electronics

Article

Predictive Modeling of Failure States in Manufacturing Systems Using Artificial Intelligence in the Context of Sustainability

Miroslav Rakyta , Peter Bubenik , Vladimira Binasova * and Martin Buzalka

Department of Industrial Engineering, Faculty of Mechanical Engineering, University of Zilina, Univerzitna 8215/1, 010 26 Zilina, Slovakia; miroslav.rakyta@fstroj.uniza.sk (M.R.); peter.bubenik@fstroj.uniza.sk (P.B.); martin.buzalka@fstroj.uniza.sk (M.B.)

* Correspondence: vladimira.binasova@fstroj.uniza.sk; Tel.: +421-041-513-2727

Abstract

This study investigates predictive modeling of failure states in manufacturing systems using artificial intelligence in the context of sustainable maintenance and Industry 4.0. The proposed methodological framework is based on historical operational and maintenance data from a single manufacturing device, encompassing multiple process and operational signals such as vibrations, temperature, electric current, and operational logs. The aim is to predict failure within a short-term horizon to support maintenance planning. The article compares Random Forest and XGBoost algorithms at different prediction horizons (8 h and 16 h) to identify the trade-off between classification accuracy and lead time for maintenance planning. Model outputs are analyzed using explainable artificial intelligence and transformed into a risk index compatible with the FMEA methodology. The practical contribution of the proposed approach is illustrated through a scenario-based what-if assessment of potential sustainability impacts, particularly in terms of estimated reductions in unplanned downtime, material waste, and energy consumption. The results point to the potential of integrating AI-supported maintenance as a tool for increasing the reliability and sustainability of manufacturing systems. The novelty of the proposed framework lies in the integration of predictive maintenance, explainable artificial intelligence, replay-based maintenance assessment, dynamic FMEA risk assessment, and sustainability impact quantification into a unified decision-support framework.

Keywords: predictive maintenance; XGBoost; random forest; SHAP; replay analysis; FMEA; sustainability; Industry 4.0

Academic Editor: Mohamed Benbouzid

Received: 30 June 2026
Revised: 13 July 2026
Accepted: 16 July 2026
Published: 21 July 2026

1. Introduction

Predictive maintenance represents a key element of modern manufacturing systems in the context of Industry 4.0 and industrial digitalization [1,2]. Traditional maintenance approaches based on reactive or preventive strategies often lead to inefficient resource utilization and increased operational costs [3].

In recent years, research attention has shifted toward the use of artificial intelligence and machine learning for predicting failure states and optimizing maintenance processes [4–7]. These approaches enable the analysis of historical data and the identification of patterns leading to failures.

Despite significant progress, most existing studies focus primarily on model accuracy, while practical implementation, interpretability, and integration with maintenance decision-making remain insufficiently developed [8–10].

 https://doi.org/10.3390/electronics15143194

Simultaneously, the need to link predictive maintenance with the concept of sustainability is growing [11–15], encompassing the optimization of energy consumption, waste reduction, and increased efficiency of manufacturing processes [16,17].

The objective of this study is to propose a methodological framework that combines:

- predictive modeling using AI,
- interpretation of results using XAI,
- integration with the FMEA methodology,
- quantification of the impact on sustainability.

The main novelty of the proposed framework lies in establishing a time-dependent maintenance decision mechanism that operationally connects AI-based failure prediction, explainable model outputs, replay-based identification of intervention timing, and dynamic FMEA-compatible risk assessment [18–24].

Unlike conventional approaches in which predictive models, explainability methods, and FMEA are applied as separate analytical components, the proposed framework continuously transforms AI-generated failure probabilities into a dynamic maintenance risk indicator and links the resulting risk level with dominant degradation factors and available intervention time. This integration creates a direct decision-support pathway from failure prediction to interpretable risk prioritization and maintenance action planning.

2. Related Work

Predictive maintenance is an intensively studied area within intelligent manufacturing. Several studies confirm that machine learning algorithms, such as Random Forest, XGBoost, or neural networks, are capable of effectively predicting failure states [3,7].

The XGBoost algorithm ranks among the most powerful methods for processing tabular data and achieves high accuracy in classification tasks [4]. Random Forest represents a robust reference model [5,6].

Model interpretability is addressed using explainable artificial intelligence (XAI) methods, such as SHAP (SHapley Additive exPlanations) values, which enable analysis of the influence of individual variables on the model output [25–28]. Recent studies emphasize the growing importance of explainability in predictive maintenance and smart manufacturing, particularly for increasing transparency, supporting standardized evaluation of explanation methods, and facilitating the practical adoption of AI-based decision-support systems [29–31].

Explainability has also gained increasing attention in prognostic applications, including remaining useful life prediction, where interpretable AI methods can support a better understanding of degradation processes and the factors influencing model predictions [30].

In the area of sustainability, research focuses on reducing energy requirements and optimizing the lifecycle of equipment [16].

Nevertheless, a research gap exists, particularly in the areas of:

- integration of AI models with FMEA,
- practical implementation of replay analysis,
- quantification of the impact on sustainability,
- distinction between technical model validation and real applicability on time-ordered data.

The methodological contribution of this study does not lie in proposing a new machine learning algorithm, explainability method, or FMEA technique in isolation. Instead, the proposed framework introduces a time-dependent maintenance decision mechanism in which AI-generated failure probabilities are continuously transformed into a dynamic FMEA-compatible risk indicator and subsequently linked with explainable model outputs, replay-based intervention timing, and maintenance decision rules. In contrast to conventional static FMEA assessment, the occurrence-related component of risk is dynamically updated according to the current predicted technical condition of the equipment. The main methodological contribution therefore lies in establishing an explicit operational link between failure prediction, explanation of dominant degradation factors, intervention timing, and dynamic maintenance risk prioritization.

3. Materials and Methods

The primary objective of the proposed methodology is to develop, validate, and interpret a predictive maintenance framework capable of forecasting industrial equipment failures before they occur. To achieve this, a comprehensive data-driven approach was designed, combining advanced machine learning, rigorous feature engineering, and explainable artificial intelligence (XAI).

3.1. Data Collection and Characterization

The proposed methodological approach is based on historical operational and maintenance data from a single manufacturing device. The analyzed period covered operational data of the manufacturing device for January, February, and March 2026. The input data consist of heterogeneous temporal and event records that include process, operational, maintenance, and energy variables. The dataset was collected from a single device over a three-month period and contained 27 registered failure events. This limited dataset size constrains the statistical robustness of model evaluation and may limit generalizability to different devices, operating regimes, and industrial conditions. Therefore, the present study should be interpreted as a case-study-based evaluation of the proposed decision-support framework. The main monitored variables include:

- process variables (vibrations, temperature, pressure, electric current),
- operational data (rotational speed, cycles, downtime),
- maintenance records (failure type, failure duration, response time, failure cause, corrective action, costs, total downtime),
- energy indicators (energy consumption, peak load),
- quality indicators (defects, OEE, process capability index).

Data were resampled to a uniform 10-min time step to ensure consistent temporal alignment of sensor measurements and subsequent feature engineering.

The proposed solution framework is illustrated in Figure 1. The diagram presents the logical flow from historical data through feature engineering and model training to interpretation, replay analysis, FMEA risk index, and sustainability impact assessment.

3.2. Definition of the Prediction Problem

The prediction problem was formulated as a classification task, the aim of which is to estimate the probability of failure occurrence within a defined time horizon. For each time instant t, a binary target attribute was defined:

$$y_t = \begin{cases} 1, & \text{if a failure occurs within the interval } [t, t + \Delta t] \\ 0, & \text{otherwise} \end{cases} \quad (1)$$

where

- y_t is the binary target variable,
- Δt represents the prediction horizon.

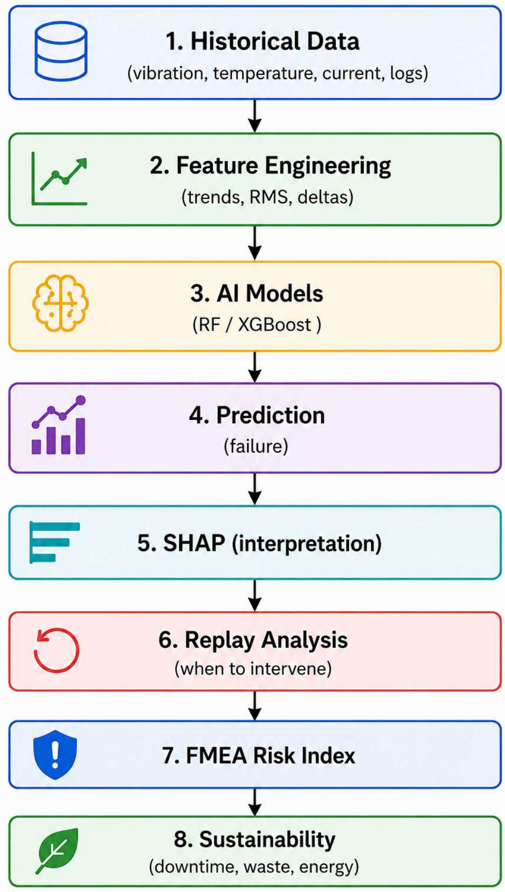

Figure 1. Schematic of the proposed solution (flow from historical data to sustainability impact assessment).

3.3. Feature Engineering

In addition to raw sensor measurements, derived features were generated to capture system dynamics and temporal evolution of operating conditions. Feature engineering was performed exclusively using historical information available at the prediction time. Future observations were not used during the construction of rolling or trend-based variables in order to eliminate information leakage and ensure realistic predictive maintenance conditions.

The generated features were designed to capture not only the instantaneous operating condition of the equipment but also short-term variations, longer-term degradation trends, and interactions between process variables. The engineered feature set included trend indicators, rolling statistics, first-order differences, and ratio-based indicators. Table 1 summarizes the complete set of engineered features used for AI model training.

The engineered features were designed to capture both short-term anomalies and long-term degradation patterns. Rolling averages represent gradual condition changes, first-order differences identify abrupt deviations, and trend features capture progressive deterioration over longer operating periods. The most important engineered features included:

- trend indicators (e.g., increase in vibration levels over the previous one hour),
- temporal differences (e.g., temperature change over one hour),
- rolling averages (rolling mean),
- statistical characteristics (e.g., RMS and variance),
- combined indicators (e.g., the ratio between vibration and temperature measurements).

Table 1. Raw and Engineered Features Used for Failure Prediction Model Development.

Feature	Feature Type	Calculation	Purpose for Failure Prediction
vibrations	Raw sensor input	Direct measurement	Mechanical stability of the equipment
bearing_temperature	Raw sensor input	Direct measurement	Thermal loading of the bearing
motor_current	Raw sensor input	Direct measurement	Electrical loading of the motor
motor_power	Raw sensor input	Direct measurement	Operating load of the system
rpm	Raw sensor input	Direct measurement	Operating regime of the equipment
pressure	Raw sensor input	Direct measurement	Process condition monitoring
flow_rate	Raw sensor input	Direct measurement	Process operating conditions
humidity_pct	Raw sensor input	Direct measurement	Environmental influence
ambient_temp_c	Raw sensor input	Direct measurement	Ambient operating conditions
health_index	Raw sensor input	Direct measurement	Overall equipment condition
vibrations_rollmean	Rolling average	$\bar{v}_t = \frac{1}{6}\sum_{i=0}^{5} v_{t-i}$	Long-term vibration trend identification
bearing_temperature_rollmean	Rolling average	$\bar{T}_t = \frac{1}{6}\sum_{i=0}^{5} T_{t-i}$	Thermal degradation trend monitoring
motor_current_rollmean	Rolling average	$\bar{I}_t = \frac{1}{6}\sum_{i=0}^{5} I_{t-i}$	Electrical loading trend identification
vibrations_delta1	First-order difference	$\Delta v_t = v_t - v_{t-1}$	Sudden vibration changes
bearing_temperature_delta1	First-order difference	$\Delta T_t = T_t - T_{t-1}$	Sudden thermal changes
motor_current_delta1	First-order difference	$\Delta I_t = I_t - I_{t-1}$	Sudden electrical load variations
vibrations_trend6	Trend feature	$TrendV_t = v_t - v_{t-6}$	Longer-term vibration development
bearing_temperature_trend6	Trend feature	$TrendT_t = T_t - T_{t-6}$	Longer-term thermal evolution
motor_current_trend6	Trend feature	$TrendI_t = I_t - I_{t-6}$	Longer-term motor loading evolution
vib_temp_ratio	Derived ratio	$\frac{v_t}{T_t}$	Mechanical–thermal degradation indicator
power_current_ratio	Derived ratio	$\frac{P_t}{I_t}$	Electrical operating efficiency indicator

v_t denotes vibration amplitude, T_t bearing temperature, I_t motor current, P_t motor power, and t the current observation time. Rolling averages represent six-step moving averages, first-order differences capture abrupt changes, and trend features describe six-step degradation development.

Rolling mean features were computed using a six-sample moving window. Considering the applied 10-min sampling interval, these variables represent approximately one-hour operating trends. First-order difference features capture short-term dynamic changes, whereas trend variables characterize longer-term degradation patterns. Six-step trend features were introduced to capture longer-term degradation behaviour by comparing the current sensor value with the value observed six time steps earlier, corresponding to approximately one hour under the applied sampling interval. Ratio-based indicators were introduced to capture interactions between mechanical, thermal, and electrical operating conditions.

Feature selection was subsequently performed using a combination of:

- statistical methods (correlation analysis, ANOVA),
- embedded methods (feature importance obtained from Random Forest and XGBoost),
- explainable artificial intelligence techniques (SHAP values).

The final feature set enabled the predictive models to capture not only current equipment conditions but also degradation dynamics and process evolution patterns preceding failure occurrence.

3.4. Model Training and Algorithm Comparison

Two types of models were used to solve the prediction task:

- Random Forest (RF) as a robust baseline model,
- XGBoost as the main model with high accuracy and interpretability.

The selection of Random Forest and XGBoost was motivated by the characteristics and scope of the available dataset. The study is based on heterogeneous tabular operational data and a limited number of recorded failure events. Under these conditions, tree-based ensemble methods provide a suitable balance between predictive capability, robustness to heterogeneous features, and model interpretability. Deep learning architectures such as LSTM, GRU, Temporal CNNs, and Transformers are particularly suitable for large-scale sequential datasets and representation learning from extensive temporal data. However, given the limited observation period and the small number of failure events in the present study, their inclusion could increase the risk of overfitting and would not allow a sufficiently robust comparison. Similarly, unsupervised anomaly detection methods address a different problem formulation from the supervised failure prediction task considered in this study. Therefore, Random Forest was selected as a robust baseline model and XGBoost as the primary predictive and interpretive model. Comparison with deep learning architectures and anomaly detection approaches on larger multi-device datasets remains an important direction for future research.

The models were trained on the same dataset with identical splits into training, validation, and test sets. Due to class imbalance between failure and non-failure states, class weighting was applied during model training. For the Random Forest model, a higher weight was assigned to the minority failure class, while for the XGBoost model, the scale_pos_weight parameter was calculated as the ratio of non-failure to failure samples in the training set. This approach increased the penalty associated with misclassifying failure events and reduced model bias toward the majority non-failure class. Models were tested at two different prediction horizons (8 h and 16 h), enabling analysis of the trade-off between model accuracy and lead time before failure. The Random Forest model was configured with 100 estimators, the Gini impurity criterion, unrestricted maximum tree depth, and a fixed random seed to ensure reproducibility. The XGBoost model was configured with 100 estimators, a learning rate of 0.1, a maximum tree depth of 6, a subsampling rate of 0.8, and a column sampling rate of 0.8. Binary logistic loss was used as the learning objective. Both Random Forest and XGBoost showed low computational requirements for the analyzed dataset and could be trained and evaluated on standard computing hardware without specialized GPU acceleration. Due to the limited dataset size and the use of tree-based ensemble models, model training was completed within a short execution time and did not represent a computational bottleneck in the proposed workflow. XGBoost required moderately higher computational effort than Random Forest due to its sequential boosting procedure; however, the computational requirements remained suitable for practical offline model retraining. During deployment, prediction inference requires substantially less computation than model training, supporting the potential use of the proposed framework for periodic or near-real-time maintenance decision support. Exact computational time may vary depending on hardware configuration, dataset size, and model implementation. Model performance was evaluated using the following metrics:

- Accuracy,
- Precision,
- Recall,
- F1-score,
- ROC-AUC.

XGBoost was used as the primary model for generating probabilistic output:

$$P_t = P(\textit{failure in interval } [t, t + \Delta t]) \tag{2}$$

where P_t is the probability of failure at time t.

To ensure objective evaluation, two validation approaches were used. The first was stratified split, which served for technical comparison of models and ensured representative representation of failure and non-failure states in both training and test sets. The second was temporal split, in which the model was trained on historical data and tested on chronologically subsequent data, simulating realistic industrial deployment of the model.

3.5. Replay Analysis and Identification of Optimal Intervention Time

To verify the practical applicability of the model, a so-called "replay analysis" was implemented, which simulates maintenance decision-making based on historical data. The replay analysis was realized by chronological replay of historical data according to the following pseudoalgorithm. For each time instant, the alarm state was defined based on a probability threshold:

$$Alert_t = 1, \; if \; P_t \geq \theta; 0, \; otherwise \tag{3}$$

where P_t is the probability of failure at time t, and θ represents the selected decision threshold (0.60 in the present study).

Algorithm 1 Replay Evaluation Algorithm:

Algorithm 1. Replay-Based Evaluation Procedure

```
Input: Historical dataset (H), alert threshold (\theta)
Output: Precision, Recall, False Alarm Rate, Average Lead Time
For each time step (t \in H) do
  Predict failure probability (P_{t})
  If (P_{t} \geq \theta) then
   Generate maintenance alert
   If a failure occurs after the alert then
     LeadTime ← FailureTime − AlertTime
     Store LeadTime
   End If
  End If
End For
Compute Precision
Compute Recall
Compute False Alarm Rate
Compute Average Lead Time
Return evaluation metrics
```

Based on this, the following were evaluated:

- lead time—time between the first alert and the failure,
- detection rate—proportion of detected failures,
- false alarm rate—number of false alarms.

The replay analysis enabled identification of the optimal maintenance intervention time and comparison with actual historical interventions.

3.6. Model Interpretation (XAI)

Explainable artificial intelligence (XAI) methods were used to interpret model decision-making, specifically SHAP and LIME. SHAP (SHapley Additive exPlanations) is based on principles from game theory and represents one of the most widely used methods for interpreting machine learning models. LIME (Local Interpretable Model-agnostic Explanations) explains the model's decision locally, i.e., in the vicinity of a specific point, using a simple approximation.

For the purposes of this article, SHAP was preferred as it provides consistent global and local explanations and is suitable for tree-based models such as Random Forest and XGBoost.

$$f(x) = \varphi_0 + \Sigma^I \varphi_i \quad (4)$$

where

- $f(x)$ is the resulting model prediction,
- φ_0 is the base value of the model,
- φ_i is the contribution of the *i*-th variable to the resulting prediction.

A positive SHAP value means that the corresponding variable increases the risk of failure, while a negative value means that it reduces the risk. Model interpretability enables transformation of AI outputs into a form understandable to maintenance personnel.

Analysis revealed the dominant factors:

- flow rate,
- pressure,
- vibrations_rollmean,
- motor current.

3.7. Uncertainty Analysis—Results

Model uncertainty was analyzed at three levels:

- data uncertainty (noise, missing values),
- model uncertainty (overfitting),
- operational uncertainty (mode changes).

For its quantification, the following were used:

- confidence intervals,
- variance of ensemble models,
- probabilistic model outputs.

Uncertainty was estimated from the variability of predictions generated by an ensemble of independently trained models using different random seeds and training subsets. The resulting prediction variance was used to derive uncertainty intervals for the predicted Dynamic FMEA Risk Index. In cases where the uncertainty exceeded a predefined threshold, the system was designed to generate a warning for the operator. This section represents a methodological framework that will be extended and validated using real enterprise data in the next phase of research.

3.8. Integration with FMEA and Risk Index

Model outputs were transformed into a risk index compatible with the FMEA methodology. The failure probability generated by the AI model was combined with the values of failure severity (Severity) and detection capability (Detection), creating a dynamic risk indicator enabling continuous risk assessment over time.

This approach enables:

- integration of AI models with existing maintenance processes,
- prioritization of maintenance interventions,
- decision support at the maintenance level,
- dynamic risk update according to the current technical condition of the equipment.

$$RI_t = P_{AI,t} \times S \times D \quad (5)$$

where:

- RI_t—dynamic risk index,
- $P_{AI,t}$—AI-predicted failure probability at time t,
- S—Severity,
- D—Detection.

Severity and Detection were taken from the enterprise FMEA database, while the failure probability was dynamically updated by the AI model based on current operational data.

The proposed approach extends the traditional static FMEA concept with time-dependent risk assessment based on the current technical condition of the equipment. The introduced dynamic risk indicator does not replace the classical RPN assessment but extends the existing FMEA framework through AI-supported updating of the failure occurrence probability over time.

The proposed solution simultaneously develops previous AI-supported FMEA concepts presented by Bubenik et al. by introducing time-dependent AI risk updating and a maintenance-oriented decision support mechanism. The result is a dynamic approach enabling more effective prioritization of maintenance interventions, better maintenance planning, and decision support in the Industry 4.0 environment.

3.9. Sustainability Impact Assessment

Based on the replay analysis, the potential impact of early intervention was assessed for:

- reduction in unplanned downtime,
- reduction in material waste,
- energy savings.

Comparison of the actual state and alternative what-if scenarios enabled scenario-based estimation of the potential effects of AI-supported maintenance. The analysis was not designed as experimental validation of realized sustainability benefits; instead, it evaluates how different assumed levels of maintenance effectiveness could influence downtime, waste, and energy consumption.

Downtime:

$$Downtime_{AI} = Downtime_{real} \times (1 - r)$$

where r = assumed downtime reduction ratio.

Waste:

$$Waste = Downtime \times k$$

where k = 0.5 pcs/min.

Energy:

$$Energy_{AI} = Energy_{real} \times (1 - e)$$

where e represents the assumed energy reduction ratio used in the scenario analysis. In the illustrative scenario, e was set to 0.08. This value is not an empirically validated energy-saving rate and is used solely as a scenario parameter to demonstrate the calculation procedure and assess the potential direction and magnitude of sustainability effects. Consequently, the resulting energy savings should be interpreted as scenario-based estimates rather than experimentally verified outcomes.

4. Results

This section presents the results obtained from the proposed AI-based predictive maintenance framework. The evaluation focuses on predictive model performance, threshold sensitivity analysis, feature importance interpretation, and replay-based validation. The results are intended to assess both the predictive capability of the developed models and their practical applicability in maintenance decision-making.

4.1. Performance of Predictive Models

Model performance was analyzed at various levels of decision threshold (ALERT_THRESHOLD), with the aim of identifying the trade-off between the ability to detect failure (Recall), alert precision (Precision), and lead time before failure (Lead Time).

4.1.1. Threshold Analysis (Sensitivity Analysis)

The decision threshold was determined empirically through sensitivity analysis on the validation data. Candidate threshold values were selected to represent progressively more conservative alerting strategies and to evaluate the trade-off between failure detection capability and alert precision. The final threshold was selected by jointly considering Precision, Recall, and F1-score rather than maximizing a single classification metric. In the analyzed case, a threshold of 0.60 provided the most balanced operating point between failure detection and false-alert control and was therefore selected for the subsequent model comparison.

System sensitivity was analyzed at three levels of decision threshold (0.55; 0.60; 0.70). The results demonstrated that a lower threshold increases the ability to detect failures (Recall), but at the cost of a higher number of alerts. The value of 0.60 represented the best compromise between Recall (0.473), Precision (0.803), F1-score (0.596), and an average lead time before failure of 14.23 h. The results of the threshold sensitivity analysis are summarized in Table 2.

Table 2. ALERT_THRESHOLD sensitivity analysis.

ALERT_THRESHOLD	Precision	Recall	F1-Score
0.70	0.878	0.384	0.534
0.60	0.803	0.473	0.596
0.55	0.678	0.527	0.593

Results showed that a lower threshold increased Recall at the cost of a higher number of alerts, while a higher threshold increased Precision. The value of 0.60 represented the best compromise between detection accuracy and practical system usability.

4.1.2. Model Comparison at Selected Threshold

Based on the threshold analysis, a threshold of 0.60 was used for further experiments. A comparison of Random Forest and XGBoost models was then performed. Tables 3 and 4 are evaluated at the threshold selected from the threshold analysis.

Table 3. Random Forest—Metric Comparison.

Accuracy	Precision	Recall	F1	ROC-AUC
0.89	1.000	0.375	0.545	0.812

Table 4. XGBoost—Metric Comparison.

Accuracy	Precision	Recall	F1	ROC-AUC
0.89	0.803	0.473	0.596	0.775

Results showed slightly different behavior between the models. Random Forest achieved higher Precision and ROC-AUC, while XGBoost provided a better compromise between Recall, F1-score, and interpretability using SHAP analysis.

4.2. Replay Analysis and Maintenance Response Time

To verify the practical applicability of the model, a replay analysis of historical failures was performed. The XGBoost model was applied to historical data and generated the failure probability over time.

Based on the defined threshold $\theta = 0.60$, the first critical alert before failure was identified, representing the optimal maintenance intervention time. The key results of the replay analysis for both prediction horizons are summarized in Table 5.

Table 5. Replay analysis results.

Parameter	Horizon 8 h	Horizon 16 h
Prediction horizon	8 h	16 h
Average lead time	~7.8 h	~15.26 h
Maximum lead time	~8 h	~24 h
Alert character	Precise, short-term	Earlier, planning-oriented
Practical use	Operational response	Maintenance planning
False alarm risk	Lower	Higher

The replay analysis demonstrated a trade-off between classification accuracy and available lead time before failure. The 8-h horizon provided more precise alerts, while the 16-h horizon provided more time for maintenance planning.

As shown in Figure 2, the model generates a time course of failure probability that progressively increases before the actual failure. After exceeding the threshold value, the first alert can be identified, representing the available lead time for performing a maintenance intervention. This result confirms the practical applicability of replay analysis in retrospective evaluation of maintenance decision-making.

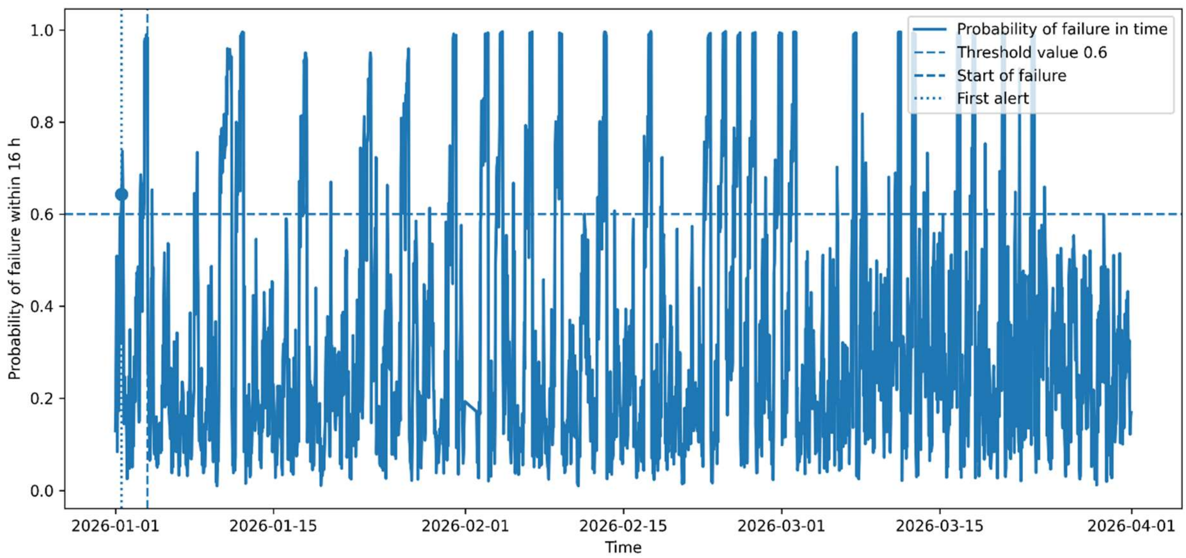

Figure 2. Evolution of failure probability over time with indication of threshold value, first alert, and moment of failure.

4.3. Probabilistic Failure Profile over Time

The XGBoost model generated a continuous time course of failure probability. A significant increase in probability was observed before failures, enabling identification of critical periods. A sample of the generated failure probability profile is presented in Table 6. The results illustrate how the model translates predicted failure probabilities into risk categories and corresponding maintenance recommendations. As the predicted failure probability increases, the assigned risk level escalates from low to critical, triggering progressively more urgent maintenance actions. The probability of failure increased progressively prior to the actual failure event, confirming the model's ability to capture the degradation process of the system.

Table 6. Sample model output.

Time	Failure Probability	Risk	Recommendation
06:00	0.18	Low	Normal operation
07:00	0.41	Medium	Monitoring
08:00	0.81	Critical	Immediate intervention

Such an output is suitable for integration into maintenance management systems (e.g., CMMS), where it can be presented as a simple traffic-light indicator system (Figure 3) for the visualization of monitored parameters and maintenance risk levels.

Figure 3. Example of a predictive maintenance dashboard based on real operating data.

4.4. Model Interpretation (XAI Analysis)

SHAP values were used to interpret the model, identifying the most significant factors influencing failure prediction. Table 7 summarizes the variables with the highest average contribution to the model output. The results indicate that flow rate, vibration trends, pressure, and motor current were the most influential factors affecting failure probability. This finding suggests that both process-related variables and condition-monitoring indicators play a significant role in predicting equipment degradation and impending failures.

Table 7. Most significant factors influencing failure prediction.

Variable	Mean Absolute SHAP Contribution
flow_rate	0.475
vibrations_rollmean	0.365
pressure	0.318
motor_current	0.253

To complement the quantitative ranking of feature importance presented in Table 7, a detailed SHAP summary plot was generated. The visualization provides insight not only into the relative importance of individual variables but also into the direction and magnitude of their influence on the model output. Positive SHAP values indicate an increased probability of failure, whereas negative values decrease the predicted risk.

Figure 4 presents the SHAP summary plot for the XGBoost model. Each point represents a single observation, with color indicating the feature value (blue = low, red = high). Features are ranked according to their average impact on model predictions, allowing identification of the variables that most strongly contribute to failure detection.

Figure 4. SHAP summary plot showing the impact of input variables on failure prediction generated by the XGBoost model.

4.5. Uncertainty Analysis

Model uncertainty was estimated using an ensemble approach based on repeated generation of prediction outputs from Random Forest and XGBoost models. The uncertainty interval was derived from the 5th and 95th percentiles of the distribution of probabilistic model outputs. Narrower intervals indicate higher model stability and greater confidence in the decision-making process.

$$CI_{90} = [P_5, P_{95}] \tag{6}$$

where

- P_5 represents the lower 5th percentile,
- P_{95} represents the upper 95th percentile.

The results suggest that greater stability of model outputs is associated with reduced uncertainty, which may increase confidence in maintenance decision-making. Conversely, a wider uncertainty interval may indicate the need for additional diagnostics or enhanced equipment monitoring.

The lower and upper bounds correspond to the 5th and 95th percentiles of the probability distribution generated by the ensemble models. The width of the interval represents the degree of model uncertainty.

The final failure probability corresponds to the output of the primary decision-making model, while the uncertainty interval characterizes the variability of predictions across the ensemble models.

To illustrate the practical interpretation of the proposed uncertainty framework, Table 8 presents representative examples of model predictions together with their corresponding uncertainty intervals. The examples demonstrate how the predicted failure probability can be complemented by confidence bounds, providing additional information for maintenance decision-making.

Table 8. Sample model uncertainty analysis.

Time	Probability	P5	P95	Interpretation
1 January 2026 02:00	0.426	0.32	0.42	Low
1 January 2026 03:00	0.509	0.39	0.51	Medium

Furthermore, an overall summary of the uncertainty analysis is provided in Table 9. The aggregated results characterize the distribution of uncertainty levels across all analyzed samples and provide insight into the overall reliability and stability of the predictive maintenance framework.

Table 9. Uncertainty analysis summary.

Parameter	Value
Number of analyzed samples	624
Average failure probability	0.53
Proportion of low uncertainty	51.9%
Proportion of medium uncertainty	47.6%
Proportion of high uncertainty	0.5%
Confidence interval used	90%

The majority of predictions (99.5%) fell into the categories of low or medium uncertainty, with only 0.5% of samples exhibiting high uncertainty. The results indicate sufficient model stability for practical use in predictive maintenance conditions.

4.6. Impact Assessment on Maintenance and Sustainability

Based on the replay analysis, the difference between the actual course of failure events and a simulated scenario utilizing early AI-prediction-supported maintenance intervention was evaluated. The assessment was performed as a what-if simulation aimed at estimating the potential impact of implementing predictive maintenance on downtime, material waste, and energy consumption. The scenario analysis was based on assumed parameters rather than experimentally observed post-implementation effects. For downtime reduction, three scenario levels were evaluated: 50%, 75%, and 90%. These values represent conservative, intermediate, and optimistic analytical scenarios, respectively, and were introduced to examine the sensitivity of the estimated sustainability effects to different assumed levels of predictive maintenance effectiveness. The 75% scenario was selected as an illustrative intermediate case for the subsequent comparison and should not be interpreted as an empirically validated expected reduction. Similarly, the assumed 8% reduction in energy consumption represents an illustrative scenario parameter associated with the potential elimination of degraded operating conditions and was not directly measured in an experimental deployment. Table 10 presents the sensitivity analysis of the simulated scenarios. The results demonstrate the dependence of the estimated downtime reduction on the assumed effectiveness of predictive maintenance intervention. The 50% scenario produces a more conservative estimate, whereas the 90% scenario represents an optimistic analytical case. The intermediate 75% scenario is used in the subsequent analysis solely to illustrate the calculation procedure and potential sustainability effects under a defined set of assumptions.

Table 10. Sensitivity analysis of simulated scenario.

Assumed Downtime Reduction	Total Downtime (min)
50%	1980
75%	991
90%	396

The overall benefits of the proposed AI-based predictive maintenance framework are summarized in Table 11 through a comparison of the actual and simulated operating scenarios.

Table 11. Comparison of actual and simulated scenarios.

Parameter	Actual State	Illustrative AI Scenario	Estimated Difference
Downtime (min)	3963	990.8	−2972.2
Waste (pcs)	1981.5	495.4	−1486.1
Energy (kWh)	25.6	23.6	−2.0

The simulation results show the potential benefit of implementing predictive maintenance from the perspective of operational efficiency and sustainability. The most significant effect was observed in the reduction in unplanned downtime, which subsequently also affected the estimated waste production. Energy savings were smaller, as energy consumption represented a parameter only partially influenced by degradation.

4.7. Integration into Maintenance Decision-Making (FMEA)

Model outputs were transformed into a dynamic risk index compatible with the logic of the FMEA methodology. The failure probability generated by the model was combined with failure severity and detection level, creating an indicator enabling continuous risk assessment over time. In contrast to the static RPN, the proposed approach enables risk updating according to the current technical condition of the equipment.

This approach enables:

- prioritization of maintenance interventions,
- more effective maintenance planning,
- earlier identification of critical states,
- maintenance decision support based on a combination of historical data, current signals, and FMEA logic.

5. Discussion

The obtained results demonstrate the potential of artificial intelligence techniques to support maintenance decision-making in manufacturing systems. Beyond achieving satisfactory predictive performance, the proposed framework integrates prediction, interpretability, uncertainty assessment, and risk evaluation into a unified approach aimed at improving equipment reliability and operational sustainability.

5.1. Interpretation of Main Results

The results of predictive modeling demonstrate the potential of Random Forest and XGBoost algorithms for supporting the identification of failure states with useful advance warning. However, the achieved Recall values indicate an important limitation of the current models. At the selected threshold of 0.60, Random Forest achieved a Recall of 0.375 and XGBoost a Recall of 0.473, meaning that a substantial proportion of actual failure states remained undetected. Therefore, the obtained predictive performance should be interpreted as promising for decision-support purposes but not sufficient for fully autonomous or safety-critical maintenance applications. In the conducted experiment, Random Forest achieved higher Precision and ROC-AUC, whereas XGBoost provided a better balance between Recall, F1-score, and interpretability through SHAP analysis.

The proposed approach represents a practically implementable maintenance decision-support framework that combines failure prediction, model interpretability, replay analysis, and dynamic risk assessment with the objective of increasing the reliability and sustainability of manufacturing systems.

Such advance warning provides sufficient time for maintenance planning, spare-part procurement, and the execution of preventive interventions before a critical failure state occurs.

Random Forest achieved higher Precision and ROC-AUC values, while XGBoost provided a better balance between Recall, F1-score, and interpretability through SHAP analysis. For this reason, XGBoost was selected as the primary interpretive model, whereas Random Forest served as a robust reference model.

From a practical perspective, a key advantage of the proposed framework is that the model does not merely provide a binary classification of equipment condition but generates a continuous failure probability profile over time. This profile enables the identification of critical periods preceding a failure and supports the determination of the optimal maintenance intervention time.

Replay analysis demonstrated that, for a prediction horizon of 16 h, the system achieved an average lead time of approximately 15.26 h, while in several cases the advance warning exceeded 20 h.

The replay analysis further confirmed that the selection of an appropriate prediction horizon is a crucial factor for the practical implementation of a predictive maintenance system.

The combination of Random Forest, XGBoost, and SHAP represents a balanced approach that integrates robustness, predictive accuracy, and interpretability. Random Forest serves as a stable reference model, XGBoost acts as the primary interpretive model for tabular operational data, and SHAP provides explanations of model outputs through the quantified contributions of individual variables. Such an approach increases the trustworthiness of artificial intelligence outputs and facilitates their adoption in maintenance decision-making.

A significant finding was the difference between results obtained using stratified and temporal data splitting. While the stratified approach produced higher classification metrics, the temporal split more accurately represented real deployment conditions in an industrial environment. The results also highlight the importance of sufficiently long data collection periods and a representative number of failure events for the reliable evaluation of predictive models under practical operating conditions.

5.2. Influence of Prediction Horizon

Results showed that the setting of the prediction horizon represents a key factor for practical implementation of a predictive maintenance system. A shorter horizon provided higher classification accuracy, while a longer horizon created more room for planning interventions.

Results therefore point to a trade-off between:

- precision,
- recall,
- lead time,
- practical system usability.

The selection of the decision threshold significantly influences the practical use of the predictive maintenance system. A higher threshold reduces the number of false alarms but also reduces the model's ability to detect emerging failures. Industrial implementation therefore requires threshold optimization according to operational requirements, availability of maintenance personnel, and economic impacts of unplanned shutdowns.

The relatively low Recall values observed in the present study also have important practical implications. A missed failure may result in unplanned downtime, secondary equipment damage, production losses, or increased maintenance costs, depending on the severity of the failure mode. Therefore, threshold selection should not be based solely on overall classification performance but should also reflect the operational and economic consequences of false-negative predictions. For critical failure modes, a lower decision threshold or a more conservative maintenance policy may be preferable, even at the cost of an increased number of false alarms. Consequently, the proposed framework should currently be considered a decision-support tool that complements, rather than replaces, conventional condition monitoring, maintenance expertise, and safety procedures.

5.3. Contribution Compared to Existing Work

The individual methodological components employed in this study, including machine learning-based failure prediction, SHAP interpretation, replay analysis, FMEA, and sustainability assessment, have been previously investigated in the literature. The contribution of the present study is therefore not attributed to these components individually. Rather, the proposed methodological contribution consists of their operational connection through a time-dependent decision mechanism in which AI-generated failure probability dynamically updates an FMEA-compatible risk indicator, while SHAP explanations identify the dominant contributing factors and replay analysis determines the available intervention time. This creates a direct methodological link between prediction, interpretation, risk prioritization, and maintenance action planning. The main contributions can be summarized as follows:

- implementation of replay analysis on historical data enabling retrospective identification of optimal intervention time,
- integration of predictive models with the FMEA methodology through a dynamic risk index,
- use of SHAP for identification of dominant factors influencing failure prediction,
- quantification of the impact on manufacturing sustainability (downtime, waste, energy).

This approach represents a shift from "prediction as a goal" to "prediction as a decision-making tool".

5.4. Significance of Model Interpretation for Maintenance

Model interpretation using SHAP analysis provides significant added value, as it enables linking artificial intelligence results with the physical behavior of the equipment.

Identification of dominant variables, such as flow rate, pressure, vibrations, and motor electrical load, enables:

- better understanding of failure mechanisms,
- design of specific preventive measures,
- increased confidence of maintenance personnel in the AI system.

Model interpretability is therefore a key factor in its implementation in an industrial environment.

5.5. Role of Uncertainty in Decision-Making

An important aspect of the proposed approach is the consideration of prediction uncertainty. Results showed that high failure probability is often accompanied by lower uncertainty, which increases the reliability of decision-making. Conversely, in cases of higher uncertainty, a conservative approach should be applied, such as increased monitoring or additional diagnostics. Consideration of uncertainty represents a significant step toward safe implementation of AI in critical industrial applications.

5.6. Impact on Manufacturing System Sustainability

The AI scenario represents a simulation estimate based on replay analysis and assumptions about downtime reduction; the results cannot be interpreted as real experimental verification. Therefore, the reported reductions in downtime, waste, and energy consumption should be interpreted as conditional scenario outcomes rather than demonstrated operational improvements. Their magnitude depends directly on the assumed effectiveness of early maintenance intervention and on the relationship between degraded operation, downtime, waste generation, and energy consumption. In particular, the energy-saving estimate requires future validation through direct measurement of equipment energy consumption before and after the implementation of AI-supported maintenance. The present analysis therefore demonstrates the assessment procedure and potential sustainability pathways rather than verified sustainability performance.

Early identification of failure enables:

- reduction in unplanned downtime,
- reduction in material waste,
- optimization of energy consumption.

These contributions have a direct impact on the environmental and economic performance of the manufacturing system. The results support the concept of sustainable maintenance as part of modern manufacturing systems [26].

5.7. Limitations and Future Research

Despite the achieved results, the proposed approach has certain limitations:

- dependence on the quality and availability of historical data,
- limited number of failure events,
- need for model adaptation when operating conditions change,
- need for verification of the proposed framework in various industrial sectors,
- simplified modeling of energy and waste impacts.

Future research should focus on:

- extending datasets to more devices and longer time periods,
- comparative evaluation of tree-based models, deep learning architectures (LSTM, GRU, Temporal CNNs, and Transformers), and anomaly detection methods on larger multi-device datasets,
- integration of digital twin models,
- advanced uncertainty modeling,
- automation of integration with CMMS systems,
- more detailed quantification of environmental impacts (e.g., CO_2).

The proposed framework is designed as a modular decision-support architecture and can be adapted to other manufacturing systems provided that relevant operational, condition-monitoring, and maintenance data are available. Application to another machine or production system would require adaptation of the input variables and feature engineering procedures to the specific equipment, retraining and validation of the predictive model using system-specific historical data, and adjustment of decision thresholds according to the operational and economic consequences of failures. The FMEA-related Severity and Detection parameters and the maintenance decision rules must likewise be adapted to the failure modes and maintenance strategy of the target system. The general workflow connecting data acquisition, failure prediction, explainability, replay-based intervention assessment, dynamic risk evaluation, and maintenance decision support remains unchanged. This modular structure enables potential application to different manufacturing equipment and production environments, although transferability must be verified through multi-device and cross-industrial validation.

The dataset used in this study was collected from a single manufacturing device over a three-month period and contained 27 recorded failure events. This limited sample size represents an important constraint on the robustness and generalizability of the reported results. In particular, the relatively small number of failure events may lead to variability in performance metrics, sensitivity to individual failure cases, and limited representation of the full range of degradation patterns and operating conditions that may occur during long-term industrial operation. Consequently, the reported comparison between Random Forest and XGBoost should be interpreted as an evaluation within the analyzed case study rather than as evidence of the general superiority of either algorithm across different machines or industrial environments. Furthermore, because the data originate from a single device, the models may capture equipment-specific relationships that may not transfer directly to other machine types, production regimes, or operating conditions. Validation using longer observation periods, larger numbers of failure events, and data from multiple devices and industrial environments is therefore required before conclusions regarding broader model robustness and generalizability can be drawn.

6. Practical Implementation

The proposed predictive maintenance system can be implemented in a real manufacturing environment as a decision support tool supporting real-time maintenance. The basis of implementation is continuous data collection from sensors such as vibrations, temperature, and electric current. These data are processed using the XGBoost model, which generates the failure probability within a defined time horizon.

Based on the model output, a threshold mechanism is defined:

- low risk ($p < 0.4$)—normal operation,
- medium risk ($0.4 \leq p < 0.7$)—increased monitoring,
- high risk ($p \geq 0.7$)—planned or immediate intervention.

The system can be integrated into existing maintenance management systems (CMMS or MES), where it can automatically generate notifications and recommendations. The deployment architecture assumes edge data collection and periodic model retraining to account for changes in operating conditions and gradual data drift.

6.1. SHAP Integration

Model interpretation using SHAP enables identification of dominant factors contributing to increased failure risk. The decision layer of the system uses maintenance rules (MAINTENANCE_RULES) that map identified dominant factors and the dynamic FMEA risk index to recommended maintenance interventions.

Maintenance rules can be defined by an enterprise MAINTENANCE_RULES database, which transforms SHAP interpretation results into specific recommended maintenance actions, e.g.,

- if vibrations > X → bearing inspection,
- if temperature > Y → lubrication,
- if current > Z → motor inspection.

This approach links model interpretability, AI-supported risk assessment, and practical maintenance decision-making into a unified decision mechanism supporting Industry 4.0 principles.

6.2. Practical Scenario

In the case where the model identifies high failure risk caused by elevated vibrations and bearing temperature, the system recommends checking lubrication or replacing bearings before failure occurs.

A significant contribution consists of visualizations indicating when to intervene (Figure 2), what level of risk exists (Figure 3), and which factors dominate the model's decision (Figure 4).

6.3. Sustainability Integration

System implementation has the potential to contribute to reduced unplanned downtime, lower material waste, and improved energy efficiency, as illustrated in the flow diagram in Figure 5. However, the magnitude of these effects depends on the effectiveness of maintenance interventions and specific operating conditions and requires validation through prospective industrial deployment.

The MAINTENANCE_RULES table implemented in the system's decision-making layer transforms the outputs of SHAP analysis and the dynamic FMEA risk index into specific recommended maintenance actions.

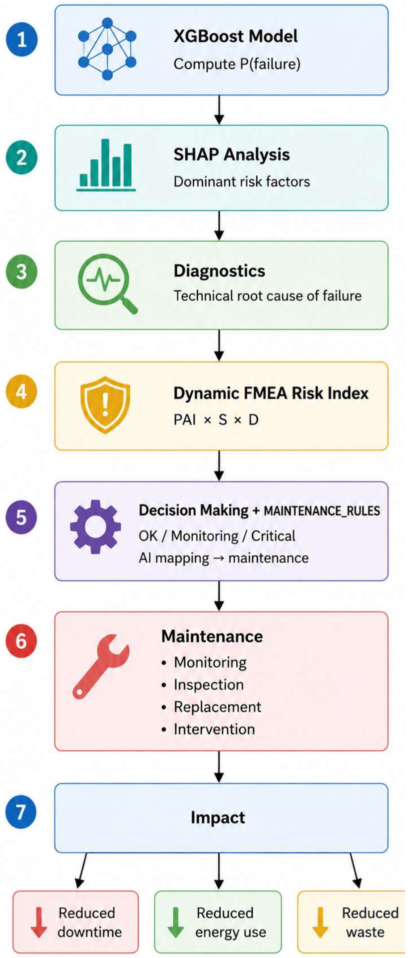

Figure 5. Flow diagram from model to maintenance decision support.

7. Conclusions

This article addressed predictive modeling of failure states in manufacturing systems using artificial intelligence in the context of sustainability. The proposed methodological framework was based on processing historical operational and maintenance data and included feature engineering, training of multiple machine learning models, their validation, interpretation, and practical verification using replay analysis.

Results showed that XGBoost and Random Forest models demonstrated potential for supporting failure-state classification, although the achieved Recall values indicate that a substantial proportion of failures remained undetected. Therefore, the current models should be regarded as components of a maintenance decision-support framework rather than as autonomous failure prevention systems. Among the evaluated approaches, XGBoost provided a suitable balance between Recall, F1-score, and interpretability in combination with explainable artificial intelligence (SHAP). The proposed approach enables generation of a time course of failure probability and supports identification of an appropriate maintenance intervention time before the occurrence of a critical operational state.

Replay analysis demonstrated the retrospective applicability of the methodological framework for simulating maintenance decision-making and enabled scenario-based estimation of the potential benefits of early intervention. Simulated scenarios suggested the possibility of reducing unplanned downtime, material waste, and energy requirements of manufacturing systems.

A significant aspect of the work is the integration of artificial intelligence outputs with the FMEA methodology through a dynamic risk index based on AI-updated failure probability. The proposed approach extends the traditional static FMEA concept with time-dependent risk assessment based on the current technical condition of the equipment. Simultaneously, the use of explainable artificial intelligence increases model transparency and enables identification of dominant factors influencing failure prediction.

The contribution of the article lies in the integration of predictive maintenance, replay analysis, explainable artificial intelligence, dynamic FMEA risk index, and sustainability impact assessment into a single decision-making framework. The proposed methodological framework represents a practically implementable maintenance decision support solution for modern manufacturing systems and the Industry 4.0 environment.

Future research will focus on extending the dataset through longer observation periods, a larger number of recorded failure events, and data collection from multiple devices and operating conditions. Such validation is necessary to assess model robustness, transferability, and generalizability across different manufacturing environments. Future work will also focus on online learning approaches enabling continuous adaptation of predictive models to changing manufacturing conditions and operating modes. Cross-industrial validation and deployment in real-time industrial environments remain important directions for future research.

Author Contributions: Conceptualization, M.R. and P.B.; methodology, M.R.; software, M.R. and M.B.; validation, M.R., P.B. and V.B.; formal analysis, M.R.; investigation, M.R.; data curation, M.R.; writing—original draft preparation, M.R.; writing—review and editing, P.B., V.B. and M.B.; visualization, M.R.; supervision, P.B.; project administration, P.B.; funding acquisition, P.B. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Agency for the Promotion of Research and Development under contract no. APVV-21-0308. This research was funded by KEGA 011ŽU-4/2025, VEGA 1/0633/24, and VEGA 1/0189/26.

Data Availability Statement: Data are available from the corresponding author upon reasonable request.

Conflicts of Interest: The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

Abbreviations

The following abbreviations are used in this manuscript:

AI	Artificial Intelligence
CMMS	Computerized Maintenance Management System
FMEA	Failure Mode and Effects Analysis
MES	Manufacturing Execution System
OEE	Overall Equipment Effectiveness
RF	Random Forest
RMS	Root Mean Square
RPN	Risk Priority Number
SHAP	SHapley Additive exPlanations
XAI	Explainable Artificial Intelligence
XGBoost	Extreme Gradient Boosting

References

1. Jardine, A.K.S.; Lin, D.; Banjevic, D. A Review on Machinery Diagnostics and Prognostics Implementing Condition-Based Maintenance. *Mech. Syst. Signal Process.* **2006**, *20*, 1483–1510. [CrossRef]
2. Lee, J.; Bagheri, B.; Kao, H.-A. A Cyber-Physical Systems Architecture for Industry 4.0-Based Manufacturing Systems. *Manuf. Lett.* **2015**, *3*, 18–23. [CrossRef]
3. Carvalho, T.P.; Soares, F.A.A.M.N.; Vita, R.; Francisco, R.P.; Basto, J.P.; Alcálá, S.G. A Systematic Literature Review of Machine Learning Methods Applied to Predictive Maintenance. *Comput. Ind. Eng.* **2019**, *137*, 106024. [CrossRef]
4. Chen, T.; Guestrin, C. XGBoost: A Scalable Tree Boosting System. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, 13–17 August 2016*; Association for Computing Machinery: New York, NY, USA, 2016; pp. 785–794.
5. Breiman, L. Random Forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]
6. Bubeník, P.; Rakyta, M.; Buzalka, M.; Biňasová, V.; Kováříková, Z. Optimization of Business Processes Using Artificial Intelligence. *Electronics* **2025**, *14*, 2105. [CrossRef]
7. Zhang, W.; Yang, D.; Wang, H. Data-driven methods for predictive maintenance of industrial equipment: A survey. *IEEE Syst. J.* **2019**, *13*, 2213–2227. [CrossRef]
8. Lei, Y.; Li, N.; Guo, L.; Li, N.; Yan, T.; Lin, J. Machinery Health Prognostics: A Systematic Review from Data Acquisition to RUL Prediction. *Mech. Syst. Signal Process.* **2018**, *104*, 799–834. [CrossRef]
9. Arez, D.; Navas, H.V.G.; Gaspar, P. The Bridge Between Artificial Intelligence and Predictive Maintenance in Industry 4.0: A Systematic Review. *Appl. Sci.* **2026**, *16*, 4882. [CrossRef]
10. Lundberg, S.M.; Lee, S.-I. A Unified Approach to Interpreting Model Predictions. *arXiv* **2017**. [CrossRef]
11. Susto, G.A.; Schirru, A.; Pampuri, S.; McLoone, S.; Beghi, A. Machine Learning for Predictive Maintenance: A Multiple Classifier Approach. *IEEE Trans. Ind. Inform.* **2015**, *11*, 812–820. [CrossRef]
12. Bousdekis, A.; Lepenioti, K.; Apostolou, D.; Mentzas, G. A review of data-driven decision-making methods for industry 4.0 maintenance applications. *Electronics* **2021**, *10*, 828. [CrossRef]
13. Ahmad, R.; Kamaruddin, S. An Overview of Time-Based and Condition-Based Maintenance in Industrial Application. *Comput. Ind. Eng.* **2012**, *63*, 135–149. [CrossRef]
14. Kang, H.S.; Lee, J.Y.; Choi, S.; Kim, H.; Park, J.H.; Son, J.Y.; Kim, B.H.; Noh, S.D. Smart Manufacturing: Past Research, Present Findings and Future Directions. *Int. J. Precis. Eng. Manuf.-Green Technol.* **2016**, *3*, 111–128. [CrossRef]
15. Lu, Y. Industry 4.0: A Survey on Technologies, Applications and Open Research Issues. *J. Ind. Inf. Integr.* **2017**, *6*, 1–10. [CrossRef]
16. Zhang, Y.; Ren, S.; Liu, Y.; Si, S. A Big Data Analytics Architecture for Cleaner Manufacturing and Maintenance Processes. *Procedia CIRP* **2017**, *61*, 56–61. [CrossRef]
17. Garetti, M.; Taisch, M. Sustainable Manufacturing: Trends and Research Challenges. *Prod. Plan. Control* **2012**, *23*, 83–104. [CrossRef]
18. Rausand, M.; Høyland, A. *System Reliability Theory: Models, Statistical Methods, and Applications*; Wiley: Hoboken, NJ, USA, 2004.
19. Stamatis, D.H. *Failure Mode and Effect Analysis: FMEA from Theory to Execution*; ASQ Quality Press: Milwaukee, WI, USA, 2003.
20. Nguyen, K.; Medjaher, K. A new dynamic predictive maintenance framework using deep learning for failure prognostics. *Reliab. Eng. Syst. Saf.* **2019**, *188*, 251–262. [CrossRef]
21. Rakyta, M.; Bubenik, P.; Binasova, V.; Gabajova, G.; Staffenova, K. The Change in Maintenance Strategy on the Efficiency and Quality of the Production System. *Electronics* **2024**, *13*, 3449. [CrossRef]
22. Fusko, M.; Rakyta, M.; Krajcovic, M.; Dulina, L.; Gaso, M.; Grznár, P. Basics of Designing Maintenance Processes in Industry 4.0. *MM Sci. J.* **2018**, 2252–2259. [CrossRef]
23. Fusko, M.; Rakyta, M.; Dulina, L'.; Sulirova, I.; Edl, M. Digitization in the Technical Service Management System. *MM Sci. J.* **2018**, 2260–2266. [CrossRef]
24. Fusko, M.; Rakyta, M.; Manlig, F. Reducing of intralogistics costs of spare parts and material of implementation digitization in maintenance. *Procedia Eng.* **2017**, *192*, 213–218. [CrossRef]
25. Pimminger, S.; Kurschl, W.; Panholzer, L.; Schönböck, J. Exploring the learnability of assembly tasks using digital work instructions in a smart factory. *Procedia CIRP* **2021**, *104*, 696–701. [CrossRef]
26. Holgado, M.; Macchi, M.; Fumagalli, L. Value-in-use of e-maintenance in service provision: Survey analysis and future research agenda. *IFAC-PapersOnLine* **2016**, *49*, 138–143. [CrossRef]
27. Molnar, C. *Interpretable Machine Learning*, 2nd ed.; Lulu.com: Morrisville, NC, USA, 2022; Available online: https://christophm.github.io/interpretable-ml-book/ (accessed on 16 June 2026).
28. Cummins, L.; Sommers, A.; Ramezani, S.B.; Mittal, S.; Jabour, J.; Seale, M.; Rahimi, S. Explainable Predictive Maintenance: A Survey of Current Methods, Challenges and Opportunities. *IEEE Access* **2024**, *12*, 57574–57602. [CrossRef]
29. Zemmouchi-Ghomari, L. Explainable AI for Predictive Maintenance: A Review and Standardized Evaluation Framework. *Manag. Sci. Lett.* **2026**, *16*, 15–36. [CrossRef]

30. Qu, Q.; Zhang, L.; Zheng, Y. Explainable AI for Remaining Useful Life Prediction in Industrial Systems: A Survey. *Meas. Sci. Technol.* **2026**, *37*, 042002. [CrossRef]
31. Puthanveettil Madathil, A.; Luo, X.; Liu, Q.; Walker, C.; Madarkar, R.; Qin, Y. A Review of Explainable Artificial Intelligence in Smart Manufacturing. *Int. J. Prod. Res.* **2025**, *63*, 8654–8697. [CrossRef]