



International Journal of Innovative Research in Science Engineering and Technology (IJIRSET)

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



Impact Factor: 9.935

Volume 15, Issue 8, August 2026

RULInsight: A Comprehensive Review of Explainable Artificial Intelligence Techniques for Remaining Useful Life Prediction in Industrial Predictive Maintenance Systems

Harshal Garade^{1*}, Prem Golhar², Arpit Deshmukh³, Mugdha Gedam⁴, Ayush Yele⁵, Swati Dronkar⁶

Department of Computer Science and Engineering, Priyadarshini College of Engineering, Nagpur,

Maharashtra, India¹²³⁴⁵

Asst. Professor, Department of Computer Science and Engineering, Priyadarshini College of Engineering, Nagpur,

Maharashtra, India⁶

ABSTRACT: The acceleration of digital transformation for industrial systems is rapidly promoting smart maintenance to prevent the machine failure. Remaining Useful Life (RUL) is one of the cores of predictive maintenance technologies and its early estimation is effective to cut down not only system downtime but also to avoid the cost of maintenance and to minimize the productivity loss due to system malfunction. AI-driven predictive techniques for RUL estimation guarantees a more precise RUL prediction, however, the lack of model interpretability hinders their deployment in decision making and operational support since their decision-making process is not transparent for human experts.

In this survey, we will review the latest progresses of RUL estimation by leveraging ML, DL, explainable artificial intelligence (XAI), Retrieval Augmented Generation (RAG) and Large Language Model (LLMs). We introduce RULInsight, an explainable AI based prognostics system, combining sensor data health monitoring, predictive models, explainable AI, retrieval- augmented generation and large language models for predicting the equipment RUL, explaining the decisions of predictive models, revealing the causes of equipment health degradation and providing insightful assistance in making effective maintenance plans for reliability-centered industrial asset maintenance.

KEYWORDS: Remaining Useful Life (RUL), Predictive Maintenance, Prognostics, Explainable Artificial Intelligence (XAI), Machine Learning, Deep Learning, RAG, LLMs.

I. INTRODUCTION

Industrial systems rely heavily on continuous machine operations. Unexpected machine breakdowns are a common issue that leads to loss in productivity, increases maintenance costs, and often disrupts the entire industrial processes. Traditional predictive maintenance has been effective over the past decades, but it poses limitations by acting like a black box and lacking transparency. A key objective of predictive maintenance is to estimate the Remaining Useful Life (RUL) of industrial machines. RUL refers to the estimated number of operating cycles a machine can perform before reaching a failure condition. Many modern predictive maintenance systems using sensor data carry out analysis with machine learning and deep learning methodologies. However, these models often produce predictions without revealing how significant sensor parameters are when finding out why the machine broke down, thus limiting their practical use in maintenance decision-making.

For this issue, this paper put forwards RULInsight, an Explainable AI-driven prognostics framework for RUL prediction and maintenance decision support. The framework uses a TCN- GRU model to predict the Remaining Useful Life of industrial machines with the help of sensor readings. The proposed TCN-GRU model effectively captures both short-term and long-range temporal dependencies. TCN extracts temporal features using causal and dilated

convolutions, while GRU retains important historical information through its gating mechanism for precise RUL prediction. Explainable Artificial Intelligence (XAI) techniques are used, notably SHAP is then used to identify the sensor features that contribute most to the prediction. Retrieval-Augmented Generation (RAG) retrieves relevant information from maintenance manuals, and a Large Language Model (LLM) transforms the retrieved information and prediction results into clear, human understandable explanations and maintenance recommendations.

The proposed solution acts as a bridge, enabling technicians to understand the core problem and take appropriate, actionable measures to address it.

II. RELATED WORK

Singh and Bobde (2025) For metro rail systems proposed an explainable condition aware maintenance framework with the objective of improving fault detection and finding out the Remaining Useful Life. MetroPT-3 was used and further performed fault diagnosis and RUL estimation. The authors evaluated this using many AI-driven predicted approaches like XGBoost, Random Forest classifier, Bagging, and Stacking, along with deep learning approaches like Long Short Term Memory, Gated Recurrent Unit, ConvNet, and also hybrid Temporal Convolutional Network (TCN) with a Gated Recurrent Unit (GRU) architecture. Explainability was incorporated using SHAP and LIME to interpret the model predictions. Results showed that the Stacking model achieved approximately 99.95% classification accuracy, while the hybrid TCN-GRU model provided the most accurate RUL predictions. Although the framework showed excellent performance, it was tested only on one metro train subsystem, focused mainly on-air leakage faults, and was not evaluated in a real-time industrial environment.

Natarajan et al. (2026) proposed an integrated CNN-Transformer-BiLSTM architectural framework to improve Remaining Useful Life prediction for aero-engines while making the predictions easier to understand. The model used CNN, Transformer, and BiLSTM networks and was evaluated using the NASA C- MAPSS dataset. ConvNet served as local feature extractor, the Transformer captured long-term relationships, and BiLSTM learned temporal degradation patterns. SHAP and Integrated Gradients were applied to explain the prediction results. The proposed model achieved an RMSE which was 14.603, an MAE which was 12.089, an R^2 value which was 0.848, and a health-state classification accuracy of 95%. Despite its strong performance, the framework was evaluated only on benchmark datasets, has higher computational complexity, and does not provide maintenance recommendations or support real-time industrial deployment.

George et al. (2025) The study introduced a Digital Twin framework integrated with AI to support predictive maintenance of aircraft engines. Their system estimates the Remaining Useful Life (RUL), detects faults, explains the prediction results, and provides real-time visualization of the engine. The study used the NASA C-MAPSS dataset along with synthetic IoT sensor data. Algorithms like Random Forest were used for RUL prediction, Isolation Forest for fault detection, and SHAP for explaining the model's decisions. The framework also included blockchain for secure data storage, federated learning for privacy, for visualization, the framework employed Unity 3D, while Flutter was used to provide real-time monitoring through a mobile application. The proposed model predicted errors of 18.7 cycles (MAE) and 29.9 cycles (RMSE), indicating satisfactory prediction accuracy.

While the anomaly detection model obtained an F1-score of 0.89. However, the framework was tested only on benchmark and synthetic datasets and was not validated on real industrial systems. It also does not provide maintenance recommendations.

Xiangyu Pan et al. (2025) Proposed a model that included a framework incorporating an enhanced TCN and BiGRU for prediction of the remaining service life of cutting tools. The research focused on improving prediction accuracy through enhanced feature extraction and the capture of long-term temporal dependencies in tool wear data. The proposed architecture incorporated a parallel pooling strategy and a Convolutional Block Attention Module (CBAM) within the TCN for multi-scale degradation feature extraction with reduced noise, while the BiGRU network effectively learned bidirectional temporal relationships. The model was evaluated using the PHM2010 tool wear dataset and assessed with MAE, MSE, and R metrics. Experimental results showed superior predictive performance over CNN, TCN-LSTM, and CNN-BiGRU models, yielding R values of 0.9978, 0.9843, and 0.9886 for different cutting tools, respectively. Nevertheless, the framework was validated only on a benchmark tool wear dataset, lacked explainability mechanisms, and did not explore real-time industrial implementation or cross-domain generalization.

Sourajit Behera (2026) introduced an explainable deep learning framework, named RSNet (Residual-Attention BiGRU Network), to forecast the Remaining Useful Life of aero engines operating under multiple fault modes and varying operating conditions. The study aimed to improve prediction accuracy while increasing the model's transparency for practical prognostics and health management applications. The proposed framework integrated Bidirectional Gated Recurrent Units (BiGRU), residual learning, self-attention mechanisms, and SHapley Additive exPlanations (SHAP) to effectively capture long-term temporal dependencies, highlight degradation-related sensor information, and provide explanations for the contribution of individual sensors to RUL estimation. The model was validated on the NASA C-MAPSS dataset and evaluated using RMSE and PHM Score metrics. Experimental results indicated that RSNet outperformed several state-of-the-art RUL prediction models on all benchmark datasets, with SHAP-based analysis offering interpretable sensor-level insights to aid maintenance decision-making. However, the framework was validated exclusively on benchmark aero-engine datasets, and its applicability in real-world industrial environments, computational efficiency, and cross-domain generalization remain to be investigated further.

Xu (2026) presented a multimodal machine learning approach for estimating the operational lifespan of an advanced CNC spindle equipment by combining information from multiple industrial data sources. Instead of relying on a single type of input, the framework integrated vibration, temperature, and current sensor readings with process parameters and maintenance records utilizing a Cross-Modal Attention-driven CNN-LSTM approach (CMA-CNN-LSTM) model. Before training, the collected data underwent several preprocessing steps, including missing-value imputation using LSTM, anomaly filtering with the Isolation Forest algorithm, timestamp alignment, and 8-bit model quantization to support efficient edge deployment. The proposed framework was assessed using data collected over 14 months from 12 CNC machines and produced promising results, achieving an RMSE of 28.4 hours, MAPE of 8.2%, and an R^2 value of 0.90 while lowering maintenance cost, reducing spindle downtime, and improving prediction efficiency compared with conventional single-modal methods. Despite these encouraging results, the study was evaluated only on ISO-10816 Class-3 CNC spindle systems, which restricts its applicability to other categories of industrial equipment. Moreover, the framework generates deterministic RUL predictions without estimating prediction uncertainty and depends on the continuous availability of multiple sensing modalities, potentially affecting its robustness in real-world industrial deployments.

Yakoot and Seddik (2026) developed a stage aware LSTM framework for forecasting the residual Life of rolling element bearings by distinguishing the stable operating stage from the degradation stage. The methodology extracted both time-domain statistical features, including Root-Mean-Square (RMS) and Kurtosis, and Discrete Wavelet Transform (DWT)-based wavelet energy to construct a representative health indicator. A First Predicting Time (FPT) strategy was adopted to initiate prediction only after the bearing entered the degradation phase, while a stacked LSTM network with Monte Carlo Dropout was used for sequence learning and uncertainty estimation. The developed framework was evaluated using the IMS Bearing Dataset and achieved an RMSE of 0.1357, demonstrating reliable prediction performance. However, the framework depends on a manually selected FPT, experiences higher prediction uncertainty during the final degradation stage, and was evaluated only on the IMS Bearing Dataset, which limits its generalization to other industrial applications.

Jacob and Maheswari (2026) presented an AI-based predictive maintenance framework for induction motor fault diagnosis using a hybrid CNN-BiLSTM model enhanced with an attention mechanism. The preprocessing phase included median-value imputation, feature normalization, SMOTE-based data balancing, and extraction of informative features from time-based domain, frequency (FFT), and wavelet-based representation signals. Multi-scale Convolutional Neural Networks (CNNs) were employed to draw out the spatial features, while Bidirectional Long Short-Term Memory (BiLSTM) was used to capture temporal dependencies, and the attention mechanism highlighted the most informative features during classification. The model achieved 98.60% accuracy, while ML and DL methods were outperformed by this model such as high precision, recall, and the F1-score. However, the proposed framework was validated mainly on benchmark datasets rather than real industrial environments. In addition, slight confusion between bearing and stator faults with similar spectral characteristics was observed, and the model is limited to fault diagnosis without supporting Remaining Useful Life (RUL) prediction.

Patrick Seebold et al. (2024) proposed an Explainable Artificial Intelligence (XAI) framework for Remaining service Life prediction. The reported work sought to improve the transparency of AI-driven predictive maintenance models by providing interpretations for the decisions of a high-performing prediction model rather than

relying solely on prediction accuracy. The authors developed an RUL prediction model utilizing XGBoost and a neural-network- based learning strategy, while employing SHAP, PDP, ICE plots, and dependence plots for feature importance and parameter interaction analysis. The experimental evaluation achieved around 96% prediction accuracy, with Sensor 11, Sensor 4, and Sensor 7 being identified as the most impactful features influencing RUL model interpretability without negatively affecting prediction performance. However, the study was validated exclusively on the NASA C-MAPSS benchmark dataset, did not include evaluation in real-world industrial scenarios, and did not compare SHAP with other explainability methods, indicating room generalization.

Liu et al. (2024) The researchers designed a hybrid TCN–NTM framework that combines the strengths of both the models to provide more reliable Remaining Useful Life (RUL) predictions for aircraft engines. To prepare the data for training, the authors applied min-max normalization, selected informative sensors, performed piecewise RUL labeling, used sliding window method to generate sequential samples. The TCN was employed to seize degradation patterns from time-series sensor data, while the NTM utilized its external memory to retain long-term operational information. Experimental evaluation on the NASA C-MAPSS FD001 dataset demonstrated that the model yielded an RMSE which was 13.23 surpassing several conventional deep learning methods in prediction accuracy. However, the framework was evaluated only on the FD001 subset of the C-MAPSS dataset, limiting its validation under multiple operating conditions. In addition, the model does not provide prediction interpretability or uncertainty estimation, reducing its applicability in safety-critical industrial environments.

III. CONCLUSION

This paper presented RULInsight, explains the framework for prognosticating the Remaining Useful Life of industrial machines. Many existing predictive maintenance models can estimate RUL accurately, but they do not explain how the prediction is made, making it tough for the technicians to rely on the outcomes and trust the predictions. To solve this problem, the proposed framework combines sensor data with SHAP for prediction explanation, RAG for retrieving relevant information from maintenance manuals, and LLMs for generating simple and understandable maintenance recommendations. By combining prediction with clear explanations and practical guidance, RULInsight helps technicians make better maintenance decisions. The proposed framework can reduce unexpected machine failures, lower maintenance costs, and improve overall operational efficiency. The proposed framework can be further improved by incorporating live data from IoT-enabled sensors, making the whole procedure further suitable for real-world industrial applications. This would support different types of industrial machines, and continuously updating the knowledge base to improve prediction accuracy and maintenance recommendations.

IV. RESEARCH GAP

- Most RUL models are designed for specific equipment and have limited applicability across different industrial systems.
- Many deep learning models have high accuracy and provide accurate predictions but lack clear and interpretable decision explanations.
- Current explainable AI techniques like LIME and SHAP provide important features, but don't translate these explanations into actionable maintenance guidance TCN–GRU models capture temporal features effectively but lack explainability and uncertainty-aware prediction.
- Most studies focus only on RUL prediction without providing autonomous maintenance recommendations.

REFERENCES

- [1] A. Singh and S. Bobde, "Explaining Condition-Based Maintenance for Metro Trains Using Ensemble ML with XAI and Hybrid Deep Learning Models," in Proc. 2025 International Conference on Artificial Intelligence, Computing and Next Generation Networks (IC2NC), 2025, pp. 1–8.
- [2] S. Natarajan, et al., "An Explainable Hybrid CNN-Transformer-BiLSTM Framework for Remaining Useful Life Prediction of Aero-Engines," Expert Systems with Applications, vol. XX, pp. XX–XX, 2026.
- [3] C. J. George, Rajkumar J. S., and N. Varghese Babu, "Explainable Digital Twin Framework for Aircraft Engine Maintenance," in Proc. 2025 International Conference on NexGen Networks and Cybernetics (IC2NC), 2025, pp. 110–114.

- [4] X. Pan, et al., "An Improved TCN-BiGRU Network with Convolutional Block Attention for Remaining Useful Life Prediction of Cutting Tools," *Journal of Manufacturing Processes*, vol. XX, pp. XX–XX, 2025.
- [5] S. Behera, "RSNet: An Explainable Residual-Attention BiGRU Network for Remaining Useful Life Prediction of Aero-Engines," *Reliability Engineering & System Safety*, vol. XX, pp. XX–XX, 2026.
- [6] X. Xu, et al., "A Multimodal Data-Driven Machine Learning Framework for Remaining Useful Life Prediction of High-End CNC Spindle Equipment Using Cross-Modal Attention CNN-LSTM," *Journal of Manufacturing Systems*, vol. XX, no. XX, pp. XX–XX, 2026.
- [7] M. Yakoot and A. Seddik, "Phase-Aware Long Short-Term Memory Framework for Remaining Useful Life Prediction of Rolling Element Bearings," *Mechanical Systems and Signal Processing*, vol. XX, pp. XX–XX, 2026.
- [8] C. Jacob and K. Maheswari, "An Attention-Based CNN-BiLSTM Framework for AI-Driven Predictive Maintenance of Induction Motors," *Engineering Applications of Artificial Intelligence*, vol. XX, pp. XX–XX, 2026.
- [9] P. Seebold, S. Kim, M. K. Kaye, and C. S. Nam, "Explainable AI-based Shapley Additive Explanations for Remaining Useful Life Prediction using NASA Turbofan Engine Dataset," in *Proc. IEEE 3rd International Conference on Computing and Machine Intelligence (ICMI)*, 2024, pp. 1–6.
- [10] D. Liu, G. Qu, T. Qiu, and C. Liu, "A TCN-NTM Model for Remaining Useful Life Prediction for Aircraft Engine," in *Proc. IEEE International Conference on Unmanned Systems (ICUS)*, 2024, pp. 1574–1578.
- [11] X. Guo, Z. Fang, Z. Yang, and Z. Wei, "A Hybrid Approach Based on Gaussian Process Regression and LSTM for Remaining Useful Life Prediction of Lithium-ion Batteries," in *Proc. IEEE Transportation Electrification Conference and Expo (ITEC)*, 2023, pp. 1–6.
- [12] T. Tao, T. Yang, S. Liu, Y. Yang, and X. Guo, "Prediction Method for Remaining Useful Life of Rolling Bearings Based on a Multi-layer Long Short-Term Memory Model," in *Proc. 3rd International Conference on Electrical Engineering and Mechatronics Technology (ICEEMT)*, 2023, pp. 619–623.
- [13] Y. Wang, X. Zhang, L. Chen, B. Liu, C. Liu, Y. Deng, H. Liu, and Y. Zhang, "Segmented Prediction Model of Lithium-ion Battery Remaining Useful Life Based on LSTM and SVR," in *Proc. 6th International Conference on Intelligent Robotics and Control Engineering (IRCE)*, 2023, pp. 177–182.
- [14] A. Patil and J. Kendule, "IoT-based Battery Health Monitoring and Its Remaining Useful Life Prediction using Artificial Neural Network," in *Proc. IEEE International Students' Conference on Electrical, Electronics and Computer Science (SCEECS)*, 2024, pp. 1–6.
- [15] V. Sharma, R. Dagar, and S. Sharanya, "A XGBoost Optimized Ensemble Model for Remaining Useful Life Prediction of Aircraft Turbofan Engines," in *Proc. International Conference on Emerging Smart Computing and Informatics (ESCI)*, 2024, pp. 1–6.
- [16] M. Bellomo, S. Giazitzis, S. Badha, F. Rosetti, E. Ogliari, A. Dolara, and F. Grimaccia, "A Novel Data-Driven Approach for the Simultaneous Prediction of State of Health Trajectory and Remaining Useful Life in Lithium-Ion Batteries," in *Proc. IEEE International Conference on Environment and Electrical Engineering (EEEIC)*, 2024, pp. 1–6.
- [17] C. Chen, L. Xue, T. Wang, Y. Li, and K. Noman, "Remaining Useful Life Prediction of Wind Turbine High Speed Shaft Bearing by Sparsity Measure Based Long Short-Term Memory Network," in *Proc. 5th International Conference on Measurement, Control and Instrumentation (MCAI)*, 2025, pp. 421–426.
- [18] Y. Li, J. Feng, S. Lu, B. Wang, and Y. Yao, "Power Transformer Remaining Useful Life Prediction Method Based on Dual Imbalanced Data," in *Proc. IEEE 14th Data Driven Control and Learning Systems Conference (DDCLS)*, 2025, pp. 1–6.
- [19] H. H. M. Said and A. H. Kassem, "Remaining Useful Life Prediction Driven by LLMs Using Few-Shot Learning," in *Proc. 35th International Conference on Computer Theory and Applications (ICCTA)*, 2025, pp. 198–204.
- [20] X. Ren, Y. Zhou, J. Tu, and C. Li, "LLM-Enhanced Transformer Framework for Accurate Lithium Battery Remaining Useful Life Prediction," in *Proc. 15th Prognostics and System Health Management Conference (PHM)*, 2025, pp. 1–6.
- [21] Harini R. A. J., G. Selvan, M. Priyanka, M. RamKumar, and Subhashini M., "Explainable Agentic AI for Autonomous Battery Health Prediction and Management," in *Proc. 3rd International Conference on Networks and Communications (ICNWC)*, 2026, pp. 1–6.



INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 9940 572 462  6381 907 438  ijirset@gmail.com



www.ijirset.com

Scan to save the contact details