applied sciences

Systematic Review

Application-Wise Review of Machine Learning-Based Predictive Maintenance: Trends, Challenges, and Future Directions

Christos Tsallis [1], Panagiotis Papageorgas [1,*], Dimitrios Piromalis [1] and Radu Adrian Munteanu [2]

[1] Department of Electrical and Electronics Engineering, University of West Attica, 12244 Athens, Greece; ctsallis@uniwa.gr (C.T.); piromali@uniwa.gr (D.P.)
[2] Electrotechnics and Measurement Department, Technical University of Cluj-Napoca, 400114 Cluj-Napoca, Romania; radu.a.munteanu@ethm.utcluj.ro
* Correspondence: ppapag@uniwa.gr

Abstract: This systematic literature review (SLR) provides a comprehensive application-wise analysis of machine learning (ML)-driven predictive maintenance (PdM) across industrial domains. Motivated by the digital transformation of industry 4.0, this study explores how ML techniques optimize maintenance by predicting faults, estimating remaining useful life (RUL), and reducing operational downtime. Sixty peer-reviewed articles published between 2020 and 2024 were selected using the preferred reporting items for systematic reviews and meta-analyses (PRISMA) 2020 guidelines, and were analyzed based on industrial sector, ML techniques, datasets, evaluation metrics, and implementation challenges. Results show that combining ML with diverse sensor data enhances predictive performance under varying operational conditions across manufacturing, energy, healthcare, and transportation. Frequently used open datasets include the commercial modular aero-propulsion system simulation (CMAPSS), the malfunctioning industrial machine investigation and inspection (MIMII), and the semiconductor manufacturing process (SECOM) datasets, though data heterogeneity and imbalance remain major barriers. Emerging paradigms such as hybrid modeling, digital twins, and physics-informed learning show promise but face issues like computational cost, interpretability, and limited scalability. The findings highlight future research needs in model generalizability, real-world validation, and explainable artificial intelligence (AI) to bridge gaps between ML innovations and industrial practice.

Keywords: machine learning; predictive maintenance; digital twins; Internet of Things; Industry 4.0; fault diagnosis; hybrid modeling

Academic Editor: Keun Ho Ryu

Received: 29 March 2025
Revised: 21 April 2025
Accepted: 25 April 2025
Published: 28 April 2025

Citation: Tsallis, C.; Papageorgas, P.; Piromalis, D.; Munteanu, R.A. Application-Wise Review of Machine Learning-Based Predictive Maintenance: Trends, Challenges, and Future Directions. *Appl. Sci.* **2025**, *15*, 4898. https://doi.org/10.3390/app15094898

1. Introduction

PdM has become an essential strategy in modern industrial settings, driven by the rapid expansion of Industry 4.0 technologies. By leveraging ML techniques, PdM aims to transition from reactive to proactive maintenance strategies, reducing unexpected failures, minimizing costs, and improving overall operational efficiency. Traditional maintenance approaches, such as run-to-failure and preventive maintenance, rely on either reacting to breakdowns or performing maintenance at fixed intervals. In contrast, PdM uses real-time data analysis and ML-based predictive modeling to optimize maintenance schedules and extend equipment lifespan [1,2].

This evolution is largely driven by AI, encompassing ML and its advanced subset, deep learning (DL). ML algorithms learn from data to make predictions or decisions without explicit programming, while DL employs multi-layered neural networks to extract features from raw unstructured data. ML techniques are typically categorized into supervised and unsupervised learning [3]. As illustrated in Figure 1, supervised learning, including support vector machines (SVMs), decision trees (DTs), and artificial neural networks (ANNs), has shown strong results in fault classification and RUL estimation, though it depends on high-quality labeled data, which is often costly or impractical in industrial settings. To address this, unsupervised methods, such as clustering algorithms and autoencoders, are increasingly applied in scenarios like anomaly detection where failure labels may be unreliable or unavailable. DL architectures like convolutional neural networks (CNNs) and long short-term memory (LSTM) networks have gained prominence for their effectiveness in handling time-series sensor data [4].

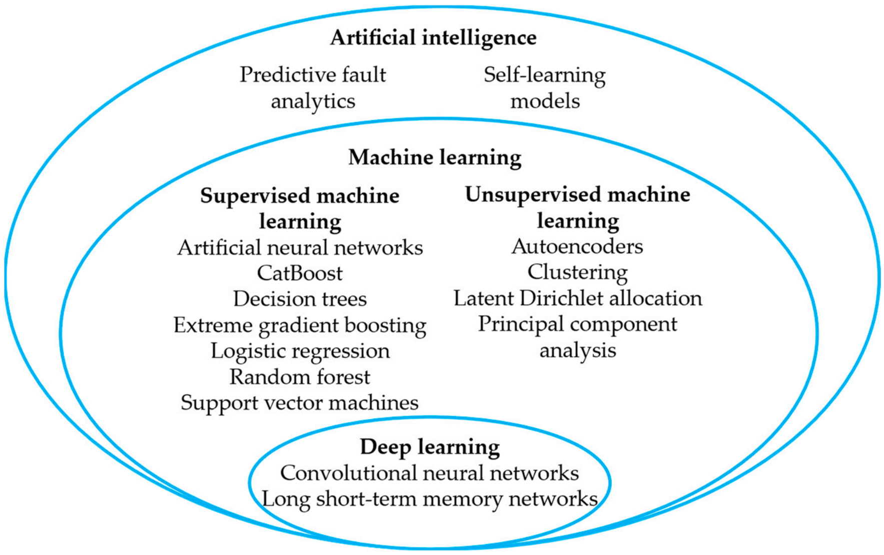

Figure 1. Conceptual hierarchy of AI, ML, and DL methods.

Recent studies affirm the growing effectiveness of AI in real-world PdM applications. Sensor-based monitoring, especially through vibration analysis, has become indispensable across industries. Techniques like ensemble learning, hybrid modeling, and transfer learning are also being explored to enhance generalizability and robustness across diverse operational conditions. However persistent challenges including data scarcity, labeling costs, and a lack of standardized benchmark datasets, continue to limit comparability and hinder large-scale adoption [5].

The integration of digital twin technology has further advanced predictive capabilities, enabling real-time synchronization between physical assets and their virtual counterparts. This supports dynamic simulations, fault diagnosis, and predictive modeling. By combining physics-based models with ML-driven analytics, these systems aim to improve prediction accuracy and reliability. Nonetheless, challenges like data heterogeneity, cloud-edge collaboration, and the computational demands of high-fidelity models remain barriers to adoption [6].

Data-driven maintenance strategies have been adopted across various industries, such as manufacturing, automotive, power electronics, and energy. In manufacturing, a major focus lies in the condition monitoring of rotating machinery, particularly bearings in grinding machines, which require continuous oversight to prevent catastrophic failures. DL-based methods, including CNNs and transfer learning, have demonstrated effectiveness in identifying fault patterns and predicting failures. Although traditional approaches have also been used, their reliance on manually extracted features often limits performance. Integrating ML with real-time sensor data has markedly improved fault classification and RUL estimation, although noise interference and limited data availability remain practical obstacles [7].

In the automotive sector, PdM has been extensively applied to critical components such as engines, batteries, and suspension systems. While supervised learning models have proven effective in fault identification, their dependence on labeled data is a limitation. Unsupervised and semi-supervised approaches are gaining traction, especially for anomaly detection in real-time vehicle monitoring. Incorporating multi-source data from onboard diagnostics, sensors, and fleet histories poses additional challenges. Concerns surrounding data privacy, computational efficiency, and DL model interpretability must also be addressed for broader implementation [8].

Power converters, which are essential for electric mobility, renewable energy, and smart grids, have also been a major focus of research. Approaches have evolved from physics-of-failure-based models to data-driven techniques like DL and probabilistic models. While model-based methods are often computationally intensive, purely data-driven ones may lack physical interpretability. Physics-informed ML is emerging as a promising compromise, incorporating domain knowledge into AI models to enhance both reliability and generalizability [9].

PdM adoption is also increasing in the energy sector, particularly for medium-voltage switchgear. Here, advanced sensor technologies such as infrared, vibration, and partial discharge sensors are combined with AI analytics. DL models like LSTMs and CNNs have improved fault diagnosis and condition monitoring, although scalability, real-time application, and sensor integration remain key hurdles [10].

A broader trend in recent research is the pursuit of hybrid models that integrate ML with physics-based simulations. In induction motors, for example, CNNs and autoencoders are widely used for fault detection. The next frontier involves combining these with multi-agent systems and digital twins. However, ensuring that AI models trained in controlled environments can generalize to dynamic real-world conditions remains a critical research challenge [11].

Therefore, this review provides a concise application-focused perspective on modern data-driven maintenance methods. By examining core algorithms, datasets, and real-world challenges in multiple industries, it aims to guide the transition from research innovation to practical adoption.

In the remainder of this paper, Section 2 details the review methodology, while Section 3 discusses domain-specific applications, key models, datasets, and metrics. Section 4 addresses recurring limitations such as data complexity and scalability, and Section 5 summarizes the main conclusions and suggests future research directions.

2. Materials and Methods

The PRISMA 2020 guidelines were followed to ensure transparency, rigor, and reproducibility. The updated framework incorporates methodological advances since PRISMA 2009, providing structured guidance for formulating research questions, designing and implementing a search protocol, defining inclusion and exclusion criteria, selecting studies, extracting data, and synthesizing findings [12].

2.1. Research Questions

To systematically guide the investigation, the following research questions (RQs) have been formulated:

RQ1: what are the industry applications and latest trends in ML-based PdM?

RQ2: which ML algorithms are commonly used for PdM, how do they perform across domains, and what evaluation metrics are used to assess their effectiveness?

RQ3: what datasets are commonly used for benchmarking PdM models?

RQ4: what challenges, and research gaps exist in implementing ML-based PdM, and what are the future directions?

2.2. Search Strategy

The literature search was conducted exclusively in Scopus, a multidisciplinary database known for its extensive peer-reviewed research coverage. With advanced tools like citation tracking, collaboration mapping, and trend analysis, Scopus enhances study accuracy and depth. Its validated reliability across disciplines makes it a robust choice for systematic literature reviews [13,14].

The final search was performed on 1 March 2025, with a keyword-based strategy designed to capture relevant studies while excluding the grey literature and unpublished works. Initially, the broad query "TITLE-ABS-KEY (predictive AND maintenance AND machine AND learning)" retrieved 4501 records. After applying the inclusion criteria, the dataset was narrowed down to 223 fully published records. To ensure both quality and temporal balance, studies were first sorted by citation count within each publication year, and a stratified selection was made by choosing an equal number of highly cited papers per year. The subsequent application of exclusion criteria, as outlined in Figure 2, resulted in a final selection of 60 articles. The exact search query, including the final set of filters and parameters applied, is detailed in Table 1.

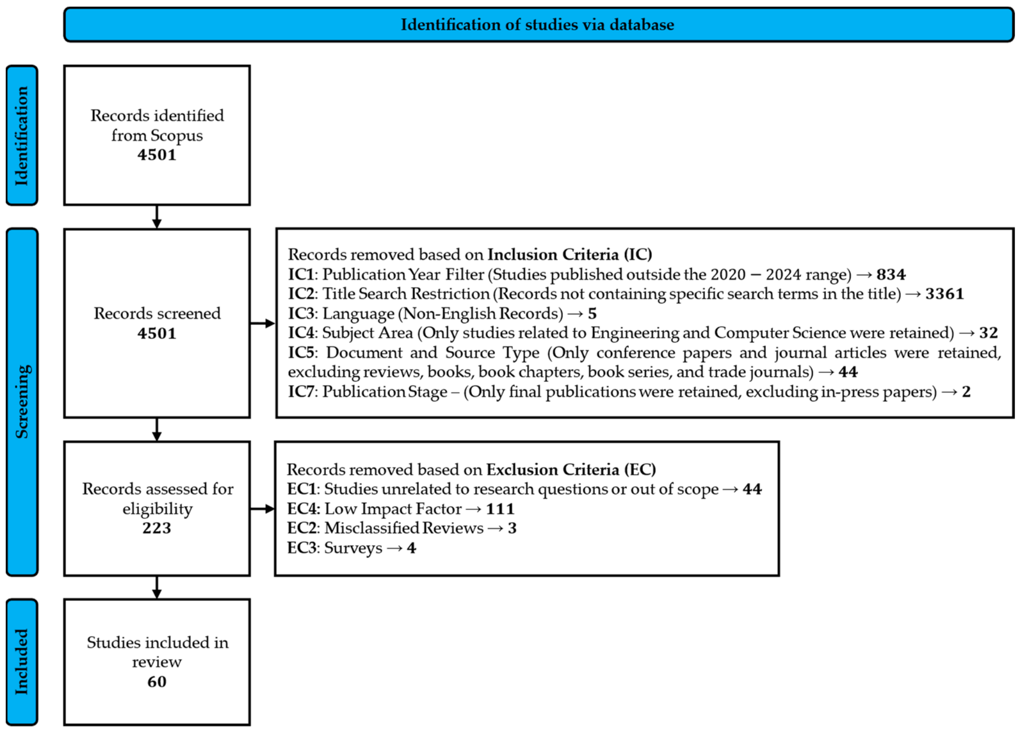

Figure 2. Flow diagram of study selection process.

Table 1. Search query used for retrieving studies from Scopus.

TITLE (predictive AND maintenance AND machine AND learning) AND PUBYEAR > 2019 AND PUBYEAR < 2025 AND (LIMIT-TO (LANGUAGE, "English")) AND (LIMIT-TO (SUBJAREA, "ENGI") OR LIMIT-TO (SUBJAREA, "COMP")) AND (LIMIT-TO (DOCTYPE, "cp") OR LIMIT-TO (DOCTYPE, "ar")) AND (LIMIT-TO (SRCTYPE, "p") OR LIMIT-TO (SRCTYPE, "j")) AND (LIMIT-TO (PUBSTAGE, "final"))

2.3. Data Extraction and Synthesis

A risk of bias assessment was conducted using a custom checklist informed by the PRISMA 2020 [12] guidelines to ensure the credibility of the findings. The assessment was applied to 60 selected studies, focusing on sample size adequacy, clarity of statistical or machine learning analysis, reproducibility of methods, and transparency of conflict of interest. Each domain was scored using a predefined three-point scale: 2 = low risk (methodologically sound, no concerns), 1 = moderate risk (some limitations present, but unlikely to alter findings), and 0 = high risk (serious methodological concerns or missing data). Total scores ranged from 0 to 8 across all domains. Studies scoring 7 or higher were classified as having a low risk of bias, while those scoring below 7 were deemed to have a moderate risk of bias. The scoring breakdown and detailed justifications for each study are provided in Supplementary Table S1.

Among the 60 studies evaluated, 55 were categorized as low risk, and 5 as moderate risk, primarily due to limitations in reproducibility or insufficient transparency regarding data or code availability. Despite these moderate concerns, no studies were identified as having a high risk of bias. Therefore, all studies were deemed methodologically sound and were retained for inclusion in this review.

Data extraction was then independently performed by the authors to minimize errors and ensure consistency. The extraction focused on key aspects of each study, including research objectives, methodological approaches, and reported outcomes. Following extraction, and in order to address the research questions, the data were synthesized according to a set of categories. These included application focus, key trends, dataset, ML techniques, evaluation metrics, key findings, challenges addressed, and suggested future directions.

3. Results

This section presents the findings in response to the first three research questions. The fourth research question, which focuses on challenges, research gaps, and future directions, is examined in detail in Section 4 and further expanded in Table A1 of Appendix A.

3.1. General Characteristics of Predictive Maintenance Studies

A text-mining approach was employed to extract and analyze the most frequently occurring terms from the reviewed studies. The dataset included titles, abstracts, and keywords from the total research papers retrieved, using the Scopus export tool. To ensure consistency and enhance data quality, a systematic preprocessing procedure was applied using Python, including the removal of non-alphabetic characters, conversion of text to lowercase, and filtering of common words, incorporating both standard and domain-specific terms. Additionally, duplicates and redundant terms were removed. A word frequency analysis was then performed to identify dominant terms. Following a similar approach to [15], a visual representation of term frequency was generated using Python's word cloud library, where the size of each word was proportional to its relative frequency within the dataset.

As shown in Figure 3, the most prominent terms highlight the strong focus on PdM and ML applications. Other key terms, such as monitoring, decision, and condition, emphasize data-driven maintenance strategies, while detection, classification, and anomaly suggest discussions on fault detection and predictive analytics. Additionally, terms like industrial, things, and internet reflect the integration of Internet of Things (IoT) technologies. Meanwhile, algorithms and forecasting point to model development considerations.

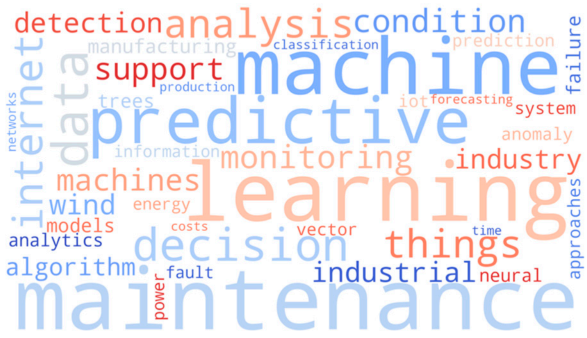

Figure 3. Word cloud of frequent terms in selected studies.

3.2. Evolution of Research Activity

Figure 4 illustrates the annual distribution of publications in the selected research domain from 2010 to 2025. To obtain a broader perspective on research trends, all filtering criteria from this study were applied except for the year filter. Additionally, the broad query, as outlined in Section 2.2, was used by considering articles that mention the key terms in their title, abstract, or keywords, rather than limiting the search to the title alone. The results reveal a steady rise in research activity, surging after 2017 and peaking in 2024, reflecting growing interest and advancements in the field. The drop in 2025 is due to the year still being in progress. The trend line confirms an overall upward trajectory, underscoring the field's increasing relevance.

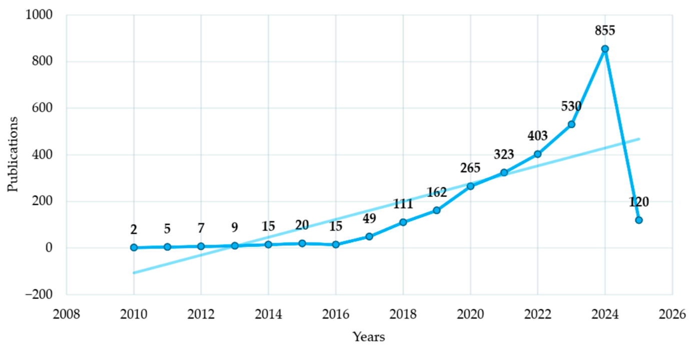

Figure 4. Annual publication trend of studies.

3.3. Applications Across Industrial Domains

To provide a structured overview, the 60 selected studies were categorized into 7 sectors based on industrial focus. Four sectors were classified based on clearly defined scopes: buildings and heating, ventilation and air conditioning (HVAC) systems, power generation and distribution, wind energy, and semiconductor manufacturing. The remaining studies were grouped into broader categories: multi-industry manufacturing, industrial equipment, and transportation. As shown in Figure 5, the number of studies included in each category is indicated in parentheses.

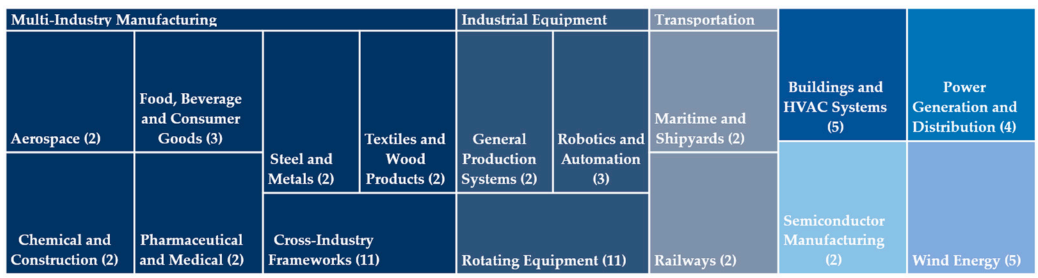

Figure 5. Categorization of reviewed studies by sector.

The multi-industry manufacturing category comprises studies whose methodologies, datasets, or implementation frameworks demonstrated relevance across at least three distinct industrial sectors. These include works utilizing cross-sector datasets, proposing generalized PdM frameworks, or presenting solutions explicitly designed for application in diverse manufacturing environments. In contrast, the industrial equipment category consists of studies focused on specific types of machinery or subsystems (such as motors, compressors, bearings, or robotic systems) regardless of the broader industrial domain. These studies typically focus on diagnostics at the equipment level using sensor-based approaches and are defined by their application within a specific industrial domain.

3.3.1. Multi-Industry Manufacturing

ML-based PdM across sectors uses IoT, real-time data, and scalable AI tools. Ensemble methods like random forest (RF) and extreme gradient boosting (XGBoost) [16] effectively process high-dimensional sensor data for fault prediction. Fog computing with logistic regression and genetic algorithms enhances real-time decision making in Industry 4.0 contexts [17]. RF supports cloud-based systems for estimating RUL [18], while, in [19], a hybrid DL model achieves high predictive performance using engine and battery data. Algorithm selection studies show that RFs and DTs excel on small datasets, while k-nearest neighbors (KNNs) suits larger ones [20]. For small and medium-sized enterprises (SMEs), a structured phased approach is recommended, emphasizing interdisciplinary teams and training [21]. Double attention or $A^2 - LSTM$ prioritizes features using attention mechanisms for improved aircraft manufacturing predictions [22], while cyber-physical systems using historical data offer cost-effective solutions for SMEs [23]. The balanced k-star model addresses class imbalance with high accuracy and interpretability [24]. Building on these advancements, gradient boosting (GB) and XGBoost remain top performers in RUL prediction and maintenance planning, offering cost and reliability benefits [25,26].

In the steel and metals industry, a cloud-based fault classification system for hot-rolling mills integrated digital twin and XGBoost [27], while LSTM autoencoders enabled early fault detection in laser welders, with 97.3% accuracy [28].

Textiles and wood product manufacturing emphasizes event-driven systems and cloud-based ML. A tree-based system using event logs achieved 98.9% accuracy in RUL prediction without hardware upgrades [29], while AdaBoost and IoT sensor data classified machine stoppages in knitting systems with 92% accuracy [30].

Food, beverage, and consumer goods applications prioritize cost-effective decision support and ensemble learning. A DT combined with a failure mode, effects, and criticality analysis system optimized food production maintenance with 96.3% accuracy [31]. Similarly, Arduino sensor systems combined with RF improved overall equipment effectiveness by 13.1% and reduced failures by 62.4% [32]. Building on this trend, ensemble learning on IoT data reached 95.9% accuracy in wafer stick production [33].

In the pharmaceutical and medical sectors, a hybrid framework using boosted decision trees and neural networks reduces production disruptions through accurate maintenance time estimation [34]. Expanding on this approach, a multi-model system combining SVMs, DTs, and ANNs predicted equipment failures across over 13,000 medical devices [35].

Chemical and construction industries apply high accuracy and interpretable models. In concrete manufacturing, CatBoost outperformed six other classifiers ($F1_{score} = 0.98$) using sensor data [36]. For petrochemical compressors, XGBoost with Shapley values enabled fault prediction and root-cause transparency (area under the curve or AUC $> 90\%$) [37].

In aerospace, federated learning models predicted anomalies and RUL in aero-engines without raw data sharing [38]. In parallel, ensemble approaches, such as RFs and random under-sampling boosted trees, were effective for RUL and fault detection in turbofan engines and pump systems [39].

3.3.2. Industrial Equipment

Industrial IoT and multi-sensor approaches enhance rotating equipment maintenance. A low-cost system used vibration, temperature, and sound signals for bearing health monitoring, with DTs outperforming other models [40]. Utility theory was integrated with GB and RF to optimize fault detection in pump systems [41]. LSTM networks outperformed traditional models for RUL prediction in compressors [42], while SVMs achieved top results for motor fault diagnosis [43]. RFs also proved effective in oil analysis for gearbox monitoring [44] and IoT driven motor diagnostics [45]. XGBoost offered the best accuracy–speed tradeoff in bearing systems [46], and both regression and ANN models predicted tool wear under varying lubrications [47]. Applications in mining, thermal systems, and cement plant fans further illustrate the adoption of ML models across industries [48–50].

General production systems combine DL and optimization. A CNN-bidirectional LSTM model, as demonstrated in [51], detected actuator and turbine faults with 96% recall using the MIMII dataset. In [52], a predictive framework combined ANNs, SVMs, and the extended great deluge (EGD) algorithm to forecast degradation and optimize multi-level maintenance scheduling.

Robotics and automation systems benefit from both historical and real-time data. An ANN model trained on enterprise management records and average time-to-failure statistics predicted packaging robot failures with 91% accuracy without an IoT infrastructure [53]. Correspondingly, a discrete Bayes filter (DBF) outperformed naïve Bayes (NB) in forecasting degradation in robot power systems [54], while, in [55], a digital twin framework using RF accurately detected conveyor belt anomalies like chain slack.

3.3.3. Transportation

Maritime and shipyard systems are using hybrid models to detect faults in low-sensor environments. A multi-model approach (GB, LSTM, one-class SVM) enabled early fault detection in vessel bearings using entropy-based features [56]. Similarly, an SVM with principal component analysis (PCA) achieved predictive accuracy for ballast pump failures with only historical data [57].

Railway applications demonstrate the value of legacy data and sensor fusion. DTs trained on historical records achieved 90% accuracy without real-time sensors [58]. Leveraging locomotive sensor data, another study proposed a hybrid framework combining SVM, RF, and digital twin technology, emphasizing the role of digital logging and IoT integration [59].

3.3.4. Power Generation and Distribution

For transformers, SVMs trained on data from 16,000 units achieved over 95% accuracy and reduced costs by 13% [60]. In hydroelectric plants, RFs predicted failures within 12–48 h at 98% accuracy [61]. LSTM autoencoders identified thermal faults in generators with 99% accuracy [62], while Gaussian process classifiers combined with IoT sensors achieved 99.56% accuracy in monitoring electrical panels [63].

3.3.5. Wind Energy

A hybrid statistical-ML approach, as presented in [64], used DTs and RFs to diagnose turbine faults with over 92% accuracy. In [65], deep neural networks applied to supervisory control and data acquisition (SCADA) data predicted anomalies up to 72 h in advance. As outlined in [66], feature selection was prioritized over model complexity in forecasting subsystem failures. A CNN-LSTM hybrid model, introduced in [67], effectively distinguished soft and hard turbine failures using IoT data. Additionally, vibration monitoring using bagged trees achieved 87% accuracy under lab conditions [68].

3.3.6. Buildings and HVAC Systems

A building information modeling and IoT system, as proposed in [69], used SVMs and ANNs to predict chiller degradation, with SVMs yielding the best performance. In [70], autoencoders detected HVAC anomalies via building automation data. As shown in [71], an LSTM model forecast heating system failures up to seven days ahead. Energy use and maintenance needs in active chilled beam systems were predicted using XGBoost and Gaussian process regression, although the study partially relied on synthetic data [72]. A more comprehensive hybrid framework combining SVM, DT, KNN, prophet, and seasonal autoregressive integrated moving average (SARIMA) was applied to hospital HVAC systems using building management system and computerized maintenance management system data [73].

3.3.7. Semiconductor Manufacturing

ML plays a critical role in supporting wafer defect detection and enhancing production reliability. Logistic regression with false discovery rate correction achieved 94.64% accuracy on the SECOM dataset, enhanced by synthetic minority oversampling technique (SMOTE), PCA, and latent Dirichlet allocation (LDA) preprocessing [74]. In [75], RF achieved 93.62% accuracy in wafer failure prediction, aiding maintenance scheduling and minimizing interventions.

3.4. Machine Learning Architectures

ML models demonstrate versatile and robust performance across industrial sectors, with each domain favoring architectures aligned to specific tasks. Tree-based models such as RFs, DTs, XGBoost, and GB are among the most frequently employed, as illustrated in Figure 6. Temporal networks like LSTM also appear prominently in commonly used architectures.

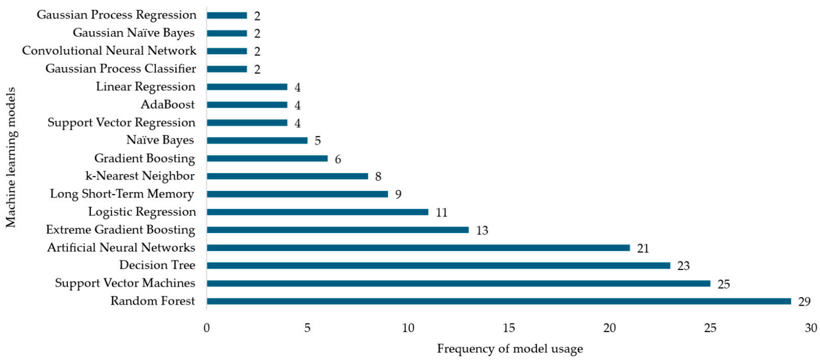

Figure 6. Frequently employed architectures for study evaluation.

In aerospace, federated models like FedLSTM offer superior RUL estimation, while traditional models show acceptable but limited accuracy. Buildings and HVAC systems benefit from SVMs and LSTM-based hybrids for early fault detection, though precision varies. CatBoost and XGBoost lead in chemical and construction domains, while ANN dominates general production for fault diagnosis and maintenance scheduling. Ensemble models consistently enhance accuracy in food, maritime, and rotating equipment sectors. Power systems, robotics, and wind energy increasingly rely on DL hybrids for anomaly detection and prediction, often surpassing conventional classifiers. These trends align with the evaluation emphasis shown in Figure 7, where accuracy, precision, recall, and F1-score are the most cited performance metrics.

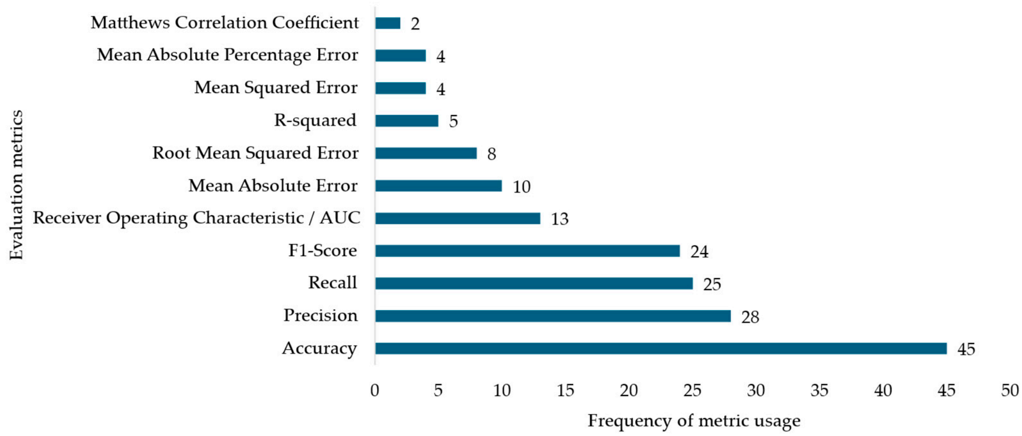

Figure 7. Frequently employed metrics for study evaluation.

However, despite the widespread reporting of these metrics, their interpretation often lacks critical depth, particularly under conditions of class imbalance or imprecise labeling. For example, several studies achieve high overall accuracy (e.g., $>$ 95% with RF or SVM models) without disclosing the distribution of faulty versus non-faulty instances. In imbalanced datasets like SECOM, this can be misleading, as high accuracy may result from bias toward the majority class. Although resampling techniques like SMOTE are sometimes used to address imbalance, many studies omit such details, raising concerns about the validity of reported results. Furthermore, inconsistencies in threshold settings and the lack of confusion matrices or AUC curves complicate meaningful cross-study comparisons and hinder critical evaluation of false positives and false negatives, which are crucial in maintenance-critical systems.

In high-stakes industrial sectors such as aerospace, energy, and manufacturing, domain-specific metrics like RUL offer a more actionable perspective than traditional classification metrics. RUL estimates the time remaining before a component fails, enabling optimized maintenance scheduling and reduced downtime. Studies leveraging datasets like CMAPSS and National Aeronautics and Space Administration (NASA) increasingly model RUL as a regression task, evaluated through root mean square error and mean absolute error, with architectures such as LSTM and CNN excelling due to their ability to capture temporal degradation patterns.

To enhance the rigor and operational relevance of predictive maintenance models, future research should consistently report dataset balance ratios, sampling methods, and error distributions, alongside standard metrics. More importantly, integrating prognostics-oriented metrics like RUL into evaluation frameworks not only improves interpretability but also ensures that model performance aligns with real-world maintenance objectives.

3.5. Commonly Used Datasets

Several open-source datasets, as shown in Table 2, are widely used in PdM and fault diagnosis due to their accessibility and relevance. CMAPSS and NASA turbofan datasets offer multivariate time-series data for RUL prediction. MIMII provides audio and vibration signals from faulty and normal industrial machines for anomaly detection. SECOM contains 591 features from semiconductor processes for quality classification. AI4I 2020 includes telemetry, error logs, and failure labels for prognostics. NASA's milling and bearing datasets deliver sensor data on tool wear and bearing degradation, supporting diagnostics and condition monitoring.

Table 2. Open-source datasets by domain.

Dataset	Domain
CMAPSS (NASA) NASA turbofan engine degradation	Aerospace
MIMII	General production systems
SECOM	Semiconductor manufacturing
AI4I 2020 PdM Milling (NASA's prognostics center)	Cross-industry frameworks
Vibration (NASA's prognostic center)	Rotating equipment

The remaining datasets are derived from real-world industrial operations and typically contain time-series sensor data such as vibration, temperature, pressure, or flow, along with maintenance logs, inspection reports, and system telemetry. Most are proprietary due to confidentiality agreements, commercial sensitivity, or infrastructure concerns, and are collected through internal monitoring systems or research collaborations. Some are synthetically generated using simulation platforms but remain restricted to private use or licensed access.

Across both public and private datasets, a key challenge lies in the aggregation and preprocessing of heterogeneous data. PdM systems, particularly in industrial Internet of Things (IIoT) and large-scale manufacturing environments, depend on effective integration of multi-source information. In modern architectures, data are typically collected at the edge (from sensors, devices, or loggers) and streamed centrally using real-time platforms like Apache Kafka and Spark. These tools support temporal synchronization via micro-batch processing and map-reduce operations, which structure raw inputs into logically ordered data streams [76].

Moreover, data privacy and consistency become central concerns, especially in federated settings. Privacy-preserving schemes such as differential privacy and homomorphic encryption are increasingly used to ensure secure data aggregation, particularly in IIoT federated learning architectures. Datasets are preprocessed using techniques like normalization and enhanced through strategies to address missing values, misaligned timestamps, and cross-sensor fusion. These steps are foundational for ensuring dataset quality, whether the source is a public benchmark or a proprietary industrial stream [77].

4. Discussion

The findings from this review underscore how ML-driven PdM has expanded across a diverse set of industrial settings, including manufacturing, power generation, buildings, healthcare, and transportation. Addressing the first research question, which investigated the industry applications and latest trends, the analysis shows that many sectors have adopted ensemble models and DL architectures to manage complex time-series data from sensors such as vibration, temperature, and pressure. Within manufacturing, there is a pronounced move from purely rule-based or preventive maintenance protocols to proactive data-rich strategies leveraging AI, while heavy industries, including steel, mining, and petrochemicals, are increasingly incorporating digital twins to enhance decision making. This broadened application scope reveals a pattern in which access to multivariate sensor data, combined with scalable ML frameworks, helps in detecting faults earlier and reducing the risk of unplanned downtime.

A key aspect of the second research question highlights that tree-based methods, including RF and GB, dominate a large part of the literature, often due to their interpretability and ease of parameter tuning when dealing with high dimensional signals. DL networks, notably LSTM and CNN architectures, are particularly prominent when analyzing noisy or unstructured data such as acoustic signals or multivariate time-series measurements in domains like wind turbines, rotating machinery, and building HVAC systems. The observed performance metrics (accuracy, precision, recall, F1-score) confirm that these advanced models outperform simpler statistical or rule-based approaches in many real-world scenarios. Still, model selection remains context-dependent. Low-sensor environments favor robust ensemble approaches, while data-rich environments use more complex neural networks for greater predictive accuracy.

Addressing the third research question on common benchmark datasets, the review underscores the repeated appearance of NASA's CMAPSS, MIMII, SECOM, and other publicly available industrial datasets. These open datasets remain essential for reproducible evaluations, although real-world applications still rely heavily on proprietary data. Data scarcity, confidentiality restrictions, and inconsistent sensor coverage frequently limit the ability to generalize findings across different industrial environments.

Building on these observations, one recurring challenge in ML-based PdM systems is the limited diversity of machines and sensor types represented in widely used datasets. Models trained on homogeneous data, such as single-type machinery or mono-sensor inputs, often underperform when deployed in complex industrial settings characterized by varied equipment and environmental conditions. Future research should prioritize the development and utilization of multi-machine multi-sensor datasets, integrating telemetry from pumps, compressors, conveyor belts, robotic actuators, and other asset types across acoustic, vibration, infrared, and temperature sensing modalities. This broader inclusion would enhance generalization and significantly strengthen models' resilience to unforeseen fault patterns. Moreover, data fusion frameworks and federated learning offer promising avenues to achieve this diversity while respecting industrial privacy and data-sharing constraints.

By examining the fourth research question, which explores challenges and future directions, this review reveals that issues with incomplete or imbalanced data, lack of standardization in data formats, and the high computational overhead of complex models continue to slow the widespread adoption of ML-based PdM. Sectors such as aerospace and automotive have demonstrated encouraging results with distributed or federated learning to reduce the burden of transferring massive amounts of sensor data, but these frameworks require more robust edge or fog computing infrastructure. Additionally, the interpretability of black-box deep networks remains a frequent point of concern for industrial practitioners who need transparent reasoning to justify significant maintenance decisions. Knowledge-driven approaches and explainable AI techniques are therefore emerging as critical trends that could bolster trust in automated PdM systems.

Overall, the review indicates that intelligent maintenance systems are a vital enabler of data-centric strategies in Industry 4.0. However, further validation in live industrial settings is necessary to address real-time constraints, deployment costs, and integration with legacy systems. Encouragingly, among the 60 studies surveyed, 40 papers have already taken their models out of the lab and into everyday service. These "yes" cases stream live sensor data from pumps, turbines, production lines, and similar assets, raise work orders automatically through the plant's maintenance software, and report tangible payoffs such as shorter downtime or higher overall equipment effectiveness. Four additional papers sit in the "partial" category; they analyzed authentic plant data retrospectively but have yet to let their models issue live alerts. The remaining 16 papers are still at the "no" stage, evaluated only on test rigs, simulations, or public benchmark datasets. Details of this deployment classification are provided in Supplementary Table S1.

A common thread runs through the fully deployed studies and offers a practical reusable roadmap for others. Each deployed study relies on an edge-to-cloud data pipeline that uses open protocols for consistent real-time collection of sensor streams. The predictive models themselves are packaged in containers, enabling hot swaps and automatic retraining whenever data drift is detected. A lightweight explainability layer helps technicians understand and trust the alerts. Finally, the prediction output is wired straight into the plant's computerized maintenance management system so that every alert is converted immediately into an actionable work order. Because this architecture has been demonstrated in wind farms, refineries, hospitals, and discrete part factories alike, it serves as an adaptable framework for moving AI models from controlled tests to reliable real-world deployments.

Research should continue to explore hybrid models that fuse domain knowledge with data-driven analytics to mitigate challenges associated with explainability, data quality, and domain adaptation. Such efforts will likely accelerate the deployment of robust PdM frameworks that can adapt to evolving operational demands while remaining economically viable and transparent to stakeholders.

5. Conclusions

By examining the intersection of ML techniques and maintenance strategies, this review highlights both significant advancements and persistent challenges across diverse industrial applications. Data-driven models, particularly ensemble methods and deep neural networks, have demonstrated superior performance compared with traditional maintenance protocols, especially in fault detection and RUL estimation. However, several critical barriers continue to limit real-world adoption.

One of the foremost challenges is the dependency on labeled data in supervised learning approaches. In PdM, obtaining high-quality labeled datasets, where each instance is annotated with a known fault or failure type, is both time-consuming and expensive. This scarcity becomes especially problematic in high-risk sectors where failures are rare but impactful. Unsupervised learning can mitigate this issue by uncovering patterns, detecting anomalies, and discovering hidden structures in unlabeled sensor data. These capabilities make unsupervised methods well suited for maintenance environments where labeled datasets are incomplete or unavailable, offering scalable alternatives for early fault detection and health monitoring [78].

In addition, the unreliability and vulnerability of sensor data, whether caused by hardware faults or cyber threats, present further challenges to the robustness of PdM systems. To address these risks, ensuring the security and integrity of data within cyber-physical systems (CPSs) is essential. Industrial environments remain particularly exposed to deception, replay, and denial-of-service attacks, all of which can distort sensor readings and degrade model performance, leading to inaccurate or missed fault predictions in critical infrastructures such as smart grids and water treatment facilities [79]. Recent research demonstrates that ensemble deep learning methods can enhance the detection of such attacks, even when cyber incidents are infrequent and datasets are imbalanced [80]. In high-stakes sectors like healthcare and energy, the consequences of undetected tampering are especially serious. Techniques based on monitoring the behavior of physical systems, such as identifying real-time deviations, have proven effective in detecting advanced threats that bypass conventional information technology-based detection systems [81].

Another recurring theme is the lack of standardized frameworks that effectively integrate domain knowledge with algorithmic intelligence. While physics-informed learning and digital twins offer promising integration strategies, their implementation is constrained by edge computing limitations, communication overheads, and model interpretability concerns. To overcome these barriers, future research should prioritize cross-disciplinary collaborations that unify data science with engineering domain expertise, alongside efforts to establish interoperable data standards and benchmarking practices.

Ultimately, as organizations move toward proactive and intelligent maintenance ecosystems, attention must shift toward building explainable, scalable, and continuously evolving ML models. Emphasizing lifelong learning, transparent decision making, and cyber-resilience will be essential to accelerate the maturity of predictive maintenance technologies and to ensure operational continuity, safety, and efficiency in increasingly complex industrial systems.

Supplementary Materials: The following Supporting Information can be downloaded at: https://www.mdpi.com/article/10.3390/app15094898/s1, Table S1: Results of the systematic literature review.

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

Funding: The APC was funded by the Special Account for Research Grants of the University of West Attica.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

AI	Artificial intelligence	ML	Machine learning
ANN	Artificial neural network	NASA	National Aeronautics and Space Administration
CMAPSS	Commercial modular aero-propulsion system simulation	NB	Naïve Bayes
CNN	Convolutional neural network	PCA	Principal component analysis
CPS	Cyber-physical system	PdM	Predictive maintenance
DBF	Discrete Bayes filter	PRISMA	Preferred reporting items for systematic reviews and meta-analyses
DL	Deep learning	RF	Random forest
DT	Decision tree	RQ	Research question
EGD	Extended great deluge	RUL	Remaining useful life
GB	Gradient boosting	SARIMA	Seasonal autoregressive integrated moving average
HVAC	Heating, ventilation, and air conditioning	SCADA	Supervisory control and data acquisition
IIoT	Industrial Internet of Things	SECOM	Semiconductor manufacturing process
IoT	Internet of Things	SLR	Systematic literature review
KNN	K-nearest neighbor	SME	Small and medium-sized enterprise
LDA	Latent Dirichlet allocation	SMOTE	Synthetic minority oversampling technique
LSTM	Long short-term memory	SVM	Support vector machine
MIMII	Malfunctioning industrial machine investigation and inspection	XGBoost	Extreme gradient boosting

Appendix A Architecture Analysis

Table A1. Summary of trends, challenges, and future directions across industrial domains.

Sector	Citation Index	Trends	Challenges	Future Directions
Cross-industry frameworks	[16]	Real-time IoT data with ensemble learning models improves factory maintenance decisions	Noisy and imbalanced data, single-site generalizability, and model retraining needs	Add rare failure types, enhance generalization, and evaluate economic impact
	[17]	Fog computing and genetic optimization support low-latency maintenance in smart factories	Real-world validation gaps, parameter tuning, and reduced failure data from undersampling	Explore DL, test in live factories, and assess cost-effectiveness
	[18]	Big data analytics with cloud-based decision systems improve PdM planning	Hard-to-label failure data, data variety issues, and limited case study generalizability	Use incremental learning, adapt to more machines, and include cost analysis
	[19]	Hybrid ML and optimization support failure forecasting for sustainable processes	Overfitting risk, dataset limitations, and high complexity in live data processing	Broaden datasets, reduce computational load, and explore external condition impact
	[20]	Comparing different learning models to match data types in predictive tasks	Model performance depends on dataset size; lacks DL and cost analysis	Combine models, use deeper networks, and address imbalance in rare failure events
	[21]	Structured adoption models help SMEs start using PdM	Limited AI skills, small budgets, and messy data make implementation difficult	Improve model training tools, generalize for more industries, and simplify integration
	[22]	Attention-based DL improves maintenance prediction accuracy	Needs lots of clean labeled data and high computing power	Use better feature selection, test across industries, and reduce complexity
	[23]	Using machine status instead of sensor data lowers cost for SMEs	Indirect data may miss detailed faults; hard to standardize models	Add sensor fusion, test double-loop CPS, and improve prediction with smart learning
	[24]	Use of interpretable ML to handle imbalanced maintenance data	Risk of losing data from undersampling and high computation for large datasets	Extend to more industries, use live data, and explore hybrid learning methods
	[25]	Real-time sensor data and model tuning to improve prediction accuracy	Limited by benchmark datasets and lack of DL for complex patterns	Apply in real-world setups and use advanced models for broader equipment types
	[26]	Using AI and cost–benefit insights to optimize failure prediction	Depends on one dataset, lacks real-time data, and ignores full cost of large-scale deployment	Add real-time monitoring, test on other industries, and refine model combinations

Table A1. *Cont.*

Sector	Citation Index	Trends	Challenges	Future Directions
Steel and metals	[27]	Combining simulation with real-time data through digital twins; cloud platforms used for real-time monitoring; selecting key sensor features for prediction	High dependency on past data and simulations; limited testing across different setups; data gaps and system noise may reduce accuracy	Improve model accuracy with more live data; expand to other industries; evaluate long-term costs and scale-up potential
	[28]	Using memory-based models to detect early failures; relying on real plant data for real-time prediction; increasing use of unsupervised learning for rare faults	Limited failure examples; trained only on one plant's data; high false positives; lacks full real-time deployment	Reduce false alerts; test in other plant settings; include full-scale live monitoring and cost evaluation
Textiles and wood products	[29]	Using machine log data instead of extra sensors; combining IoT with learning models to predict failure time; running systems on big data platforms for many machines at once	Log data may miss failure signs; processing large event files is complex; models tested only on woodworking machines	Test in other industries; use more types of failure predictions; improve models for real-time use and general use
	[30]	Real-time machine tracking with connected devices; using boosting models to predict machine stops; combining old and live data for better accuracy	Some failures are harder to predict; model only trained on one type of machine; system relies mainly on one data type	Add more machine types and data sources; improve how minor failure types are handled; check how system performs in new settings
Food, beverage, and consumer goods	[31]	Use of decision trees with cost and risk analysis to guide PdM strategies	Requires complete and accurate data; complex scaling in large operations; relies on expert input	Apply to more industries, enhance cost modeling, and integrate smarter algorithms for broader decision making
	[32]	Focus on low-cost sensors and models to improve equipment uptime in small-scale setups	Limited model diversity, short testing period, and lower sensor accuracy may affect long-term performance	Improve model range, use higher quality sensors, and test system stability over time
	[33]	Integration of sensor data and ensemble models for real-time fault detection in food manufacturing	Limited features, trained on one setup, and computing demands may hinder fast decision making	Add more machine variables, adapt model for real-time use, and evaluate financial benefits of deployment

Table A1. *Cont.*

Sector	Citation Index	Trends	Challenges	Future Directions
Pharmaceutical and medical	[34]	Combining PdM with production scheduling using simulation and smart models	High computing needs, limited testing beyond one factory, and data issues like noise and missing values	Apply to more industries, include deeper models, and improve real-time performance and adaptability
	[35]	Using data-driven models to predict failure types and timing in healthcare equipment	Hard to detect sudden failures, some data gaps, and unclear long-term costs	Add real-time data sources, test newer models, and explore economic impact for wider use
Chemical and construction	[36]	Use of advanced learning models to predict failures in construction machinery; growing role of real-time sensor data in planning maintenance	Models depend on a few sensor indicators; no deep models tested; results from one site may not apply broadly	Explore deeper models, test in varied environments, and assess long-term costs and benefits
	[37]	Increased focus on making prediction models explainable; combining ML with diagnostic tools in heavy industries	Hard to generalize from one refinery case; interpreting results in real time is resource-heavy	Add more real-time data, simplify models for faster use, and adapt the approach to different equipment
Aerospace	[38]	Growing use of decentralized learning to protect data and reduce network use; use of edge–fog–cloud models; lightweight models for faster, safer updates	Lower accuracy due to uneven data; resource limits on edge devices; risk of misleading results from combined data	Improve model handling for uneven data; reduce edge computing demands; test in real industrial setups
	[39]	Adoption of ML for predicting failures; preference for ensemble models; strong focus on data preparation before modeling	Use of only simulated data limits real-world relevance; no DL tested; manual feature work is time-consuming	Use real-time industrial data; try advanced models; automate feature selection for better performance
Rotating equipment	[40]	Use of real-time monitoring with wireless data transmission in low-cost industrial systems	Small experimental dataset, increased complexity with more features, and lack of cost analysis	Improve speed and accuracy of models and explore DL in real environments
	[41]	Use of utility theory with ML for better decision making in maintenance	Limited by binary models, small dataset from one site, and no cost–benefit analysis	Apply to diverse settings, test adaptive models, and evaluate financial viability

Table A1. *Cont.*

Sector	Citation Index	Trends	Challenges	Future Directions
Rotating equipment	[42]	Adoption of LSTM models and Grafana dashboards for time-based maintenance insights	Sensor noise, missing data, model complexity, and lack of transformer model exploration	Improve data quality, integrate hybrid models, and explore edge computing solutions
	[43]	Two-phase detection and classification of motor faults using vibration data and SVM	Small and narrow dataset, reliance on one type of sensor, inconsistent fault classification	Expand datasets, add more sensors, and use digital twins for model training
	[44]	Shift from vibration to oil analysis with ML for fault detection	Imbalanced data, no real-time monitoring, and limited model diversity	Use real-time data, combine methods like vibration and oil, explore deeper models
	[45]	Real-time motor monitoring using IoT sensors and ML	Small dataset, limited sensor variety, and communication delays when scaling	Add more sensor types, test on larger systems, and assess financial feasibility
	[46]	Use of ensemble and DL models to analyze vibration data for bearing faults	Small dataset, missing real-world validation, and no advanced feature extraction	Expand datasets, improve signal processing, and test in real-time factory environments
	[47]	AI-driven modeling of machining force and tool wear under different lubrication methods	Hard to generalize, ANN tuning issues, and limited to lab conditions	Validate in real plants, add real-time monitoring, and test more learning techniques
	[48]	Cloud-based ML and sensor integration for predicting mining equipment faults	Single-site data, manual data entry, and lack of DL models	Improve real-time input, expand model range, and assess economic feasibility
	[49]	Multi-sensor fusion and data preprocessing to improve motor condition classification	Controlled test setting, no DL models, and lack of real-world deployment	Test in diverse plants, explore advanced algorithms, and evaluate cost–benefit
	[50]	Semi-automated diagnostics with frequency domain vibration analysis and ensemble models	Imbalanced data, limited feature methods, and no external factor handling	Use richer datasets, refine features, and explore real-time economic implementation
General production systems	[51]	Rise of hybrid DL models for accurate fault detection; growing use of real-time data from industrial sensors	High model complexity, limited real-world testing, and narrow fault coverage from one dataset	Test in real factories, reduce computing demands, and explore models that predict time to failure
	[52]	Data-driven decision making combining ML and optimization; use of smart scheduling to reduce costs	Model assumptions limit real-time use, optimization is slow for big systems, and lacks detailed cost info	Add live data from sensors, improve speed for large setups, and study financial impact across industries

Table A1. *Cont.*

Sector	Citation Index	Trends	Challenges	Future Directions
Robotics and automation	[53]	Using past failure data from internal systems instead of real-time sensors; applying neural networks without IoT devices	Limited failure records; no real-time updates; basic models may miss complex patterns	Collect more data; connect predictions to live systems; explore newer learning methods
	[54]	Real-time data and ML used to track slow wear in smart factories; models adapt to uncertain sensor readings	Hard to combine many sensor types; limited use outside tested factory; matching score still low	Improve model accuracy; test in new settings; manage data gaps better
	[55]	Digital twins closely mirror real systems for early warning; hybrid learning methods used to balance speed and accuracy	High data volume slows systems; model fits one setup only; DL not used due to cost	Use smarter models; improve speed for real-time use; test long-term costs and broader use
Maritime and shipyards	[56]	Use of real-time ship data and multiple ML models to detect engine anomalies early	Limited sensors on vessels, difficulty in spotting slow damage, and high false alarms from some models	Improve feature design, add models to predict remaining part life, and apply to more ship systems
	[57]	Shift to predictive methods without sensors using historical pump data for early failure alerts	Small data size, lack of sensors, and missing outside factors like temperature or water quality	Add sensor technology, collect better data, and apply to other shipyard systems
Railways	[58]	Move from manual work to data-driven maintenance; use of historical records instead of live sensors; open-source tools used for modeling	Hard to connect models to current systems; no real-time data used; data quality and general use across countries not tested	Add real-time data; test with more systems; use smarter models for better results
	[59]	Shift toward digital checks with sensor data; mix of history and real-time used for training; focus on improving maintenance through simple models	Records still on paper; few data types collected; model not yet tested in real use	Test models with real data; collect wider data types; plan full digital upgrade

Table A1. *Cont.*

Sector	Citation Index	Trends	Challenges	Future Directions
Power generation and distribution	[60]	Shift from reactive to PdM using historical data; increasing use of supervised models for maintenance planning	Integrating prediction models into existing workflows; lack of some condition data; model assumes similar environmental conditions	Add more condition variables; explore advanced models; apply to broader networks
	[61]	Use of real-time sensor data for prediction; comparison of different learning models; growing use of smart technology in energy systems	Selecting key sensor inputs; model relies on past data only; limited to one plant setting	Expand to other industries; refine variable selection; integrate real-time data streams
	[62]	Adoption of DL for early fault detection; use of time-series data from SCADA systems; handling rare failures with data balancing techniques	Dataset imbalance; limited sensor types used; only six months of data; high computational needs	Collect longer term data; use more sensor types; improve real-time deployment; assess cost-effectiveness
	[63]	IoT-enabled anomaly detection in electrical panels using sensor fusion and lightweight ML	Sensor sensitivity, thermal camera cost, and integration issues	Scale to large systems, optimize sensor design, and improve real-time processing
Wind energy	[64]	Combining sensor data with ML for early fault prediction	Imbalanced data, limited dataset size, and high computation needs	Real-time adaptation, wider turbine coverage, and cost-effectiveness analysis
	[65]	DL used with condition monitoring to predict faults in advance	Data imbalance, limited by SCADA data, and inconsistent turbine behavior	Improve real-time detection and expand sensor integration
	[66]	Emphasis on data preprocessing and feature selection over complex models	Small dataset, missing values, and limited generalizability	Test on larger datasets, explore DL, and assess economic impact
	[67]	IoT and hybrid DL models used for predictive analytics	Sensor reliability, high data volume, and high model complexity	Improve scalability, reduce processing time, and test across diverse environments
	[68]	Vibration monitoring paired with ML in controlled settings	Not tested in real-world conditions and moderate prediction accuracy	Field validation, deeper models, and multi-sensor approaches

Table A1. *Cont.*

Sector	Citation Index	Trends	Challenges	Future Directions
Buildings and HVAC systems	[69]	Growing use of real-time building data, integration of building models and sensors for smart upkeep	Limited integration in large-scale systems, sensor reliability, and data quality	Test in bigger buildings, improve data handling, explore deeper learning for better accuracy
	[70]	Use of deep models and smart sensors in buildings, shift toward anomaly detection in maintenance	Small datasets, missing sensor data, unclear model results, narrow case testing	Collect more data over time, validate in different buildings, and assess financial impact
	[71]	Adoption of time-based models to predict failures early using real-world heating system data	Imbalanced data, inconsistent device settings, and early signs of failure are hard to detect	Balance datasets better, test deeper models, and expand data coverage for long-term performance
	[72]	Combining maintenance planning with energy saving in HVAC systems using smart prediction models	Dependence on synthetic data, limited real-world diversity, and high computing needs	Add real sensor input, test across climates, and streamline models for wider building applications
	[73]	Use of short- and long-term models in hospital maintenance, combining building and maintenance data	Small data window, only tested on one type of HVAC system, and missing broader testing	Expand to more systems, gather data longer, and test across hospital equipment for broader use
Semiconductor manufacturing	[74]	Rise in use of ML to handle complex manufacturing data; use of data balancing and feature reduction to improve model accuracy	Too many features and too few failure cases; limited to past data; lacks testing in real factories	Add real-time data; test models in working factories; explore deeper learning methods
	[75]	More advanced models used for predicting equipment issues; shift to data-driven planning using manufacturing sensor data	Data does not reflect real-world factory conditions; oversampling may reduce real-world accuracy; results not tested live	Use real factory data; include more sensor types; study cost and real-world impact

References

1. Carvalho, T.P.; Soares, F.A.A.M.N.; Vita, R.; Francisco, R.D.P.; Basto, J.P.; Alcalá, S.G.S. A systematic literature review of machine learning methods applied to predictive maintenance. *Comput. Ind. Eng.* **2019**, *137*, 106024. [CrossRef]
2. Çinar, Z.M.; Nuhu, A.A.; Zeeshan, Q.; Korhan, O.; Asmael, M.; Safaei, B. Machine learning in predictive maintenance towards sustainable smart manufacturing in industry 4. *Sustainability* **2020**, *12*, 8211. [CrossRef]
3. Liu, R.; Yang, B.; Zio, E.; Chen, X. Artificial intelligence for fault diagnosis of rotating machinery: A review. *Mech. Syst. Signal Process.* **2018**, *108*, 33–47. [CrossRef]
4. Nacchia, M.; Fruggiero, F.; Lambiase, A.; Bruton, K. A systematic mapping of the advancing use of machine learning techniques for predictive maintenance in the manufacturing sector. *Appl. Sci.* **2021**, *11*, 2546. [CrossRef]
5. Dalzochio, J.; Kunst, R.; Pignaton, E.; Binotto, A.; Sanyal, S.; Favilla, J.; Barbosa, J. Machine learning and reasoning for predictive maintenance in Industry 4.0: Current status and challenges. *Comput. Ind.* **2020**, *123*, 103298. [CrossRef]

6. Chen, C.; Fu, H.; Zheng, Y.; Tao, F.; Liu, Y. The advance of digital twin for predictive maintenance: The role and function of machine learning. *J. Manuf. Syst.* **2023**, *71*, 581–594. [CrossRef]
7. Schwendemann, S.; Amjad, Z.; Sikora, A. A survey of machine-learning techniques for condition monitoring and predictive maintenance of bearings in grinding machines. *Comput. Ind.* **2021**, *125*, 103380. [CrossRef]
8. Theissler, A.; Pérez-Velázquez, J.; Kettelgerdes, M.; Elger, G. Predictive maintenance enabled by machine learning: Use cases and challenges in the automotive industry. *Reliab. Eng. Syst. Saf.* **2021**, *215*, 107864. [CrossRef]
9. Fassi, Y.; Heiries, V.; Boutet, J.; Boisseau, S. Toward Physics-Informed Machine-Learning-Based Predictive Maintenance for Power Converters-A Review. *IEEE Trans. Power Electron.* **2024**, *39*, 2692–2720. [CrossRef]
10. Hoffmann, M.W.; Wildermuth, S.; Gitzel, R.; Boyaci, A.; Gebhardt, J.; Kaul, H.; Amihai, I.; Forg, B.; Suriyah, M.; Leibfried, T.; et al. Integration of novel sensors and machine learning for predictive maintenance in medium voltage switchgear to enable the energy and mobility revolutions. *Sensors* **2020**, *20*, 2099. [CrossRef]
11. Drakaki, M.; Karnavas, Y.L.; Tziafettas, I.A.; Linardos, V.; Tzionas, P. Machine Learning and Deep Learning Based Methods Toward Industry 4.0 Predictive Maintenance in Induction Motors: A State of the Art Survey. *J. Ind. Eng. Manag.* **2022**, *15*, 31–57. [CrossRef]
12. Page, M.J.; Moher, D.; Bossuyt, P.M.; Boutron, I.; Hoffmann, T.C.; Mulrow, C.D.; Shamseer, L.; Tetzlaff, J.M.; Akl, E.A.; Brennan, S.E.; et al. PRISMA 2020 explanation and elaboration: Updated guidance and exemplars for reporting systematic reviews. *BMJ* **2021**, *372*, 160. [CrossRef]
13. Pech, G.; Delgado, C. Assessing the publication impact using citation data from both Scopus and WoS databases: An approach validated in 15 research fields. *Scientometrics* **2020**, *125*, 909–924. [CrossRef]
14. Alryalat, S.A.S.; Malkawi, L.W.; Momani, S.M. Comparing bibliometric analysis using pubmed, scopus, and web of science databases. *J. Vis. Exp.* **2019**, *2019*, 58494. [CrossRef]
15. Barrera Castro, G.P.; Chiappe, A.; Ramírez-Montoya, M.S.; Alcántar Nieblas, C. Key Barriers to Personalized Learning in Times of Artificial Intelligence: A Literature Review. *Appl. Sci.* **2025**, *15*, 3103. [CrossRef]
16. Ayvaz, S.; Alpay, K. Predictive maintenance system for production lines in manufacturing: A machine learning approach using IoT data in real-time. *Expert Syst. Appl.* **2021**, *173*, 114598. [CrossRef]
17. Teoh, Y.K.; Gill, S.S.; Parlikad, A.K. IoT and Fog-Computing-Based Predictive Maintenance Model for Effective Asset Management in Industry 4.0 Using Machine Learning. *IEEE Internet Things J.* **2023**, *10*, 2087–2094. [CrossRef]
18. Rosati, R.; Romeo, L.; Cecchini, G.; Tonetto, F.; Viti, P.; Mancini, A.; Frontoni, E. From knowledge-based to big data analytic model: A novel IoT and machine learning based decision support system for predictive maintenance in Industry 4. *J. Intell. Manuf.* **2023**, *34*, 107–121. [CrossRef]
19. Abidi, M.H.; Mohammed, M.K.; Alkhalefah, H. Predictive Maintenance Planning for Industry 4.0 Using Machine Learning for Sustainable Manufacturing. *Sustainability* **2022**, *14*, 3387. [CrossRef]
20. Ouadah, A.; Zemmouchi-Ghomari, L.; Salhi, N. Selecting an appropriate supervised machine learning algorithm for predictive maintenance. *Int. J. Adv. Manuf. Technol.* **2022**, *119*, 4277–4301. [CrossRef]
21. Welte, R.; Estler, M.; Lucke, D. A method for implementation of machine learning solutions for predictive maintenance in small and medium sized enterprises. In *Procedia CIRP*; Elsevier: Amsterdam, The Netherlands, 2020; pp. 909–914.
22. Jiang, Y.; Dai, P.; Fang, P.; Zhong, R.Y.; Zhao, X.; Cao, X. A2-LSTM for predictive maintenance of industrial equipment based on machine learning. *Comput. Ind. Eng.* **2022**, *172*, 108560. [CrossRef]
23. Putnik, G.D.; Manupati, V.K.; Pabba, S.K.; Varela, L.; Ferreira, F. Semi-Double-loop machine learning based CPS approach for predictive maintenance in manufacturing system based on machine status indications. *CIRP Ann.* **2021**, *70*, 365–368. [CrossRef]
24. Ghasemkhani, B.; Aktas, O.; Birant, D. Balanced K-Star: An Explainable Machine Learning Method for Internet-of-Things-Enabled Predictive Maintenance in Manufacturing. *Machines* **2023**, *11*, 322. [CrossRef]
25. Sharma, R.K. Improving quality of predictive maintenance through machine learning algorithms in industry 4.0 environment. *Proc. Eng. Sci.* **2023**, *5*, 63–72. [CrossRef]
26. Satwaliya, D.S.; Thethi, H.P.; Dhyani, A.; Kiran, G.R.; Al-Taee, M.; Alazzam, M.B. Predictive Maintenance using Machine Learning: A Case Study in Manufacturing Management. In Proceedings of the 2023 3rd International Conference on Advance Computing and Innovative Technologies in Engineering, ICACITE, Greater Noida, India, 12–13 May 2023; pp. 872–876.
27. Panagou, S.; Fruggiero, F.; Lerra, M.; Vecchio, C.D.; Menchetti, F.; Piedimonte, L.; Natale, O.R.; Passariello, S. Feature investigation with Digital Twin for predictive maintenance following a machine learning approach. In Proceedings of the IFAC-PapersOnLine, Hamburg, Germany, 24–27 July 2022; pp. 132–137.
28. Choi, J.S.; Choi, S.W.; Lee, E.B. Modeling of Predictive Maintenance Systems for Laser-Welders in Continuous Galvanizing Lines Based on Machine Learning with Welder Control Data. *Sustainability* **2023**, *15*, 7676. [CrossRef]
29. Calabrese, M.; Cimmino, M.; Fiume, F.; Manfrin, M.; Romeo, L.; Ceccacci, S.; Paolanti, M.; Toscano, G.; Ciandrini, G.; Carrotta, A.; et al. SOPHIA: An event-based IoT and machine learning architecture for predictive maintenance in industry 4. *Information* **2020**, *11*, 202. [CrossRef]

30. Elkateb, S.; Métwalli, A.; Shendy, A.; Abu-Elanien, A.E.B. Machine learning and IoT—Based predictive maintenance approach for industrial applications. *Alex. Eng. J.* **2024**, *88*, 298–309. [CrossRef]
31. Arena, S.; Florian, E.; Zennaro, I.; Orrù, P.F.; Sgarbossa, F. A novel decision support system for managing predictive maintenance strategies based on machine learning approaches. *Saf. Sci.* **2022**, *146*, 105529. [CrossRef]
32. Natanael, D.; Sutanto, H. Machine Learning Application Using Cost-Effective Components for Predictive Maintenance in Industry: A Tube Filling Machine Case Study. *J. Manuf. Mater. Process.* **2022**, *6*, 108. [CrossRef]
33. Mujib, A.; Djatna, T. Ensemble learning for predictive maintenance on wafer stick machine using IoT sensor data. In Proceedings of the 2020 International Conference on Computer Science and Its Application in Agriculture, ICOSICA, Online, 16–17 September 2020.
34. Azab, E.; Nafea, M.; Shihata, L.A.; Mashaly, M. A machine-learning-assisted simulation approach for incorporating predictive maintenance in dynamic flow-shop scheduling. *Appl. Sci.* **2021**, *11*, 11725. [CrossRef]
35. Zamzam, A.H.; Hasikin, K.; Wahab, A.K.A. Integrated failure analysis using machine learning predictive system for smart management of medical equipment maintenance. *Eng. Appl. Artif. Intell.* **2023**, *125*, 106715. [CrossRef]
36. Alshboul, O.; Al Mamlook, R.E.; Shehadeh, A.; Munir, T. Empirical exploration of predictive maintenance in concrete manufacturing: Harnessing machine learning for enhanced equipment reliability in construction project management. *Comput. Ind. Eng.* **2024**, *190*, 110046. [CrossRef]
37. Steurtewagen, B.; Van den Poel, D. Adding interpretability to predictive maintenance by machine learning on sensor data. *Comput. Chem. Eng.* **2021**, *152*, 107381. [CrossRef]
38. Bemani, A.; Björsell, N. Aggregation Strategy on Federated Machine Learning Algorithm for Collaborative Predictive Maintenance. *Sensors* **2022**, *22*, 6252. [CrossRef]
39. Sahasrabudhe, N.; Asegaonkar, R.; Deo, S.; Umredkar, S.; Mundada, K. Experimental analysis of machine learning algorithms used in predictive maintenance. In Proceedings of the 3rd International Conference on Smart Systems and Inventive Technology, ICSSIT, Online, 20–22 August 2020; pp. 1302–1308.
40. Cakir, M.; Guvenc, M.A.; Mistikoglu, S. The experimental application of popular machine learning algorithms on predictive maintenance and the design of IIoT based condition monitoring system. *Comput. Ind. Eng.* **2021**, *151*, 106948. [CrossRef]
41. Khorsheed, R.M.; Beyca, O.F. An integrated machine learning: Utility theory framework for real-time predictive maintenance in pumping systems. *Proc. Inst. Mech. Eng. Part B J. Eng. Manuf.* **2021**, *235*, 887–901. [CrossRef]
42. Achouch, M.; Dimitrova, M.; Dhouib, R.; Ibrahim, H.; Adda, M.; Sattarpanah Karganroudi, S.; Ziane, K.; Aminzadeh, A. Predictive Maintenance and Fault Monitoring Enabled by Machine Learning: Experimental Analysis of a TA-48 Multistage Centrifugal Plant Compressor. *Appl. Sci.* **2023**, *13*, 1790. [CrossRef]
43. Nikfar, M.; Bitencourt, J.; Mykoniatis, K. A Two-Phase Machine Learning Approach for Predictive Maintenance of Low Voltage Industrial Motors. In *Procedia Computer Science*; Elsevier: Amsterdam, The Netherlands, 2022; pp. 111–120.
44. Keartland, S.; Van Zyl, T.L. Automating predictive maintenance using oil analysis and machine learning. In Proceedings of the 2020 International SAUPEC/RobMech/PRASA Conference, SAUPEC/RobMech/PRASA, Cape Town, South Africa, 29–31 January 2020.
45. Mohammed, N.A.; Abdulateef, O.F.; Hamad, A.H. An IoT and Machine Learning-Based Predictive Maintenance System for Electrical Motors. *J. Eur. Des Syst. Autom.* **2023**, *56*, 651–656. [CrossRef]
46. Farooq, U.; Ademola, M.; Shaalan, A. Comparative Analysis of Machine Learning Models for Predictive Maintenance of Ball Bearing Systems. *Electronics* **2024**, *13*, 438. [CrossRef]
47. Singh, G.; Appadurai, J.P.; Perumal, V.; Kavita, K.; Ch Anil Kumar, T.; Prasad, D.; Azhagu Jaisudhan Pazhani, A.; Umamaheswari, K. Machine Learning-Based Modelling and Predictive Maintenance of Turning Operation under Cooling/Lubrication for Manufacturing Systems. *Adv. Mater. Sci. Eng.* **2022**, *2022*, 320. [CrossRef]
48. Guerroum, M.; Zegrari, M.; Elmahjoub, A.A.; Berquedich, M.; Masmoudi, M. Machine Learning for the Predictive Maintenance of a Jaw Crusher in the Mining Industry. In Proceedings of the 2021 IEEE International Conference on Technology Management, Operations and Decisions, ICTMOD, Marrakech, Morocco, 24–26 November 2021.
49. Kammerer, C.; Gaust, M.; Küstner, M.; Starke, P.; Radtke, R.; Jesser, A. Motor classification with machine learning methods for predictive maintenance. In *IFAC-PapersOnLine*; Elsevier: Amsterdam, The Netherlands, 2021; pp. 1059–1064.
50. Khalil, A.F.; Rostam, S. Machine Learning-based Predictive Maintenance for Fault Detection in Rotating Machinery: A Case Study. *Eng. Technol. Appl. Sci. Res.* **2024**, *14*, 13181–13189. [CrossRef]
51. Justus, V.; Kanagachidambaresan, G.R. Machine learning based fault-oriented predictive maintenance in industry 4. *Int. J. Syst. Assur. Eng. Manag.* **2024**, *15*, 462–474. [CrossRef]
52. Achour, A.; Kammoun, M.A.; Hajej, Z. Towards Optimizing Multi-Level Selective Maintenance via Machine Learning Predictive Models. *Appl. Sci.* **2024**, *14*, 313. [CrossRef]

53. Koca, O.; Kaymakci, O.T.; Mercimek, M. Advanced Predictive Maintenance with Machine Learning Failure Estimation in Industrial Packaging Robots. In Proceedings of the 2020 15th International Conference on Development and Application Systems, DAS 2020—Proceedings, Suceava, Romania, 21–23 May 2020; pp. 1–6.
54. Chakroun, A.; Hani, Y.; Elmhamedi, A.; Masmoudi, F. A predictive maintenance model for health assessment of an assembly robot based on machine learning in the context of smart plant. *J. Intell. Manuf.* **2024**, *35*, 3995–4013. [CrossRef]
55. Pulcini, V.; Modoni, G. Machine learning-based digital twin of a conveyor belt for predictive maintenance. *Int. J. Adv. Manuf. Technol.* **2024**, *133*, 6095–6110. [CrossRef]
56. Makridis, G.; Kyriazis, D.; Plitsos, S. Predictive maintenance leveraging machine learning for time-series forecasting in the maritime industry. In Proceedings of the 2020 IEEE 23rd International Conference on Intelligent Transportation Systems, ITSC, Rhodes, Greece, 20–23 September 2020.
57. Kimera, D.; Nangolo, F.N. Predictive maintenance for ballast pumps on ship repair yards via machine learning. *Transp. Eng.* **2020**, *2*, 100020. [CrossRef]
58. Kalathas, I.; Papoutsidakis, M. Predictive maintenance using machine learning and data mining: A pioneer method implemented to greek railways. *Designs* **2021**, *5*, 5. [CrossRef]
59. Putra, H.G.P.; Supangkat, S.H.; Nugraha, I.G.B.B.; Hidayat, F.; Kereta, P.T. Designing Machine Learning Model for Predictive Maintenance of Railway Vehicle. In Proceedings of the 8th International Conference on ICT for Smart Society: Digital Twin for Smart Society, ICISS 2021—Proceeding, Online, 17–19 May 2021.
60. Alvarez Quiñones, L.I.; Lozano-Moncada, C.A.; Bravo Montenegro, D.A. Machine learning for predictive maintenance scheduling of distribution transformers. *J. Qual. Maint. Eng.* **2023**, *29*, 188–202. [CrossRef]
61. Vallim Filho, A.R.A.; Farina Moraes, D.; Bhering de Aguiar Vallim, M.V.; da Silva, L.S.; da Silva, L.A. A Machine Learning Modeling Framework for Predictive Maintenance Based on Equipment Load Cycle: An Application in a Real World Case. *Energies* **2022**, *15*, 3724. [CrossRef]
62. Velasquez, V.; Flores, W. Machine Learning Approach for Predictive Maintenance in Hydroelectric Power Plants. In Proceedings of the 2022 IEEE Biennial Congress of Argentina, ARGENCON, San Juan, Argentina, 7–9 September 2022.
63. Pekşen, M.F.; Yurtsever, U.; Uyaroğlu, Y. Enhancing electrical panel anomaly detection for predictive maintenance with machine learning and IoT. *Alex. Eng. J.* **2024**, *96*, 112–123. [CrossRef]
64. Hsu, J.Y.; Wang, Y.F.; Lin, K.C.; Chen, M.Y.; Hsu, J.H.Y. Wind turbine fault diagnosis and predictive maintenance through statistical process control and machine learning. *IEEE Access* **2020**, *8*, 23427–23439. [CrossRef]
65. Chen, H.; Hsu, J.Y.; Hsieh, J.Y.; Hsu, H.Y.; Chang, C.H.; Lin, Y.J. Predictive maintenance of abnormal wind turbine events by using machine learning based on condition monitoring for anomaly detection. *J. Mech. Sci. Technol.* **2021**, *35*, 5323–5333. [CrossRef]
66. Garan, M.; Tidriri, K.; Kovalenko, I. A Data-Centric Machine Learning Methodology: Application on Predictive Maintenance of Wind Turbines. *Energies* **2022**, *15*, 826. [CrossRef]
67. Gong, L.; Chen, Y. Machine Learning-enhanced loT and Wireless Sensor Networks for predictive analysis and maintenance in wind turbine systems. *Int. J. Intell. Netw.* **2024**, *5*, 133–144. [CrossRef]
68. Granados, D.P.; Ruiz, M.A.O.; Acosta, J.M.; Lara, S.A.G.; Domínguez, R.A.G.; Kañetas, P.J.P. A Wind Turbine Vibration Monitoring System for Predictive Maintenance Based on Machine Learning Methods Developed under Safely Controlled Laboratory Conditions. *Energies* **2023**, *16*, 2290. [CrossRef]
69. Cheng, J.C.P.; Chen, W.; Chen, K.; Wang, Q. Data-driven predictive maintenance planning framework for MEP components based on BIM and IoT using machine learning algorithms. *Autom. Constr.* **2020**, *112*, 103087. [CrossRef]
70. Bouabdallaoui, Y.; Lafhaj, Z.; Yim, P.; Ducoulombier, L.; Bennadji, B. Predictive maintenance in building facilities: A machine learning-based approach. *Sensors* **2021**, *21*, 44. [CrossRef]
71. Fernandes, S.; Antunes, M.; Santiago, A.R.; Barraca, J.P.; Gomes, D.; Aguiar, R.L. Forecasting appliances failures: A machine-learning approach to predictive maintenance. *Information* **2020**, *11*, 208. [CrossRef]
72. Hajimirza Amin, N.; Etemad, A.; Abdalisousan, A. Data-driven performance analysis of an active chilled beam air conditioning system: A machine learning approach for energy efficiency and predictive maintenance. *Results Eng.* **2024**, *23*, 102747. [CrossRef]
73. Al-Aomar, R.; AlTal, M.; Abel, J. A data-driven predictive maintenance model for hospital HVAC system with machine learning. *Build. Res. Inf.* **2024**, *52*, 207–224. [CrossRef]
74. Chazhoor, A.; Mounika, Y.; Vergin Raja Sarobin, M.; Sanjana, M.V.; Yasashvini, R. Predictive Maintenance using Machine Learning Based Classification Models. In Proceedings of the IOP Conference Series: Materials Science and Engineering, Chennai, India, 16–17 September 2020.
75. Pradeep, D.; Vardhan, B.V.; Raiak, S.; Muniraj, I.; Elumalai, K.; Chinnadurai, S. Optimal Predictive Maintenance Technique for Manufacturing Semiconductors using Machine Learning. In Proceedings of the 2023 3rd International Conference on Intelligent Communication and Computational Techniques, ICCT, Jaipur, India, 19–20 January 2023.
76. Morariu, C.; Morariu, O.; Răileanu, S.; Borangiu, T. Machine learning for predictive scheduling and resource allocation in large scale manufacturing systems. *Comput. Ind.* **2020**, *120*, 103244. [CrossRef]

77. Jia, B.; Zhang, X.; Liu, J.; Zhang, Y.; Huang, K.; Liang, Y. Blockchain-Enabled Federated Learning Data Protection Aggregation Scheme With Differential Privacy and Homomorphic Encryption in IIoT. *IEEE Trans. Ind. Inform.* **2022**, *18*, 4049–4058. [CrossRef]
78. Naeem, S.; Ali, A.; Anam, S.; Ahmed, M.M. An Unsupervised Machine Learning Algorithms: Comprehensive Review. *Int. J. Comput. Digit. Syst.* **2023**, *13*, 911–921. [CrossRef] [PubMed]
79. Ding, D.; Han, Q.L.; Xiang, Y.; Ge, X.; Zhang, X.M. A survey on security control and attack detection for industrial cyber-physical systems. *Neurocomputing* **2018**, *275*, 1674–1683. [CrossRef]
80. Al-Abassi, A.; Karimipour, H.; Dehghantanha, A.; Parizi, R.M. An ensemble deep learning-based cyber-attack detection in industrial control system. *IEEE Access* **2020**, *8*, 83965–83973. [CrossRef]
81. Cárdenas, A.A.; Amin, S.; Lin, Z.S.; Huang, Y.L.; Huang, C.Y.; Sastry, S. Attacks against process control systems: Risk assessment, detection, and response. In Proceedings of the 6th International Symposium on Information, Computer and Communications Security, ASIACCS, Hong Kong, China, 22–24 March 2011; pp. 355–366.