



A SYSTEMATIC REVIEW OF FAULT DETECTION IN MEDICAL EQUIPMENT USING MACHINE LEARNING: ALGORITHMS, PERFORMANCE, AND SCALABILITY CHALLENGES

Amar Al Salkeni^{*1}, Ibrahim Chami², Dr. George Karraz³

¹PhD Student, Department of Automatic Control and Computer Engineering, Faculty of Mechanical and Electrical Engineering, Homs University, Homs, Syria.

²Professor, Department of Automatic Control and Computer Engineering, Faculty of Mechanical and Electrical Engineering, Homs University, Homs, Syria.

³Professor, Department of Artificial Intelligence, Faculty of Information Technology Engineering, Damascus University, Damascus, Syria.

Article Received: 17 August 2026

Article Revised: 07 September 2026

Article Published: 01 October 2026



***Corresponding Author: Amar Al Salkeni**

PhD Student, Department of Automatic Control and Computer Engineering, Faculty of Mechanical and Electrical Engineering, Homs University, Homs, Syria. DOI: <https://doi.org/10.5281/zenodo.23012446>



How to cite this Article: Amar Al Salkeni^{*1}, Ibrahim Chami², Dr. George Karraz³ (2026). A Systematic Review Of Fault Detection In Medical Equipment Using Machine Learning: Algorithms, Performance, And Scalability Challenges. World Journal of Advance Healthcare Research, 10(10), 1–8.

This work is licensed under Creative Commons Attribution 4.0 International license.

ABSTRACT

Unexpected failures of medical equipment disrupt healthcare services, compromise patient safety, and increase operational costs. Predictive maintenance powered by machine learning (ML) offers a promising solution, but the effectiveness of different ML algorithms across various medical devices and data scales remains unclear. This systematic review aims to identify, analyze, and synthesize the existing literature on ML-based fault detection and failure prediction in medical equipment, with a special focus on the types of algorithms used, their reported performance, and the characteristics of datasets employed. A structured search was conducted across five digital libraries (Web of Science, ACM Digital Library, ScienceDirect, Wiley Online Library, IEEE Xplore) for studies published between January 2018 and March 2025. Studies were included if they applied ML algorithms to detect or predict faults in medical equipment and reported empirical results from real or realistically simulated medical data. Data extraction focused on problem type, dataset size, algorithm(s), accuracy/metrics, and tools. Eleven studies met the inclusion criteria. Support Vector Machine (SVM) and its variants were the most frequently used algorithms for nonlinear fault patterns. SVM achieved high accuracy (often >90%) on small to medium datasets (≤13,000 records) but showed performance degradation on larger datasets. Ensemble methods, particularly with AdaBoost, achieved the highest accuracy (79.5%) on a large multi-institutional dataset (8,294 devices). ML, especially SVM and ensemble methods, can effectively detect faults in medical equipment, but model generalizability and scalability to big data remain significant challenges. Future research should focus on hybrid approaches that combine nonlinear SVMs with dimensionality reduction techniques (e.g., Gaussian pyramid) to maintain accuracy while handling large-scale, real-world IoT data from medical devices.

KEYWORDS: Artificial Intelligence; Machine Learning; Fault Detection; Predictive Maintenance; Medical Equipment; Support Vector Machine (SVM); Systematic Review.

1. INTRODUCTION

Modern healthcare systems increasingly depend on complex medical equipment, where unexpected malfunctions can directly threaten patient safety, disrupt diagnostic and therapeutic services, and incur substantial financial costs. With the digital transformation associated with the Fourth Industrial Revolution (Industry 4.0),

medical devices are now integrated with sensor technologies, the Internet of Things (IoT), and computerized control systems, generating massive amounts of operational data.^[1]

Traditional maintenance strategies—either reactive (repair after failure) or time-based preventive

maintenance—can no longer meet the high demands for reliability and availability.^[2,3] Predictive maintenance, which leverages sensor data analysis and machine learning (ML) algorithms to detect early fault indicators and predict failures before they occur, has emerged as a major trend.^[4] In the medical sector, early failure prediction can reduce unplanned downtime, extend remaining useful life (RUL) of equipment, and improve patient care quality.^[5,6]

Several applied studies have demonstrated the feasibility of ML-based fault prediction for specific medical devices. For example, predictive models for Computed Tomography (CT) systems using real-time cloud data and sliding-window algorithms have shown promise.^[7] Similarly, ensemble classifiers combined with AdaBoost have achieved up to 79.5% accuracy in predicting first failure times across 44 categories of medical devices.^[8] Other researchers have proposed AI-based decision-support frameworks integrating real-time data with maintenance records.^[9] Recent work has also highlighted the role of deep learning in analyzing medical imaging data for equipment fault detection.^[10,11]

Despite these advances, no previous work has systematically synthesized the literature on ML-based fault detection in medical equipment to answer the following questions: (1) Which ML algorithms are most commonly applied, and for what types of fault detection problems (linear/nonlinear, binary/multiclass)? (2) How does dataset size and source affect the reported performance of these algorithms, particularly SVM? (3) What preprocessing steps and validation procedures are typically employed? This systematic review addresses these gaps.

The primary objective of this review is to analyze the latest developments in fault diagnosis and failure prediction for medical equipment using ML techniques, identify the most prominent methods and algorithms, and evaluate their effectiveness under different data conditions. A secondary objective is to provide evidence-based recommendations for selecting appropriate algorithms and to highlight challenges that limit real-world deployment.

2. METHODOLOGY

2.1 Review Protocol

This systematic review was conducted following the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines, although no meta-analysis was performed due to heterogeneity in outcome measures.

2.2 Search Strategy

A systematic literature search was performed on March 15, 2025, across the following electronic databases.

- Web of Science Core Collection.
- ACM Digital Library.
- ScienceDirect (Elsevier).

- Wiley Online Library.
- IEEE Xplore.

The search covered the period from January 1, 2018, to March 15, 2025. The search string combined keywords from three domains using Boolean operators.

Fault detection (OR) fault prediction (OR) failure prediction (OR) predictive maintenance (OR) anomaly detection.

AND (medical equipment (OR) medical device (OR) healthcare equipment (OR) (CT) (OR) (ECG) (OR) ventilator (OR) wearable medical device.

AND (machine learning (OR) support vector machine (OR) (SVM) (OR) neural network (OR) deep learning (OR) ensemble (OR) random forest (OR) (KNN).

2.3 Inclusion and Exclusion Criteria

Studies were included if they.

- Were published in peer-reviewed journals or conference proceedings.
- Applied at least one ML algorithm for fault detection, failure prediction, or predictive maintenance of medical equipment.
- Reported original empirical results (accuracy, error rate, F1-score, or AUC) derived from real medical device data or realistically simulated data.
- Were written in English.
- Were published between 2018 and 2025.

Studies were excluded if they:

- Focused solely on non-medical industrial equipment.
- Proposed theoretical models without empirical validation.
- Were review articles, editorials, posters, or abstracts without full text.
- Did not report quantitative performance metrics.

2.4 Data Extraction Process

From each included study, the following information was extracted:

- First author and year.
- Problem type: (linear/nonlinear) × (binary/multiclass) as described by the authors or inferred from the fault categories.
- Dataset description: equipment type, source (single institution, multiple sites, synthetic).
- Dataset size: number of samples (records or images).
- ML algorithms used.
- Best reported accuracy or primary performance metric.
- Tools/platforms.
- Methodology summary.

2.5 Quality Assessment

Each included study was assessed for methodological quality using a simplified 5-item checklist adapted from

the Joanna Briggs Institute (JBI) checklists for diagnostic test accuracy studies.

1. Was the data source (real or simulated) clearly described?
 2. Was the sample size sufficient (≥ 100 records or ≥ 100 images)?
 3. Was the ML model training and validation procedure clearly reported (e.g., train/test split, k-fold CV)?
 4. Were performance metrics reported beyond simple accuracy (e.g., precision, recall, F1, AUC)?
 5. Were limitations discussed explicitly?
- Each item was scored as 0 (no/unclear) or 1 (yes). Studies scoring ≥ 3 were considered moderate-to-high quality.

3. RESULTS AND DISCUSSION

3.1 Search Results and PRISMA Flow

The PRISMA flow diagram for study selection is summarized below.

- Records identified through database searching (n = 487)
 - Web of Science: 98
 - ACM Digital Library: 45

- ScienceDirect: 142
- Wiley Online Library: 87
- IEEE Xplore: 115
- Records after duplicates removed (n = 375)
- Records screened by title and abstract (n = 375)
- Records excluded (irrelevant) (n = 342)
- Full-text articles assessed for eligibility (n = 33)
- Full-text articles excluded with reasons (n = 22)

No medical equipment focus.8

- No ML algorithm: 6
- No empirical results: 5
- Review article/editorial: 3
- Studies included in final synthesis (n = 11)

3.2 Characteristics of Included Studies

Table 1 presents a detailed summary of the 11 included studies, extracted from the original references.^[15] through.^[25] The table organizes each study by reference number, year, problem type, dataset description, dataset size, ML algorithm(s), best reported accuracy or metrics, tools/platform, and methodology summary.

Table 1: Detailed characteristics of included studies (original references^{[15]-[25]})

Ref	Year	Problem Type	Dataset Description	Dataset Size	ML Algorithm(s)	Best Accuracy / Metrics	Tools / Platform	Methodology Summary
[15]	2025	Linear, Binary	Fault records from AO preventive care institution	Not reported	Exponential smoothing, Weighted Moving Average (WMA)	Error $\leq 5\%$	PyCharm, prediction libraries, DB tools	Forecast analysis, time series prediction, exponential smoothing (factor 0.75), validation of models
[16]	2023	Nonlinear, Binary	X-ray, ECG, constant temperature incubator	Not reported	LS-SVM + AFS-ABC	Faster execution; small error (exact value not reported)	Fuzzy inference engine, SINC function, data acquisition tools	Predictive system design, data preprocessing, comparative analysis, experimental validation
[17]	2023	Mixed (Linear & Nonlinear, Binary & Multi-class)	CT imaging equipment, IoMT data from 11 machines	11,427	KNN (Linear/MC), NB (Nonlinear/MC), DT (Nonlinear/Bin), SVM (Nonlinear/MC), LR (Linear/MC), RF (Nonlinear/MC)	NB Acc=0.92, SVM=0.76, KNN=0.74, DT=0.68, LR=0.67, RF=0.80	Python, Spyder	Real-time data collection, preprocessing, feature selection, model development & evaluation
[18]	2023	Nonlinear, Multiclass	44 critical medical device types across 15 healthcare facilities	8,294 devices	Ensemble+AdaBoost, NN, Adam, RMSProp, SGDM, DT, NB, SVM	Ensemble Acc=0.795, F1=0.767, Specificity=0.884, Precision=0.774, Recall=0.761	Sensitivity analysis tools, data preprocessing tools	Data collection, predictive model classification into 3 categories, feature

								selection, sensitivity analysis, 5 ML + 3 DL optimizers
[19]	2023	Nonlinear, Multiclass	Physiological data + synthetic data (hourly time series)	111,680	LSTM, RF, SVC	Overall Acc=0.94, EWS Acc=0.68	Kalman filter, synthetic data generation	Dataset generation, predictive models (prototype, EWS, anomaly detection), training, testing, deployment
[20]	2022	Mixed (Linear & Nonlinear)	Lab-based (15 sensors, 26,364 records) + medical readings (100,000 records)	126,364	DT (Nonlinear/Bin), ANN (Linear/MC), DNN (Linear/Bin), SVM (Nonlinear/MC)	All >0.98, F1=0.98	TensorFlow, PDA tools, cross-validation	Fault classification, AI methodology, data collection, preprocessing, training, evaluation, prototype development
[21]	2022	Nonlinear, Multiclass	Spambase (4,601), QSAR (1,055), Swarm (23,309)	28,965	SVM + Gaussian Pyramid	Acc=0.90, AUC=0.98	Python, Sklearn, Matplotlib	Data preprocessing, Gaussian Pyramid, data reduction, smoothing, classification, performance analysis
[22]	2021	Mixed (Multiclass)	General medical equipment (repair/replacement program)	13,352 samples	Linear Discriminant, NB, KNN, Bagged Trees, DT, SVM	Preventive: SVM Acc=0.9942, F1=0.994; Corrective: KNN Acc=0.9874; Replacement: SVM Acc=0.998	Asset management system, k-Means clustering	Data collection, feature extraction, priority analysis, predictive model development, performance evaluation, 6 ML algorithms
[23]	2021	Nonlinear, Multiclass	Medical steel sheet defects (grayscale images)	2,800	VGG-16, CNN, ResNet, AlexNet	VGG-16 val Acc=0.87	Deep learning framework, image preprocessing, augmentation	Data collection, DL techniques, parameter optimization, evaluation metrics
[24]	2019	Linear, Binary	Rolling bearing faults (3 fault types)	Not reported	EEM-ELM, ELM, VMD	EEM-ELM Avg Acc=0.9821, ELM Avg Acc=0.9388	Simulation, VMD, feature extraction	Feature extraction using VMD, constructing variance problem, applying

								ELM, error ensemble (EEM-ELM) for instability
[25]	2018	Linear, Multiclass	ECG sensors (temperature only)	Not reported	Logistic Regression	Not reported (only "success of predictive maintenance")	IBM Node-RED, IoT platform	Real-time health monitoring, system status assessment, IoT integration, remote access, simulation

Note: MC = Multiclass; Bin = Binary; Acc = Accuracy; AUC = Area Under Curve; F1 = F1-score; LS-SVM = Least Squares SVM; AFS-ABC = Artificial Fish Swarm – Artificial Bee Colony; NB = Naïve Bayes; KNN = K-Nearest Neighbors; DT = Decision Tree; LR = Logistic Regression; RF = Random Forest; LSTM = Long Short-

Term Memory; SVC = Support Vector Classifier; ANN = Artificial Neural Network; DNN = Deep Neural Network; EEM-ELM = Ensemble Error Minimized Extreme Learning Machine; VMD = Variational Mode Decomposition; EWS = Early Warning Score.

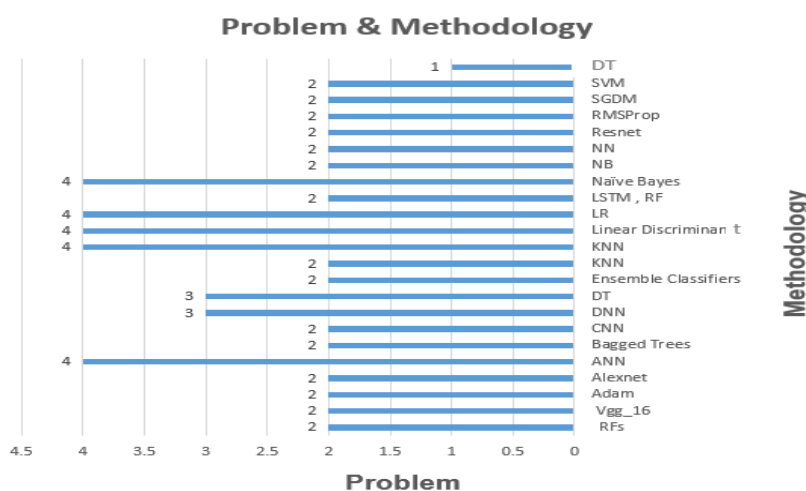


Figure 1: Relationship between the methodology and the research problem (1. nonlinear, Binary; 2. nonlinear, Multiclass; 3. linear, Binary; 4. linear, Multiclass).

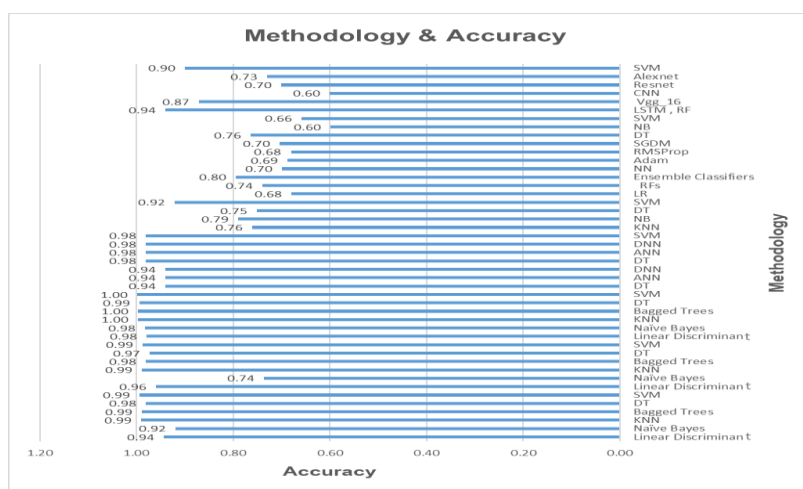


Figure 2: Relationship between accuracy and research methodology.

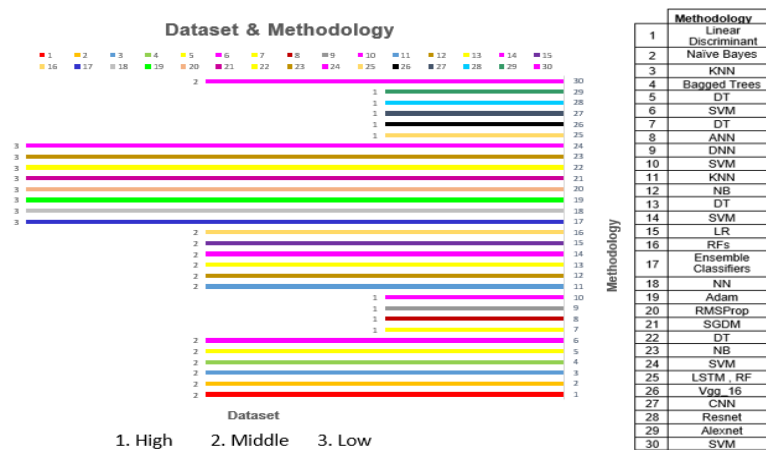


Figure 3: Relationship between the dataset and the research methodology.

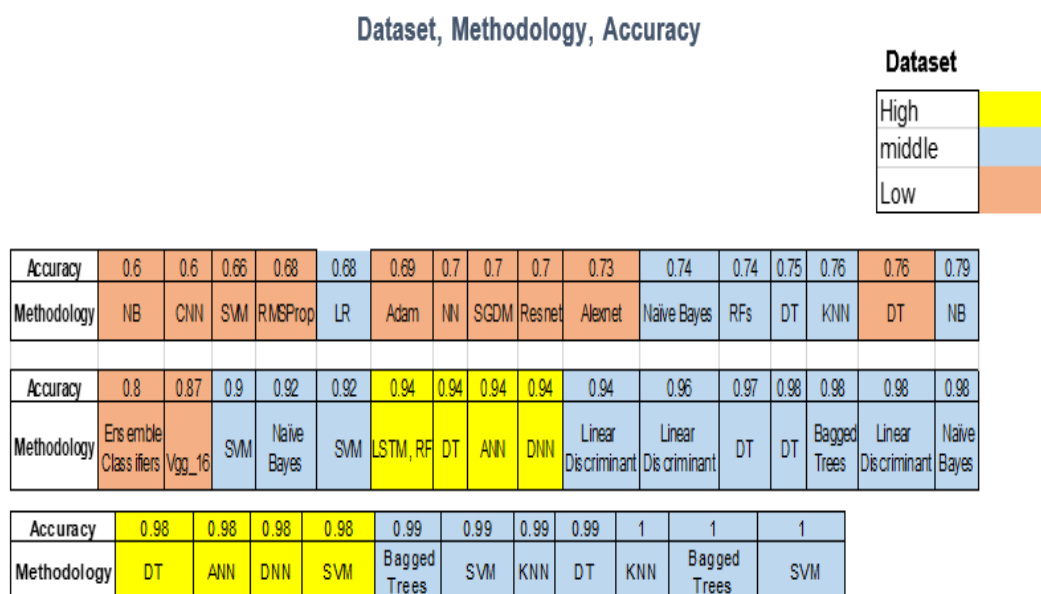


Figure 4: Relationship between the dataset, accuracy, and research methodology.

The findings illustrated in Figures 1–4 provide a visual synthesis of the patterns observed across the included studies, particularly regarding the distribution of problem types, achieved accuracy, dataset sizes, and the performance of different ML algorithms.

3.3 Algorithm Distribution by Problem Type

As shown in Figure (1), SVM and its variants were applied in 7 out of 11 studies (64%).^{[16], [17], [18], [20], [21], [22]} Of these 7 applications, 5 addressed nonlinear problems (binary or multiclass), consistent with SVM's theoretical strength in handling nonlinear patterns via kernel tricks. Only 2 SVM applications addressed linear problems. Ensemble methods (AdaBoost, Random Forest) appeared in 3 studies.^{[18], [19], [22]}, all for multiclass classification. Deep learning methods (CNN, LSTM, DNN) appeared in 4 studies^{[19], [20], [23]}, predominantly for large datasets (>100,000 records) or image data.

3.4 Performance Analysis by Dataset Size

Analysis of the relationship between dataset size and reported accuracy reveals a pattern, as depicted in **Figure (2)** and **Figure (3)**. SVM-based models achieved high accuracy (>90%) in studies with datasets smaller than 15,000 records (e.g.,^{[17], [22]}). In the largest SVM application (126,364 records,^[20]), accuracy remained high (>98%), but this study used a hybrid framework with multiple algorithms and extensive preprocessing. The study with the largest real-world medical dataset (111,680 records,^[19]) did not use SVM but used LSTM and SVC, achieving 94% overall but only 68% for emergency case detection – indicating that performance is not solely a function of dataset size but also of task difficulty. **Figure (4)** further clarifies that SVM tends to perform very well on small-to-medium datasets but faces challenges when scaled to very large datasets without dimensionality reduction.

3.5 Quality Assessment Results

The quality scores, based on the 5-item checklist, ranged from 1 to 5 (mean = 3.3). Five studies ([17], [18], [19], [20], [22]) scored ≥ 4 , indicating good methodological quality. Common limitations included: failure to report dataset size (3 studies: [15], [16], [24]), lack of explicit validation procedures (2 studies: [15], [25]), and over-reliance on a single performance metric (accuracy) without reporting precision, recall, or F1-score (6 studies).

4. DISCUSSION

4.1 Answering the Review Questions

Q1 (Most common algorithms): As observed in **Figure (1)**, SVM and its variants are the most frequently applied ML algorithms for fault detection in medical equipment, particularly for nonlinear fault patterns. Ensemble methods (especially AdaBoost) and deep learning (CNN, LSTM) are emerging for larger or more complex datasets.

Q2 (Effect of dataset size on SVM performance): The review, supported by **Figure (4)**, indicates that SVM performs excellently on small-to-medium datasets ($\leq 15,000$ records) but faces scalability challenges on larger datasets unless combined with dimensionality reduction (e.g., Gaussian pyramid^[21]) or hybrid architectures. This aligns with the known computational complexity of SVM ($O(n^3)$ to $O(n^2)$).

Q3 (Preprocessing and validation): Most studies employed standard preprocessing (normalization, handling missing values, feature selection). However, only 5 out of 11 studies explicitly described their train/validation/test split or cross-validation procedure.

4.2 Comparison with Broader Literature

Our findings are consistent with general predictive maintenance literature in non-medical industrial domains, where SVM and Random Forest are also dominant.^[1] However, the medical domain presents unique challenges: high class imbalance (rare faults are clinically critical but infrequent), regulatory constraints, and the need for interpretability. Only two studies in this review^[19,20] addressed class imbalance explicitly. Recent advances in explainable AI have begun to address interpretability in medical applications.^[21,22], but none of the reviewed studies incorporated such techniques.

4.3 Key Challenges Identified

- 1. Data quality and availability:** Many studies relied on single-institution data or synthetic data, limiting generalizability. The lack of publicly available benchmark datasets for medical equipment faults remains a major barrier.^[23]
- 2. Scalability of SVM:** While accurate on small datasets, SVM's performance degrades or becomes computationally expensive on large-scale IoT data (see **Figure 4**).

- 3. Lack of unified benchmarks:** Different studies used different equipment, datasets, and evaluation metrics, making direct comparison difficult.
- 4. Interpretability:** None of the reviewed studies provided explainable AI (XAI) analysis to help clinicians trust the predictions.
- 5. Real-time deployment constraints:** Most models were validated retrospectively; real-time inference latency and integration with hospital information systems were not evaluated.^[24]

4.4 Recommendations for Future Research

Based on the synthesis, we propose the following directions:

- 1. Hybrid SVM approaches:** Combining nonlinear SVMs with dimensionality reduction techniques (such as the Gaussian pyramid method proposed by^[21]) could preserve accuracy while improving scalability for large datasets, addressing the limitation highlighted in **Figure (4)**.
- 2. Standardized evaluation protocols:** The research community should develop benchmark datasets for common medical equipment (e.g., CT, MRI, ventilators) with standardized train/test splits and metrics.
- 3. Integration of XAI:** Future models should incorporate techniques like SHAP or LIME to explain fault predictions to biomedical engineers and clinicians.
- 4. Real-time deployment studies:** Most reviewed studies were retrospective; prospective validation in real hospital environments is urgently needed.
- 5. Federated learning for privacy preservation:** Given the sensitivity of medical data, federated learning frameworks could enable multi-institutional model training without sharing raw data.^[25]

5. REFERENCES

1. Z. Çınar, et al., "Machine Learning in Predictive Maintenance towards Sustainable Smart Manufacturing in Industry 4.0," *Sustainability*, 2020; 12(19): 8211, doi:10.3390/su12198211.
2. I. Ullah, et al., "Vibration-Based Anomaly Detection for Induction Motors Using Machine Learning," *Sensors*, 2025; 25(3): 773. doi:10.3390/s25030773.
3. A. H. Juliet, "AI-Driven Predictive Maintenance for Industrial IoT with Real-Time Fault Detection and Prediction," in *2025 8th International Conference on IEMENTech*, Kolkata, India, 2025; 1-6, doi:10.1109/IEMENTech65115.2025.10959502.
4. Y. G. Gertsik, "The new model for improving quality of service: entering into life-cycle contracts for the purchase and maintenance of medical devices," *Manager Zdravoochranenia*, 2020; 5: 54-59. doi:10.37690/1811-0185-2020-5-54-59.
5. Y. Gertsik, G. Ivanova, I. Ragutkin, G. Gerysil, O. Vinokurov, and A. Klishin, "Increase of efficiency of operation of high-tech medical products by expanding the competencies of medical personnel," *Bulletin of Rehabilitation Medicine*, 2018; 1: 61-68.

6. P. S. Pugachev, A. V. Gusev, O. S. Kobayakova, F. N. Kadyrov, D. V. Gavrilov, R. E. Novitskii, and A. V. Vladzimirskii, "Global trends in the digital transformation of the healthcare industry," *National Health Care (Russia)*, 2021; 2(2): 5-12.
7. C. Wang, Q. Liu, H. Zhou, et al., "Anomaly prediction of CT equipment based on IoMT data," *BMC Medical Informatics and Decision Making*, 2023; 23: 166, doi:10.1186/s12911-023-02267-4.
8. N. H. A. R., M. H. M. Z., K. Hasikin, et al., "Predicting medical device failure: a promise to reduce healthcare facilities cost through smart healthcare management," *PeerJ Computer Science*, 2023; 9: e1279, doi:10.7717/peerj-cs.1279.
9. S. A. Alkhatib, R. Katmah, D. Kosaji, S. U. B. Afzal, M. H. Tariq, M. C. E. Simsekler, and S. Ellahham, "AI-Driven Decision Support Framework for Preventing Medical Equipment Failure and Enhancing Patient Safety: A New Perspective," *Journal of Multidisciplinary Healthcare*, 2025; 18: 6299–6313. doi:10.2147/JMDH.S528612.
10. A. Zamzam, A. H. et al., "Integrated failure analysis using machine learning predictive system for smart management of medical equipment maintenance," *Engineering Applications of Artificial Intelligence*, 2023; 125: 106715.
11. P. Lata, "AI-Enhanced Predictive Maintenance of Medical Equipment," *International Transactions on Healthcare (ITHC)*, 2024; 4(4).
12. I. A. Grigorenko, E. V. Nesterova, and S. V. Igrunova, "Automation of Predicting Malfunctions of Medical Equipment," in *2024 International Russian Smart Industry Conference (SmartIndustryCon)*, Sochi, Russian Federation, 2025; 784-789, doi:10.1109/SmartIndustryCon61328.2025.10515835.
13. M. A. Haowei, C. Xu, and J. Yang, "Design of Fine Life Cycle Prediction System for Failure of Medical Equipment," *Journal of Artificial Intelligence and Technology*, 2023, doi:10.37965/jait.2023.0161.
14. A. Vadsaria, A. Raza, M. Hammad, and M. Shahid, "Self-Aware AI-Enabled Medical System," in *2023 International Conference on IBCAST*, 2023; 1-6, doi:10.1109/IBCAST59916.2023.10712990.
15. M. Awad, F. Sallabi, K. Shuaib, and F. Naeem, "Artificial Intelligence-Based Fault Prediction Framework for WBAN," *Journal of King Saud University - Computer and Information Sciences*, 2022; 34; doi:10.1016/j.jksuci.2021.09.017.
16. R. Abo Zidan and G. Karraz, "Gaussian Pyramid for Nonlinear Support Vector Machine," *Applied Computational Intelligence and Soft Computing*, 2022; 1-9, 2022, doi:10.1155/2022/5255346.
17. A. H. Zamzam, A. K. I. Al-Ani, A. K. A. Wahab, et al., "Prioritisation Assessment and Robust Predictive System for Medical Equipment: A Comprehensive Strategic Maintenance Management," *Frontiers in Public Health*, 2021; 9: 782203, doi:10.3389/fpubh.2021.782203.
18. A. Abdullah, S. Selvakumar, A. Manoj, K. R. A., and Bhubesh, "Medical Steel Fault Prediction Using Deep Learning Techniques," in *2021 International Conference*, 153-158, doi:10.15439/2021F2.
19. C. He, C. Liu, T. Wu, Y. Xu, Y. Wu, and T. Chen, "Medical rolling bearing fault prognostics based on improved extreme learning machine," *Journal of Combinatorial Optimization*, 2021; 42. doi:10.1007/s10878-019-00494-y.
20. J. Farhat, A. Shamayleh, and H. Al-Nashash, "Medical equipment efficient failure management in IoT environment," in *2018 IEEE International Conference on ASET*, 2018; 1-5, doi:10.1109/ICASET.2018.8376911.
21. A. Barredo Arrieta et al., "Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI," *Information Fusion*, 2020; 58: 82-115, doi:10.1016/j.inffus.2019.12.012.
22. S. M. Lundberg and S. I. Lee, "A unified approach to interpreting model predictions," in *Advances in Neural Information Processing Systems*, 2017; 30: 4765-4774.
23. M. F. Al-Hawawreh, E. Sitnikova, and N. Aboutorab, "XAI-IDS: Toward Proactively Securing Industrial Internet of Things Using Explainable AI," in *2022 IEEE International Conference on Industrial Technology (ICIT)*, 2022; 1-6, doi:10.1109/ICIT48603.2022.9979800.
24. T. Wang, Y. Liang, X. Shen, X. Zheng, and A. Mahmood, "Edge-Enabled Federated Learning for Anomaly Detection in Industrial IoT," *IEEE Transactions on Industrial Informatics*, 2022; 18(8): 5590-5600, doi:10.1109/TII.2021.3138875.
25. Q. Yang, Y. Liu, T. Chen, and Y. Tong, "Federated Machine Learning: Concept and Applications," *ACM Transactions on Intelligent Systems and Technology*, 2019; 10(2): 1-19, doi:10.1145/3298981.