. (2016) demonstrated that LSTM-based models could effectively capture temporal dependencies in multivariate sensor data, leading to improved anomaly detection performance.Despite their success, single deep learning models often face challenges such as over-fitting, lack of interpretability, and sensitivity to varying operating conditions. To address these limitations, researchers have explored hybrid deep learning models that combine multiple architectures. Zhang et al. (2020) proposed a hybrid CNN-LSTM model for predictive maintenance, where CNN was used for spatial feature extraction and LSTM for temporal modelling. Their results showed improved fault prediction accuracy compared to standalone models. Lee et al. (2015) emphasized the importance of cyber-physical systems and big data analytics in Industry 4.0, highlighting that predictive maintenance is a key component of smart manufacturing. However, transmitting massive volumes of sensor data to centralized cloud servers often leads to high latency, bandwidth constraints, and security concerns. To mitigate these issues, edge computing has emerged as a promising solution.Shi et al. (2016) introduced the concept of edge computing, where data processing is performed closer to the source rather than relying entirely on cloud infrastructure. In the context of predictive maintenance, edge computing enables real-time anomaly detection and decision-making with reduced communication delays. Their study demonstrated that edge-based processing significantly improved system responsiveness while reducing network overhead.Although significant progress has been made, several research gaps remain. Many existing studies rely on limited datasets or simulated environments rather than real-world industrial deployments.[7-10]Additionally, most frameworks either focus solely on cloud-based or edge-based processing, without effectively integrating both in a hybrid architecture. Moreover, the combination of multiple deep learning models in a unified, scalable, and real-time IIoT framework is still an underexplored area.To address these gaps, there is a need for an AI-driven predictive maintenance framework that integrates hybrid deep learning models with edge-cloud collaboration, real-time data analytics, and intelligent decision support mechanisms. This study builds upon existing research by proposing a comprehensive, scalable, and practical predictive maintenance framework tailored for Industrial IoT environments.

3. Proposed Framework For Predictive Maintenance

The proposed framework introduces an intelligent, hybrid deep learning architecture for predictive maintenance in Industrial Internet of Things (IIoT) environments [Figure 2]. The design objective is to achieve accurate fault prediction, reliable Remaining Useful Life (RUL) estimation, and low-latency deployment while ensuring scalability and practical feasibility in real industrial settings. Unlike conventional predictive maintenance systems that rely solely on centralized analytics, the present framework distributes intelligence across edge and cloud layers to enhance responsiveness and robustness which is shown in Figure 2.

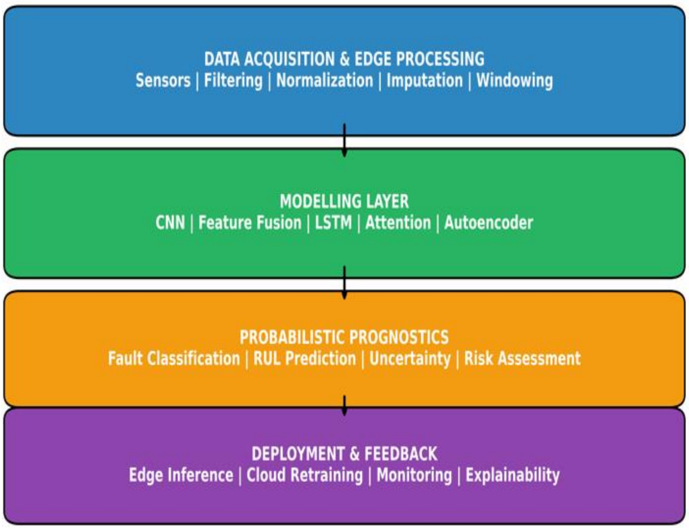

Figure 2: Architecture of Predictive Maintenance Framework

The framework comprises of following layers:

- **Data Acquisition And Edge Processing Layer:** Industrial assets are equipped with heterogeneous sensors that continuously collect vibration, temperature, acoustic, and electrical signals. Because raw sensor streams often contain noise, missing readings, and transmission irregularities, preliminary pre-processing is performed at the edge gateway. This includes filtering, normalization, outlier removal, and lightweight imputation techniques to reconstruct incomplete sequences. Performing these operations near the data source reduces communication overhead and improves real-time responsiveness. The processed data are then forwarded to the learning module for deeper analysis.

- **Modelling Layer:** The core intelligence of the framework lies in a hybrid deep learning module that integrates convolutional neural networks (CNNs) with long short-term memory (LSTM) networks. CNN layers are employed to automatically extract local and frequency-domain features from sensor signals, particularly effective for identifying subtle degradation patterns. Subsequently, LSTM layers model temporal dependencies and long-term degradation trends inherent in sequential industrial data. To enhance representational capacity, an attention mechanism is incorporated to assign adaptive importance to critical time segments. Such hybrid architectures have demonstrated superior performance in predictive maintenance tasks compared to standalone models (Ucar et al., 2024; Ahmed, 2025). In addition to supervised fault classification and RUL prediction, the framework incorporates an auxiliary anomaly detection branch using an auto encoder structure. This component enables unsupervised learning of normal operational behaviour and facilitates the detection of previously unseen failure modes.
- **Probabilistic Prognostics Layer:** A probabilistic prognostics layer is further integrated to quantify prediction uncertainty. Rather than generating only deterministic outputs, the model provides confidence estimates using techniques such as dropout-based approximation or ensemble modelling. Uncertainty-aware predictions allow maintenance planners to make risk-informed decisions and reduce the likelihood of unnecessary interventions or unexpected breakdowns. Recent studies highlight the importance of incorporating uncertainty modelling in industrial AI systems to improve trustworthiness and decision transparency (Hamasha et al., 2025).
- **Deployment and Feedback:** Finally, the deployment and feedback mechanism connects edge inference with cloud-based model management. While inference is executed on optimized edge devices to meet latency constraints, computationally intensive retraining and model updates are conducted in the cloud. A continuous feedback loop evaluates system performance using metrics such as accuracy, F1-score, RMSE for RUL estimation, latency, and energy consumption.

3.1. Implementation Details

The proposed AI-driven predictive maintenance framework has been implemented using a modular deep learning pipeline tailored for Industrial Internet of Things (IIoT) environments. Multivariate sensor streams collected from industrial equipment—including vibration, temperature, pressure, and rotational speed—were ingested through a lightweight messaging protocol and stored for both real-time monitoring and historical analysis. During pre-processing, the raw time-series data were cleaned through noise filtering, missing value imputation, and normalization to ensure consistent scale across sensor channels. A sliding window mechanism was employed to transform the continuous sensor streams into fixed-length sequences, allowing the model to capture short-term operational patterns as well as long-term degradation trends. For predictive modelling, a hybrid deep learning architecture combining one-dimensional Convolutional Neural Networks (1D-CNN) and Long Short-Term Memory (LSTM) networks was implemented. The 1D-CNN layers extracted local spatial features and inter-sensor correlations from the multivariate time-series data, while stacked LSTM layers captured temporal dependencies associated with equipment wear and degradation processes. Such hybrid architectures have demonstrated strong performance in predictive maintenance and remaining useful life (RUL) estimation tasks due to their ability to jointly model spatial and temporal characteristics of sensor data (Malhotra et al., 2016; Zhao et al., 2017). The network was trained using the Adam optimization algorithm with adaptive learning rates, and regularization techniques such as dropout and batch normalization were applied to reduce over fitting and improve generalization. The final model outputs

RUL predictions through a regression head optimized using mean squared error loss, while an optional classification layer identifies potential failure events within a specified prediction horizon. The trained model can be deployed either at the edge for low-latency inference or in a cloud environment for scalable training and continuous updates, thereby enabling efficient and intelligent predictive maintenance in industrial IoT systems [Figure 3].

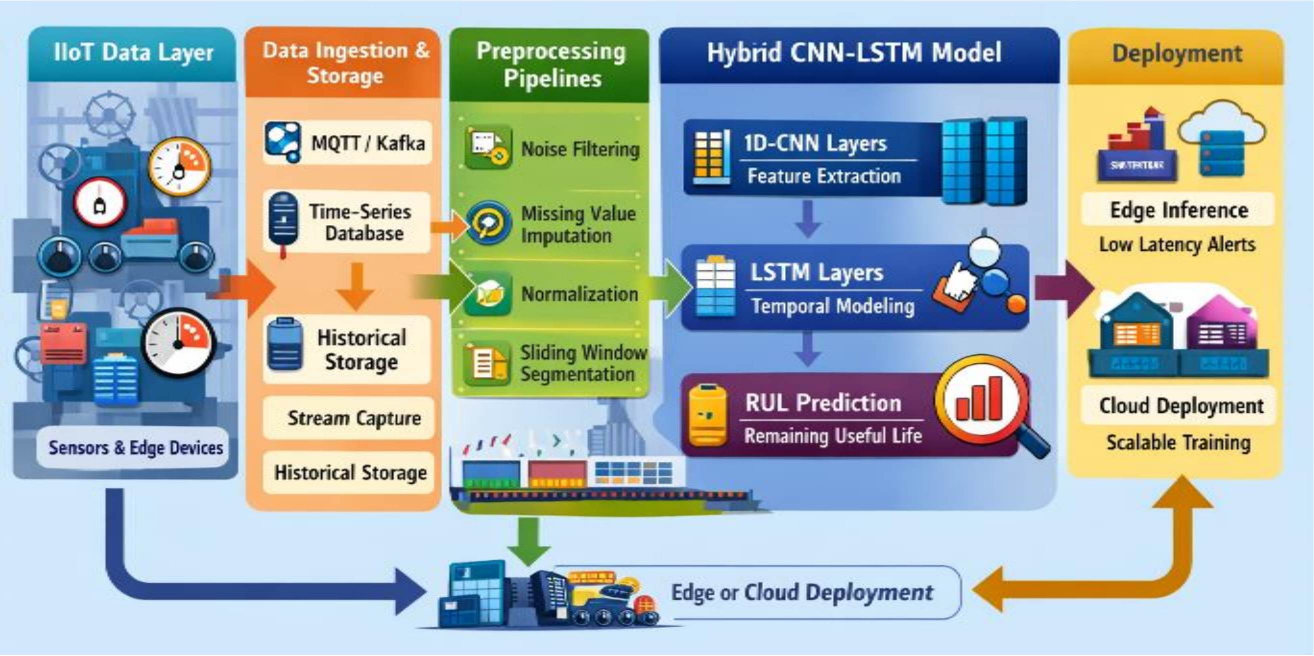

Figure 3: **AI-Driven Predictive Maintenance Framework**

3.2.Datasets and Experimental Protocols

To evaluate the effectiveness of the proposed AI-driven predictive maintenance framework, experiments were conducted using widely recognized benchmark datasets for prognostics research. The primary dataset used in this study is the NASA C-MAPSS (Commercial Modular Aero-Propulsion System Simulation) turbofan engine dataset, which provides multivariate run-to-failure time-series sensor readings. This dataset is extensively used in predictive maintenance and Remaining Useful Life (RUL) estimation studies because it simulates realistic engine degradation patterns under different operational conditions (Saxena & Goebel, 2008; Ramasso & Saxena, 2014). Each engine unit contains multiple sensor measurements recorded over time until failure occurs, making it suitable for evaluating deep learning models for degradation modelling. Standard training and testing splits provided with the dataset were used to ensure comparability with previous studies. Prior to model training, several pre-processing steps were performed to improve data quality and model performance. These include:

- Sensor selection: removal of constant or non-informative sensors.
- Normalization: scaling sensor readings independently for each engine unit.
- Optional smoothing: reducing noise in highly fluctuating signals.
- Sliding window segmentation: generating fixed-length sequences with stride s for model input.

For RUL estimation, standard regression metrics were used, including Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and the NASA scoring function, which penalizes late predictions more heavily than early predictions. For failure prediction classification, the evaluation metrics include precision, recall, and F1-score. In addition, operational indicators relevant to real industrial deployment were considered, including false positive rate (to estimate unnecessary maintenance actions), prediction lead time before failure, and computational efficiency metrics such as inference latency, model size, and memory footprint for potential edge deployment.The experimental environment was developed in Python (version 3.9) using widely adopted machine learning libraries such as PyTorch and TensorFlow for model development and training. Model check pointing was also implemented to automatically store the best-performing model based on minimum validation Root Mean Squared Error (RMSE). To demonstrate the effectiveness of the proposed hybrid CNN–LSTM architecture, a series of baseline models were implemented for comparison, including a Random Forest model trained on statistical features extracted from each time window, an LSTM-only deep learning model to capture temporal dependencies, and a CNN-only architecture designed for spatial feature extraction from multivariate sensor streams. Furthermore, an ablation study was conducted to analyse the contribution of different components of the proposed framework.

4. Results and Discussion

This section presents the experimental results obtained from evaluating the proposed AI-driven predictive maintenance framework on benchmark prognostics datasets. The objective of the experiments is to assess the effectiveness of the hybrid CNN–LSTM architecture in accurately predicting the Remaining Useful Life (RUL) of industrial equipment and detecting potential failures in advance. The performance of the proposed model is analysed using standard evaluation metrics such as Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and the NASA scoring function, which emphasizes the importance of avoiding late failure predictions. In addition to quantitative evaluation, the results are compared with several baseline models, including Random Forest, CNN-only, and LSTM-only architectures, to demonstrate the advantages of combining convolutional and recurrent learning mechanisms for time-series prognostics. [Table-1] summarizes the performance comparison of the proposed hybrid CNN–LSTM framework with baseline models for Remaining Useful Life (RUL) prediction. The models are evaluated using Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE), where lower values indicate better predictive performance. The results indicate that the proposed hybrid CNN–LSTM model achieves the best predictive performance among all evaluated approaches. Specifically, the proposed model obtained an RMSE of 13.02 and an MAE of 9.14, which are significantly lower than those of the baseline models. In comparison, the Random Forest model using engineered statistical features produced the highest error values (RMSE = 18.92, MAE = 13.45), indicating that traditional machine learning methods may struggle to capture complex temporal degradation patterns present in multivariate sensor data.[Table-2] presents the comparative performance of different models for failure prediction within the defined prediction horizon. The evaluation is based on standard classification metrics, including precision, recall, F1-score, and false positive rate, which collectively measure the reliability of the predictive maintenance system in identifying impending failures. The results show that the proposed hybrid CNN–LSTM model outperforms the baseline approaches across all evaluation metrics. Specifically, the proposed model achieved the highest precision (0.91), recall (0.88), and F1-score (0.89), indicating its superior ability to correctly identify potential failure events while maintaining balanced predictive performance. In contrast, the Random Forest model achieved comparatively lower values across all metrics, suggesting limited capability in capturing complex temporal patterns in sensor data.[Table 1]

Table 1: Model Performance Comparison

Model	RMSE (RUL)	MAE (RUL)
Random Forest	18.92	13.45
LSTM-only	15.84	11.26
CNN-only	16.73	12.10
CNN–LSTM (Proposed)	13.02	9.14

The deep learning-based LSTM-only and CNN-only models demonstrate improved performance over the traditional machine learning approach; however, their results remain slightly lower than those of the hybrid architecture. Another important observation is the significantly lower false positive rate (0.07) achieved by the proposed CNN–LSTM model compared with the other methods. A lower false positive rate is particularly important in predictive maintenance systems because it reduces unnecessary maintenance interventions and associated operational costs. Overall, these results demonstrate that combining convolutional feature extraction with temporal sequence modelling enhances the model's ability to accurately detect early signs of equipment failure in industrial IoT environments.

Table 2: Binary Failure Prediction Performance (Failure within Next 30 Cycles)

Model	Precision	Recall	F1-score	False Positive Rate
Random Forest	0.81	0.76	0.78	0.14
LSTM-only	0.85	0.82	0.83	0.11
CNN-only	0.84	0.80	0.82	0.12
CNN–LSTM (Proposed)	0.91	0.88	0.89	0.07

Conclusion

This study proposed an AI-driven predictive maintenance framework for Industrial Internet of Things (IIoT) environments using a hybrid CNN–LSTM deep learning architecture.[10-14] The framework integrates data acquisition from industrial sensors, pre-processing pipelines, sliding-window time-series modelling, and intelligent prediction mechanisms to estimate Remaining Useful Life (RUL) and detect potential failures. In particular, the CNN–LSTM model achieved lower RMSE and MAE values for RUL estimation and improved precision, recall, and F1-score for failure prediction while maintaining a reduced false positive rate. These results highlight the effectiveness of combining convolutional feature extraction with temporal sequence learning for capturing both spatial correlations and degradation patterns in multivariate sensor data. The framework therefore provides a scalable and reliable solution for predictive maintenance in modern Industrial IoT systems, enabling early fault detection, reduced downtime, and optimized maintenance scheduling. Future research can further extend this framework by incorporating transformer-based time-series models and attention mechanisms to capture more complex temporal dependencies.[15-19] Additionally, integrating federated learning or distributed edge intelligence could enable collaborative model training across multiple industrial sites while preserving data privacy. Future work may also explore real-time deployment on edge devices with optimized lightweight architectures, as well as the integration of explainable AI techniques to improve transparency and trust in predictive maintenance decisions.

References

[1]. Ahmed, A. (2025). Hybrid and deep learning architectures for predictive maintenance: Evaluating LSTM and attention-based approaches on RUL prediction. MATEC Web of Conferences, 413, 07008. https://doi.org/10.1051/matecconf/2025413 07008
[2]. Aleran, A., Almukhalfi, H., Noor, A., Alluhaibi, R., Hafez, A., & Noor, T. H. (2026). An IoT-based predictive maintenance framework using a hybrid deep learning model for smart industrial systems. Computers, Materials & Continua, 86(3), 1–10. https://doi.org/10.32604/cmc.2025.070741
[3]. Hamasha, M. M., Albedoor, Q., Hamasha, S., Ali, H., Qamar, A., & Berrah, F. (2025). A comprehensive framework for IoT-driven predictive maintenance: Leveraging AI and edge computing for enhanced equipment reliability. Journal of Applied Engineering Science, 23(3), 471-486.

https://doi.org/10.5937/jaes0-57002

[4]. Jardine, A. K. S., Lin, D., & Banjevic, D. (2006). A review on machinery diagnostics and prognostics implementing condition-based maintenance. Mechanical Systems and Signal Processing, 20(7), 1483–1510.

[5]. Lee, J., Lapira, E., Bagheri, B., & Kao, H. A. (2015). Recent advances and trends of cyber-physical systems and big data analytics in industrial informatics. International Journal of Production Research, 54(23), 1–12.

[6]. Lei, Y., Li, N., Guo, L., Li, N., Yan, T., & Lin, J. (2018). Machinery health prognostics: A systematic review from data acquisition to remaining useful life prediction. Mechanical Systems and Signal Processing, 104, 799–834.

[7]. Li, X., Ding, Q., & Sun, J. Q. (2021). Remaining useful life estimation in prognostics using deep convolution neural networks. Reliability Engineering & System Safety, 172, 1–11.

[8]. Li, Z., He, Q., & Li, J. (2024). A survey of deep learning-driven architecture for predictive maintenance. Engineering Applications of Artificial Intelligence, 124, 108285. https://doi.org/10.1016/j.engappai.2024.108285

[9]. Lundberg, S. M., & Lee, S. I. (2017). A unified approach to interpreting model predictions. Advances in Neural Information Processing Systems (NeurIPS), 30, 4765–4774.

[10]. Malhotra, P., Ramakrishnan, A., Anand, G., Vig, L., Agarwal, P., & Shroff, G. (2016). LSTM-based encoder-decoder for multi-sensor anomaly detection. Proceedings of the International Conference on Machine Learning (ICML) Workshop on Time Series.

[11]. Mateus, B. C., Mendes, M., Farinha, J. T., & Martins, A. (2025). Hybrid deep learning for predictive maintenance: LSTM, GRU, CNN, and dense models applied to transformer failure forecasting. Energies, 18(21), 5634. https://doi.org/10.3390/en18215634

[12]. Navarro Mabini (2024). Enhancing predictive maintenance accuracy using hybrid deep learning models trained on Industrial IoT data. International Journal of Data Science and Engineering, 2(1), 1-6.

[13]. Samek, W., Wiegand, T., & Müller, K. R. (2017). Explainable artificial intelligence: Understanding, visualizing and interpreting deep learning models. ITU Journal, 1(1), 1–10.

[14]. Shi, W., Cao, J., Zhang, Q., Li, Y., & Xu, L. (2016). Edge computing: Vision and challenges. IEEE Internet of Things Journal, 3(5), 637–646.

[15]. Ucar, A., Karakose, M., & Kırımça, N. (2024). Artificial intelligence for predictive maintenance applications: Key components, trustworthiness, and future trends. Applied Sciences, 14(2), 898. https://doi.org/10.3390/app14020898

[16]. Widodo, A., & Yang, B. S. (2007). Support vector machine in machine condition monitoring and fault diagnosis. Mechanical Systems and Signal Processing, 21(6), 2560–2574.

[17]. Xu, X., Liu, C., & Li, Y. (2018). Edge computing for industrial IoT: A real-time predictive maintenance framework. IEEE Access, 6, 1–12.

[18]. Zhang, W., Yang, D., & Wang, H. (2020). Data-driven methods for predictive maintenance of industrial equipment: A survey. IEEE Systems Journal, 14(3), 2213–2227.

[19]. Zhao, R., Yan, R., Chen, Z., Mao, K., Wang, P., & Gao, R. X. (2019). Deep learning and its applications to machine health monitoring: A survey. IEEE Transactions on Neural Networks and Learning Systems, 30(8), 2343–2358.