



Artificial Intelligence-Based Predictive Maintenance Framework for Industrial Automation Systems

Farheen Leroux-Martin

Department of Computer Science and Engineering, Daehan Institute of Management and Logistics, South Korea

sudarshan.rafizadeh@misem-ph.org

Peer Review Information

Submission: 30 Jan 2020

Revision: 09 Feb 2020

Acceptance: 26 March 2020

Keywords

*Artificial Intelligence,
Predictive Maintenance,
Industrial Automation,
Machine Learning,
Condition Monitoring*

Abstract

Predictive maintenance has become a key requirement in industrial automation because conventional corrective and time-based preventive maintenance may either react too late or replace components before their useful life is exhausted. This paper proposes an artificial intelligence (AI)-based predictive maintenance framework that integrates multi-sensor condition monitoring, data preprocessing, feature engineering, machine-learning classification, maintenance decision support, and feedback to programmable logic controller (PLC) or supervisory control and data acquisition (SCADA) environments. The methodology is designed for rotating and process equipment monitored through vibration, temperature, current, pressure, acoustic, operating-load, and runtime variables. A controlled simulation is used to demonstrate the evaluation procedure without presenting synthetic measurements as real plant data. Four models—logistic regression, support vector machine, random forest, and a multilayer perceptron neural network—are evaluated using accuracy, precision, recall, and F1-score. The proposed workflow emphasizes explainable maintenance triggers, periodic retraining, and integration with existing industrial automation architecture. The study indicates that ensemble and neural approaches are suitable for detecting nonlinear degradation patterns, while data quality, class imbalance, model drift, and integration with maintenance practice remain important deployment constraints.

Introduction

Industrial automation systems depend on motors, pumps, bearings, gearboxes, conveyors, actuators, drives, and process-control equipment whose unexpected failure can interrupt production and affect safety, product quality, and cost. Condition-based maintenance established the principle that maintenance decisions should be linked to observed equipment condition rather than fixed intervals; diagnostics and prognostics consequently became central to maintenance planning (Jardine, Lin, & Banjevic,

2006). Later work on machine prognostics emphasized remaining useful life (RUL) as a decision variable for anticipating maintenance before functional failure (Peng, Dong, & Zuo, 2010).

The growth of industrial sensing, embedded computation, cyber-physical systems, and Industry 4.0 has made continuous equipment data available at a scale suitable for AI. Machine learning can learn relationships between sensor patterns and failure states that are difficult to represent with fixed engineering thresholds. Susto et al. (2015) demonstrated that multiple

classifiers can support predictive maintenance decisions across different prediction horizons. More broadly, modern prognostics pipelines combine acquisition, health indicators, degradation assessment, prediction, and decision support (Lei et al., 2018).

For industrial automation, however, predictive accuracy alone is insufficient. A deployable framework must fit within the control and maintenance environment, transform heterogeneous sensor streams into stable features, identify abnormal operating states early enough for intervention, and communicate actionable information to maintenance personnel. It must also avoid excessive false alarms because unnecessary maintenance can consume the same resources that predictive maintenance is intended to save. Data-driven surveys have therefore highlighted the importance of selecting algorithms according to available data, failure modes, prediction horizon, and industrial constraints (Carvalho et al., 2019; Zhang, Yang, & Wang, 2019).

This paper addresses these requirements through a structured AI-based predictive maintenance framework for industrial automation systems. The objective is not to claim a new production-scale deployment, but to present a reproducible methodology that can be adapted to a real plant. The framework combines sensing, edge or gateway acquisition, data cleaning, feature extraction, model training, online fault-risk estimation, maintenance prioritization, and feedback. It also distinguishes fault detection from maintenance action: a model estimates risk, while a decision layer converts risk into alarms, inspection requests, or planned maintenance according to asset criticality and confidence.

Literature Review

Jardine et al. (2006) provided a foundational review of machinery diagnostics and prognostics for condition-based maintenance, organizing the field around data acquisition, data processing, diagnostics, and prognostics. Their structure remains directly relevant because AI does not eliminate the need for reliable sensing and signal processing; instead, it adds data-driven inference to those stages. Peng et al. (2010) reviewed machine prognostics and emphasized the difficulty of estimating future degradation under uncertain operating conditions, motivating approaches that learn from historical condition data.

The use of machine learning for predictive maintenance became more explicit as industrial datasets expanded. Susto et al. (2015) proposed a multiple-classifier methodology that relates prediction horizons to maintenance costs. Lei et al. (2018) later summarized machinery health prognostics from data acquisition through RUL prediction and highlighted data-driven methods as a major branch of modern prognostics. These studies support a methodology in which sensor selection, feature representation, model choice, and decision criteria are designed together rather than as separate tasks.

Carvalho et al. (2019) systematically reviewed machine-learning methods used in predictive maintenance and showed that algorithm selection is strongly dependent on the type and quality of available industrial data. Deep models also became important for extracting degradation features directly from high-dimensional measurements. Yang, Liu, and Zio (2019), for example, demonstrated a double-convolutional architecture for RUL prediction, illustrating how deep learning can model nonlinear degradation without relying entirely on handcrafted indicators. Zhang et al. (2019) similarly surveyed data-driven predictive-maintenance methods and linked their growth to the broader development of smart industrial systems.

By 2020, predictive maintenance was increasingly framed as an Industry 4.0 capability rather than an isolated analytics task. Zonta et al. (2020) catalogued methods, standards, and applications in a systematic review and noted the heterogeneity of predictive-maintenance implementations. Dalzochio et al. (2020) focused on machine learning and reasoning frameworks, including the role of structured knowledge in maintenance decision support. Çınar et al. (2020) discussed machine learning in predictive maintenance for sustainable smart manufacturing and emphasized both the potential of AI and the importance of suitable industrial data. Together, these studies motivate an integrated framework that connects AI analytics to automation and maintenance operations rather than treating model training as the endpoint.

Methodology

The proposed methodology follows a design-and-evaluation approach. It defines an architecture for acquiring machine-condition data, constructing fault-prediction models, and converting model outputs into maintenance

decisions. The method is intentionally modular so that the same pipeline can be adapted to motors, pumps, bearings, conveyors, industrial robots, or process equipment by changing sensors, features, and fault labels while preserving the decision flow.

1. Data Acquisition and Variables

A typical automated asset is instrumented with vibration, temperature, motor-current, pressure, acoustic, runtime, and operating-load measurements. Signals can be obtained from dedicated condition-monitoring sensors, PLC analog inputs, smart drives, or IIoT gateways. Data are time-stamped and synchronized to create fixed observation windows. For a real deployment, the sampling rate is selected according to the physics of the fault: high-frequency vibration may require kilohertz sampling, whereas temperature, pressure, and load can often be sampled at lower rates. The proposed framework stores both raw or aggregated condition measurements and maintenance-event labels so that the model can learn the relationship between degradation and future failure.

2. Data Preprocessing and Feature Engineering

Preprocessing contains five operations: missing-value handling, sensor-range validation, noise filtering, synchronization, and normalization. Outliers that clearly represent sensor errors are removed or imputed, while legitimate abnormal values are preserved because they may contain fault information. For rotating equipment, time-domain features such as root-mean-square vibration, kurtosis, crest factor, mean temperature, current variation, and load are calculated per observation window. Frequency-domain features can be added when bearing or gearbox signatures are important. Feature

scaling is used for algorithms such as support vector machines and neural networks.

3. Target Definition and Dataset Split

The classification target is defined as whether a maintenance-relevant failure is expected within a selected prediction horizon, for example the next 24 operating hours. Historical maintenance records are aligned with condition windows to create positive and negative labels. To prevent information leakage, training and test partitions should be separated by machine or by chronological period rather than randomly mixing highly correlated samples from the same failure episode. In the illustrative validation used in this paper, 7,000 synthetic multivariate observations are generated from controlled degradation and operating-load variables, then split into 75% training and 25% testing using a fixed random seed. The synthetic experiment is used only to demonstrate the evaluation procedure.

4. Candidate AI Models

Four representative models are evaluated: logistic regression as an interpretable linear baseline, a radial-basis-function support vector machine for nonlinear class boundaries, random forest as an ensemble model with robust nonlinear interactions, and a multilayer perceptron (MLP) neural network. The baseline is useful for judging whether complex AI is actually required. The random forest also supports feature-importance analysis, while the MLP represents a compact neural model suitable for structured sensor features. In a deployment with long sequential windows, recurrent or convolutional models may be added for direct temporal learning, consistent with the deep prognostic's direction reported by Yang et al. (2019).

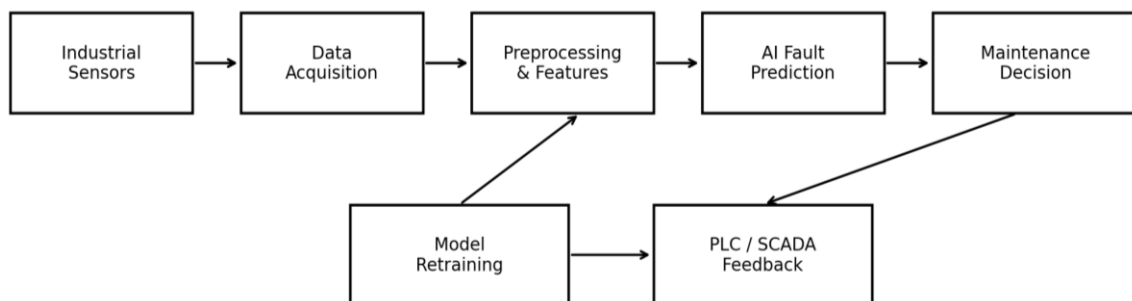


Figure 1. Proposed AI-based predictive maintenance framework Online Inference and Maintenance Decision Logic

After model validation, the selected model is deployed at an edge computer, industrial PC, server, or gateway connected to the automation network. Each new sensor window receives a fault-risk score. A risk score below the normal threshold produces no action; an intermediate score creates an inspection recommendation; and a high score sustained across multiple windows generates a maintenance alarm. Asset criticality modifies this policy: a moderate probability on a safety-critical machine may justify action sooner than the same probability on a redundant low-cost asset. This separation between prediction and decision prevents a single model probability from automatically stopping production.

The maintenance result is fed back into the data store. Confirmed faults, replaced components, false alarms, and no-fault inspections become new labels for later model retraining. This feedback loop addresses model drift caused by aging equipment, process changes, new products, sensor replacement, or maintenance practice. The framework therefore treats predictive maintenance as a lifecycle rather than a one-time model-building activity, which is consistent with the integrated Industry 4.0 perspective summarized by Zonta et al. (2020), Dalzochio et al. (2020), and Çınar et al. (2020).

5. Experimental Evaluation Plan

Table 1. Evaluation workflow for the proposed framework.

Stage	Input	Method	Output
Acquisition	Sensor and PLC signals	Windowing and synchronization	Condition dataset
Preprocessing	Raw/aggregated signals	Cleaning, validation, scaling	Model-ready features
Training	Features + failure labels	LR, SVM, RF, MLP	Candidate models
Validation	Held-out test data	Accuracy, precision, recall, F1	Model comparison
Deployment	Live sensor windows	Risk scoring + persistence rule	Maintenance alert
Feedback	Inspection/repair outcome	Relabeling and retraining	Updated model

Performance is evaluated with accuracy, precision, recall, and F1-score. Accuracy summarizes total correct predictions, but it can be misleading when failures are rare. Precision measures how many predicted failures are true failures and is therefore important for controlling unnecessary maintenance. Recall measures how many actual failures are detected and is important for avoiding unexpected breakdowns. F1-score provides a balanced summary of precision and recall. In a plant deployment, additional measures should include warning lead time, false alarms per operating week, downtime avoided, maintenance cost, and model-inference latency.

Results

The results below are obtained from the controlled simulation described in the methodology. They are not measurements from a specific factory and should be interpreted as a demonstration of how the proposed framework can be evaluated before deployment. The synthetic dataset contains correlated degradation, operating-load, vibration, temperature, current, pressure, acoustic, and runtime variables. Failure labels are generated from a nonlinear degradation-risk rule with noise, which prevents perfect separation and allows meaningful model comparison.

Table 2. Simulation-based comparison of candidate predictive-maintenance models.

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
Logistic Regression	95.31	86.34	76.62	81.19
SVM (RBF)	95.03	86.00	74.46	79.81
Random Forest	95.89	85.33	83.12	84.21
MLP Neural Network	95.77	86.85	80.09	83.33

The random forest produced the strongest balance between missed failures and false alarms, with an accuracy of 95.89% and F1-score of 84.21%. The MLP neural network achieved a similar overall accuracy but a slightly lower F1-

score. Logistic regression obtained high overall accuracy because the non-failure class is more frequent, yet its lower recall indicates that a larger fraction of failure cases remained undetected. This illustrates why predictive-

maintenance studies should not select models using accuracy alone.

Conclusion and Discussion

This paper developed a methodology for an AI-based predictive maintenance framework for industrial automation systems. The framework combines condition sensing, synchronized data acquisition, preprocessing, feature engineering, AI-based fault prediction, maintenance decision logic, and feedback to the plant data environment. Its main methodological contribution is the separation of three functions that are often mixed together: condition measurement, failure-risk estimation, and maintenance action. This separation makes the approach easier to integrate with PLC/SCADA systems and safer for industrial use because AI predictions support, rather than replace, established control and maintenance procedures. The simulation-based validation demonstrates how candidate models can be compared using metrics that reflect maintenance priorities. In the controlled experiment, random forest provides the best F1-score, while the neural network performs closely and offers a path toward more complex temporal models. The result is consistent with the broader literature showing that no single algorithm is universally best for predictive maintenance; performance depends on the equipment, failure mode, available sensors, label quality, and prediction horizon (Susto et al., 2015; Carvalho et al., 2019). Consequently, the proposed framework requires model benchmarking rather than selecting an algorithm in advance based only on popularity. The literature up to 2020 also indicates that predictive maintenance should be understood as an Industry 4.0 system capability. Reviews by Zonta et al. (2020) and Dalzochio et al. (2020) show that successful solutions require more than machine learning: standards, knowledge representation, data engineering, integration, and maintenance reasoning all influence deployment. Çınar et al. (2020) similarly links predictive maintenance to smart and sustainable manufacturing, where reduced unplanned downtime and better use of component life can improve operational efficiency. Several implementation challenges remain. First, industrial data are frequently imbalanced because normal operation is much more common than failure. Second, sensor drift, process changes, and maintenance interventions can change the statistical distribution of data, requiring periodic validation and retraining. Third, labels may be uncertain when maintenance records describe symptoms rather

than confirmed root causes. Fourth, false alarms can reduce operator trust, so model outputs should include confidence, trend persistence, and interpretable evidence wherever possible. Fifth, cybersecurity and network segmentation must be considered when condition data cross operational-technology and information-technology boundaries. Future work should validate the framework on real industrial assets using chronologically separated training and test periods, compare fault classification with RUL regression, and include cost-sensitive decision thresholds. Deep convolutional or recurrent models may be evaluated for raw vibration and multivariate time-series data, while explainability methods can be used to show which sensor patterns drive each maintenance recommendation. A further extension is to connect predictions to a computerized maintenance management system so that risk scores can support inspection scheduling, spare-parts planning, and work-order prioritization. With these additions, the proposed framework can serve as a practical bridge between AI research and industrial maintenance execution.

References

- 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. <https://doi.org/10.1016/j.ymssp.2005.09.012>
- Peng, Y., Dong, M., & Zuo, M. J. (2010). Current status of machine prognostics in condition-based maintenance: A review. *The International Journal of Advanced Manufacturing Technology*, 50, 297–313. <https://doi.org/10.1007/s00170-009-2482-0>
- Susto, G. A., Schirru, A., Pampuri, S., McLoone, S., & Beghi, A. (2015). Machine learning for predictive maintenance: A multiple classifier approach. *IEEE Transactions on Industrial Informatics*, 11(3), 812–820. <https://doi.org/10.1109/TII.2014.2349359>
- Lei, Y., Li, N., Guo, L., Li, N., Yan, T., & Lin, J. (2018). Machinery health prognostics: A systematic review from data acquisition to RUL prediction. *Mechanical Systems and Signal Processing*, 104, 799–834. <https://doi.org/10.1016/j.ymssp.2017.11.016>
- Carvalho, T. P., Soares, F. A. A. M. N., Vita, R., Francisco, R. P., Basto, J. P. V., & Alcalá, S. G. S.

(2019). A systematic literature review of machine learning methods applied to predictive maintenance. *Computers & Industrial Engineering*, 137, 106024. <https://doi.org/10.1016/j.cie.2019.106024>

Yang, B., Liu, R., & Zio, E. (2019). Remaining useful life prediction based on a double-convolutional neural network architecture. *IEEE Transactions on Industrial Electronics*, 66(12), 9521–9530. <https://doi.org/10.1109/TIE.2019.2924605>

Zhang, W., Yang, D., & Wang, H. (2019). Data-driven methods for predictive maintenance of industrial equipment: A survey. *IEEE Systems Journal*, 13(3), 2213–2227. <https://doi.org/10.1109/JSYST.2019.2905565>

Zonta, T., da Costa, C. A., da Rosa Righi, R., de Lima, M. J., da Trindade, E. S., & Li, G. P. (2020).

Predictive maintenance in the Industry 4.0: A systematic literature review. *Computers & Industrial Engineering*, 150, 106889. <https://doi.org/10.1016/j.cie.2020.106889>

Dalzochio, J., Kunst, R., Pignaton, E., Binotto, A., Sanyal, S., Favilla, J., & Barbosa, J. (2020). Machine learning and reasoning for predictive maintenance in Industry 4.0: Current status and challenges. *Computers in Industry*, 123, 103298. <https://doi.org/10.1016/j.compind.2020.103298>

Çınar, Z. M., Nuhu, A. A., Zeeshan, Q., Korhan, O., Asmael, M., & Safaei, B. (2020). Machine learning in predictive maintenance towards sustainable smart manufacturing in Industry 4.0. *Sustainability*, 12(19), 8211. <https://doi.org/10.3390/su12198211>