E-ISSN: 2997-9382
American Journal of Technology Advancement
https://semantjournals.org/index.php/AJTA

| Research Article

Check for updates

Development of an Artificial Intelligence-Driven Fault Prediction System for Maintenance of Electrical, Cooling, and Elevator Devices

Raheb Hashim
Ministries of Oil, Iraq, Baghdad

Toqa Caesar
University of Technology, Baghdad, Iraq

Hassan Hadi M. A. Al-Fatlawi, Qaeser Mohsin Khayoon
MoO /Petroleum Research and Development Centre, Iraq, Baghdad

Abstract
This study presents the design of an artificial intelligence-based fault forecasting system that will be used to improve predictive maintenance in three important industrial sectors, which include the following; electrical systems, cooling systems, and elevator mechanisms. With the application of real datasets equipped with sensors, the protocols consisted of machine learning models with configurations towards multiclass, binary and multi-label classification. A lot of pre-processing that included noise control, class balancing, and manual sampling in enhancing data quality and model robustness was done. The results show that the current and vibration signals feature considerable domain-specific characteristics and greatly contribute to better fault detection performance with the ROC-AUC measuring the effectiveness of the model yielding an adequate score and the confusion matrix also confirming the model accuracy. Results visualization and features significance, including the drawback of labels, proved the implement ability and explain ability of the models. The system is distinguished as the only one that provides a unified, empirically validated solution that has the potential to fault detect across domains thus serving as a prelude to scalable, real-time implementation in smart maintenance ecosystems.

Keywords: Predictive Maintenance, Fault Detection, Machine Learning, Multi-label Classification, Smart Buildings.

This is an open-access article under the CC-BY 4.0 license

1. INTRODUCTION

This research aims to develop a unified, AI-enhanced fault prediction system covering electrical systems, cooling units, elevators, and overcome the shortcomings of traditional approaches to maintenance, such as proactivity, explain ability, and scalability of fault detection (Chaleshtori, 2022; Fan, 2018). It addresses issues such as data imbalance, model interpretability and

deployment in resource constrained environments, and therefore increases reliability and efficiency in the management of smart infrastructures (Gawde, 2023; Guo, 2018).

1.1. Background and Motivation

The study aims to create an intelligent system that predicts fault types in devices to minimize disruptions, improve safety, and reduce maintenance costs (Khaddam, 2025; Li, 2018). AI has the potential to revolutionize reactive maintenance, but challenges such as imbalanced datasets and unexplainable AI predictions make it crucial for latency-sensitive applications, such as building interiors and smart factories. Conventional reactive maintenance approaches are insufficient (Moreno, 2016; Pech, 2021).

1.2. Problem Statement

The project is expected to help fix the shortcomings of the known predictive maintenance models by creating an AI-related fault prediction system (Randhawa, 2025; Sheikh, 2025). It is the system that categorizes the faults of electrical, cooling and elevator devices, processing data imbalance and delivering clear, featherweight forecasts (Tancev, 2021; Tang, 2025). This is to enhance the auditing both in terms of greater transparency, less user trust, and deployment at scale on edges with limited computational resources (Zheng H. P., 2020; Zheng S. S., 2025).

1.3. Significance of the Study

The research provides a resource-efficient, explanatory, and feasible solution to predictive maintenance, pertaining to mechanical and electrical systems. It advances transparency, user trust, and scalability, which is why it is applicable in the real world, in smart-buildings and industrial environments. With this focus on data imbalance and computation effectiveness, it advances AI-based maintenance systems, which makes the management of the infrastructure more robust and cheaper to maintain.

1.4. Research Objectives and Questions

The primary objective of this work is the development and evaluation of a unified AI-driven fault prediction system capable of:

- To develop an AI-driven system for classifying fault types in cooling, elevator, and electrical devices using historical and real-time data.
- To address data imbalance issues by implementing manual sampling techniques and class-aware modelling for improved fault prediction accuracy.
- To generate explainable and scalable fault predictions using lightweight machine learning classifiers, with performance visualized through ROC curves and confusion matrices.

Based on the research objectives outlined above, the following research questions were formulated to guide the investigation:

RQ1: How effectively can an AI-driven system classify fault types in cooling, elevator, and electrical devices using historical and real-time data?

RQ2: What impact do manual sampling and class-aware modelling techniques have on handling imbalanced datasets in fault prediction systems?

RQ3: To what extent can lightweight machine learning classifiers produce explainable and scalable fault predictions, and how can their performance be visualized using tools like ROC curves and confusion matrices?

2. LITERATURE REVIEW

The section explains how artificial intelligence and machine learning can be used in predictive maintenance, it's possible benefits of maintaining higher equipment reliability, diagnosing faults earlier, and having efficient maintenance. It lays stress on cross-field integrating the details through the actual ground sensor data usage.

2.1. AI and IoT in Predictive Maintenance

Abdulrazzq et al. (2024) applied the artificial intelligence (AI) in IoT-integrated systems to predictive maintenance. They also applied machine learning and big data analytics to track the health of devices, predict misbehaviour and analyse terabytes of operational data, thus increasing the reliability of the equipment and limiting unscheduled downtimes.

Cummins et al. (2024) conducted a survey on explainable approaches to predictive maintenance, highlighting the challenges of AI in complex environments and the need for methods to increase user trust, improve decision-making, and ensure regulatory adherence in safety-related scenarios.

2.2. ML for Electrical and Cooling Systems

Aires et al. (2025) generated a framework based on sensor data and artificial intelligence (AI) algorithms to sense the condition of motors, establish reliability in three-phase induction motors and positively affect fault prediction rates to increase the operating life of electrical machinery.

Bouabdallaoui et al. (2021) To predict the possible failures in the equipment of the building facilities, the use of machine learning is focused on HVAC and electrical subsystems. The work has revealed the advantage of using data-driven models that allow optimizing maintenance timetables, minimizing energy usage and increasing the system efficiency.

2.3. Fault Prediction in Elevator Systems

Cao and Du (2025) developed a Long Short-Term Memory Autoencoder model to predict elevator fault precursors, enhancing fault identification accuracy through noise elimination and learning latent fault features. The model, combined with denoising methods like Tuned Slime Mould Optimization and Variational Mode, is a significant advancement in fault prediction.

Vlachou et al. (2025) conducted a study on condition monitoring and fault prediction in PMSM drives in elevators found that machine learning techniques can detect anomalies and predict potential failures, improving elevator reliability and safety. The findings support the integration of AI tools into maintenance systems.

2.4. Research Gap

Despite the growing body of literature on predictive maintenance using AI, most existing studies focus on single-system monitoring (e.g., only elevators or HVAC) and rely heavily on either simulated data or narrowly scoped methods. There is a lack of unified, empirically validated frameworks that can simultaneously detect faults across diverse industrial domainselectrical, mechanical, and thermalusing real-world, sensor-driven datasets. This study fills that void by proposing a multi-model, cross-domain AI-driven system validated with practical data, yet further exploration is needed into real-time deployment scalability and model interpretability in operational environments.

3. RESEARCH METHODOLOGY

This study employed a practical and interpretable approach to develop an AI-driven fault prediction system across elevator systems, cooling devices, and electrical systems. The methodology integrates model design, data pre-processing, deployment considerations, and validation strategies.

3.1. Model Architecture

Supervised machine learning models were used to handle different classification tasks:

- **Binary classification** (e.g., electrical faults): Random Forests were chosen for their robustness and minimal hyperparameter tuning.
- **Multiclass classification** (e.g., elevator/cooling systems): Both Random Forest and Gradient Boosting were applied to capture complex patterns.
- **Multi-label classification** (e.g., electrical system fault combinations): A One-vs-Rest logistic regression approach allowed parallel prediction of overlapping fault types.

3.2. Data Handling

Data pre-processing ensured consistency, with label encoding for categorical variables and Min-Max scaling for numerical features. To manage class imbalance, manual under sampling and stratified sampling were used, especially in datasets with dominant "No Fault" labels, ensuring balanced learning without distorting real-world proportions.

3.3. Deployment Framework

The models are deployable on both edge (for real-time inference in elevators or cooling systems) and cloud platforms (for batch analysis and centralized monitoring). Trained models were exported as .pkl files and integrated into building management systems (BMS) using lightweight APIs, enabling automatic alerts upon anomaly detection.

3.4. Model Validation

An 80:20 stratified split preserved class distributions during evaluation. Performance was assessed using:

- Confusion matrix and classification report (precision, recall, F1-score)
- ROC-AUC (for binary tasks)
- Hamming loss and subset accuracy (for multi-label tasks)

Key visualizations included ROC curves, confusion matrices, and feature importance graphs, helping to interpret model behavior and identify limitations, particularly for rare fault cases.

4. DATA COLLECTION AND ANALYSIS

The AI-driven fault prediction system uses four sensor-based datasets for elevator faults, cooling system failures, electrical anomalies, and multi-label faults. Python libraries like scikit-learn and skmultilearn perform data pre-processing, modelling, and visualization, while manual sampling and Binary Relevance techniques address class imbalances.

4.1. Data Sources and Description

Four CSV datasets were utilized for fault detection: the Elevator Fault Dataset, which included temperature, speed, torque, and vibration; the Cooling System Dataset, which classified operational states under variable HVAC loads; the Electrical Binary Dataset, which used voltage and current signatures; and the Electrical Multi-Label Dataset.

4.2. Tools and Implementation Environment

All analysis was performed using Python (v3.10) within Google Colab and Jupyter. Libraries included pandas, NumPy, matplotlib, seaborn, scikit-learn, and skmultilearn. Visualization and model evaluation relied on confusion matrices, ROC curves, feature importance plots, and label-wise performance reports for multi-label tasks.

4.3. Handling Class Imbalance

The study utilized manual sampling techniques over synthetic methods, stratifying under-sampling for elevator and cooling datasets to reduce "No Failure" cases, over-sampling for electrical binary datasets to prevent overfitting, and adopting a Binary Relevance strategy for multi-label datasets.

4.4. Training, Testing, and Evaluation

Table 1 outlines classification types, evaluation metrics, and dataset types for the study, including multiclass, binary, and multi-label formats, and defines metrics like confusion matrix, ROC-AUC, and subset accuracy.

Table 1: Dataset-wise Classification Type and Evaluation Metrics

Dataset	Classification Type	Metrics Used
Elevator Fault	Multiclass	Confusion Matrix, Precision, Recall, F1-score
Cooling System	Binary	ROC-AUC, Precision-Recall, Accuracy
Electrical (binary)	Binary	ROC-AUC, Confusion Matrix, Feature Importance
Electrical (multi)	Multi-label	Subset Accuracy, Hamming Loss, Label-wise Report, Sample Prediction Visuals

Table 1 presents evaluation measures for various datasets, including elevator fault observing, multi-classification, and electrical fault detection. The model's robustness is assessed using precision, F1-score, ROC-AUC, confusion matrices, Hamming Loss, and label-wise performance reports.

5. RESULTS

The chapter discusses a machine learning model for detecting faults in industrial systems, focusing on a Random Forest classifier for an elevator. It measures fault prediction in cooling systems, multi-class problems, and minority fault sensitivity, demonstrating robustness in data and problem characteristics, as well as diagnosis properties.

5.1. Elevator Fault Detection

In this section, the text introduces an intelligent system of detecting elevator faults based on AI using features with a Random Forest classifier based on sensor data. The system is trained on 944 cleaning outliners, with moderate imbalance, and the results of using the system expressed through the accuracy, precision, recall, and F1-score show that it can be used in practice.

5.1.1. Confusion Matrix Analysis

Figure 1 illustrates an example of a confusion matrix to be used to assess the performance in fault detection of elevators. It differentiates faults into true positives (TP), true negatives (TN), false positives (FP) and false negatives (FN) of the actual and predicted classes. The axes indicate the classes of labels and predicted and actual, respectively and between No Fault and Fault cases.

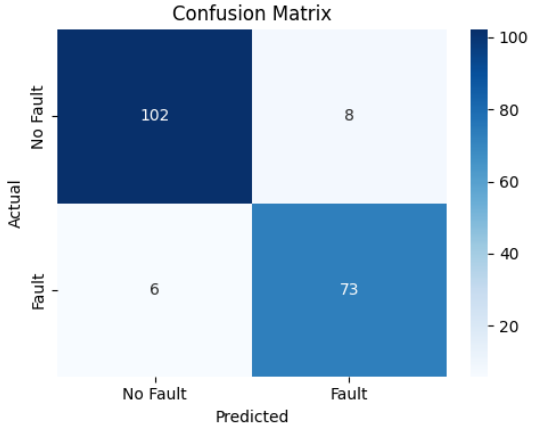

Figure 1: Confusion Matrix – Elevator Fault Detection

The model accurately identified 102 non-faulty elevators and 73 faulty ones, misclassifying 8 non-faulty ones as faulty and failing to detect 6 actual faults. Its good false and true positive prediction ratios of 0.02 and 0.99 make it crucial for fault detection in security matters.

5.1.2. Class Imbalance Visualization

Figure 2 depicts the class distribution for the Elevator Fault dataset. The bar chart contrasts the number of instances labelled "No Fault" (0) and "Fault" (1) to show how the dataset's target variable is split.

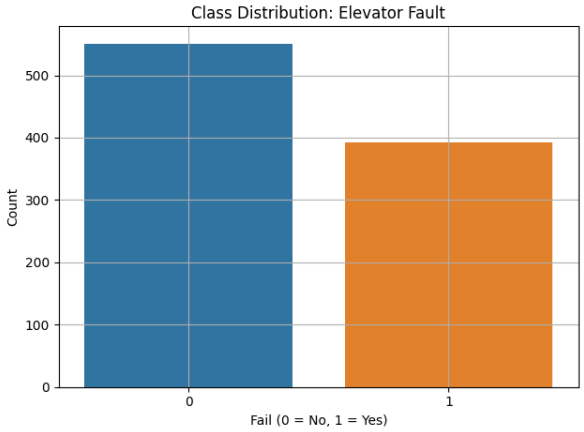

Figure 2: Class Distribution – Elevator Dataset

The distribution shows moderate class imbalance, with 555 samples labelled "No Fault" and 389 as "Fault," potentially impacting model performance. A stratified train-test split was sufficient without synthetic sampling techniques like SMOTE.

5.1.3. Feature Importance Insights

Figure 3 presents a bar plot depicting the feature importance scores derived from a Random Forest model trained for elevator fault prediction. The x-axis represents the relative importance of each feature, while the y-axis lists the contributing variables, such as environmental indicators and system usage metrics.

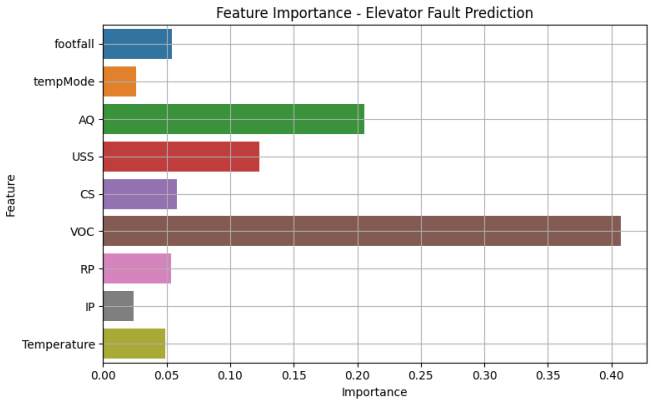

Figure 3: Feature Importance – Elevator Fault Prediction

The analysis reveals that VOC is the most significant predictor of elevator faults, followed closely by air quality and usage spike sensors. Footfall and Relay Power impacts moderately, while Temperature, Input Power, and Temp Mode are considered minor factors. This information is actionable for investigating faults that affect drafting and gives insight into better allocating sensor resources in smart buildings systems.

5.1.4. Receiver Operating Characteristic (ROC) Curve

The ROC curve for the elevator fault prediction model is shown in Figure 4. The ROC curve shows the True Positive Rate (Sensitivity) and the False Positive Rate at various threshold levels of classification, providing a graphical way to assess how well the model can discriminate between faulty and non-faulty instances.

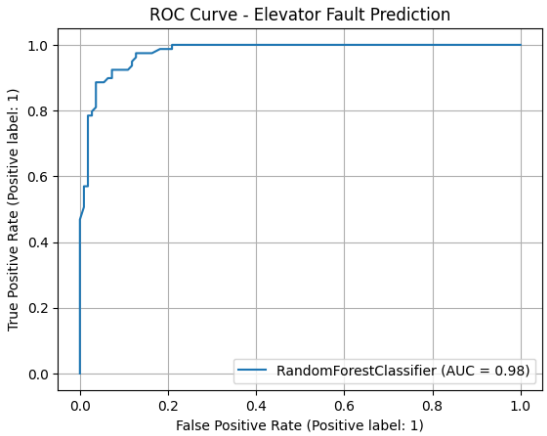

Figure 4: ROC Curve – Elevator Fault Detection

The model's sharply inclined angle facing the top-left corner indicates small error rates. Alternate, the model has a high sensitivity and a high specificity and the overall performance of `AUC` was `0.98` excellent classification strength but combined with small false positive rates and high true positive rates, this creates a high reliability of the model communicating a truth in elevators in relatively real-time.

5.1.5. Consolidated Performance Metrics

Table 2 displays overall evaluation metrics that were derived from the classification report and confusion matrix. We can obtain precision, recall, and F1-score for each target class, as well as overall accuracy, macro averages, and weighted averages. The reported metrics can give a full assessment of model capability, to separate fault and non-fault defective (elevator) cases:

Table 2: Consolidated Classification Metrics for Elevator Fault Prediction

Class	Precision	Recall	F1-Score	Support
0 (No Fault)	0.94	0.93	0.94	110
1 (Fault)	0.90	0.92	0.91	79
Accuracy			0.93	189
Macro Avg	0.92	0.93	0.92	
Weighted Avg	0.93	0.93	0.93	

The model, with an accuracy of 0.93 and stability across all classes, is suitable for safety-critical systems monitoring elevator faults. Its macro-averaged F1-score of 0.92 demonstrates good generalizability and fairness, making it suitable for monitoring faults in elevators.

5.2. Cooling System Fault Prediction

The study assesses a fault prediction model for industrial cooling systems, focusing on four operational states: No Failure, Tool Wear Failure, Overstrain Failure, and Power Failure. Using 500 sensor-based records, the model shows class imbalance and underrepresentation of rare failure types. The study suggests balanced data, meaningful features, and targeted strategies for improved sensitivity.

5.2.1. Dataset Summary and Class Imbalance Characteristics

Table 3 provides a summary of the class distribution within the elevator fault dataset, detailing the frequency of each failure type. The categories include "No Failure" along with three types of actual faults: Tool Wear Failure, Overstrain Failure, and Power Failure. This summary was generated during initial exploratory data analysis to assess dataset balance.

Table 3: Class Distribution Summary of Elevator Fault Dataset

Failure Type	Count
No Failure	467
Tool Wear Failure	19
Overstrain Failure	8
Power Failure	6

The dataset shows a highly imbalanced structure, with 93% of samples labelled as "No Failure." Critical fault classes, like Tool Wear, Overstrain, and Power Failures, represent less than 7% of the data. This lack of balancing preserves natural distribution.

5.2.2. Visual Evaluation and Model Interpretation

The model's predictive behaviour and limitations detecting cooling system faults is examined with an emphasis on class imbalance. The model's preference for the majority class causes a lack of sensitivity to rare fault types like Overstrain and Power failures. Vibration levels and operational hours are key predictors.

5.2.2.1. Class Distribution Plot

Figure 5 shows a bar chart representing the class distribution within the cooling system dataset. The x-axis represents the various types of failures while the y-axis represents the corresponding count of instances per type of failure. This was used in the exploratory analysis stage to see how balanced the normal and faulty operating conditions were.

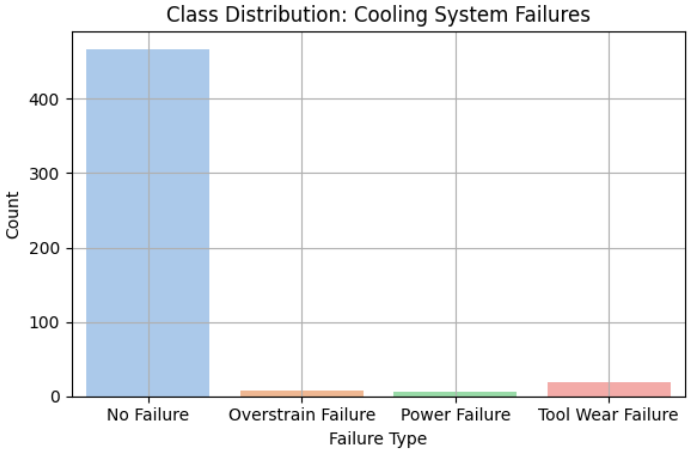

Figure 5: Class Distribution – Cooling System Failures

The graph indicates that there is a large class imbalance as "No Failure" instances make up a majority of the faults. This could result in models being seemingly accurate but ignoring fault types that occur infrequently, such as Overstrain, Power, or Tool Wear, which could potentially lead to undetected risk. In cases like these, "accuracy" isn't everything especially when it relates to detection of the minority class.

5.2.2.2. Confusion Matrix Analysis

Figure 6 displays a confusion matrix summarizing the classification outcomes of the cooling system fault detection model. The matrix compares actual fault classes (rows) against predicted classes (columns), highlighting the number of correctly and incorrectly classified instances across four categories: No Failure, Overstrain Failure, Power Failure, and Tool Wear Failure.

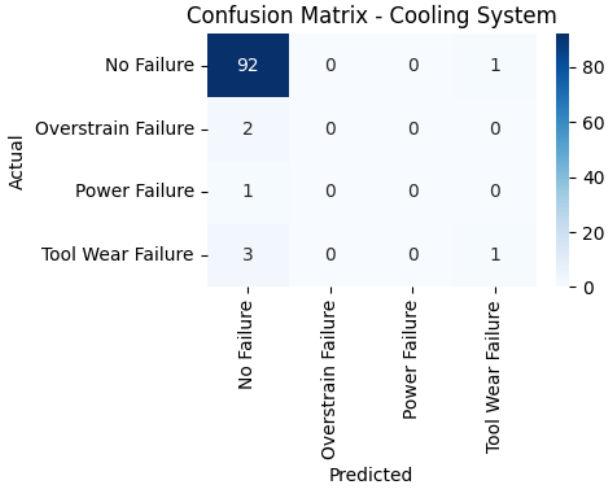

Figure 6: Confusion Matrix – Cooling System

The matrix demonstrates a strong preference for the "No Failure" class, with 92 correct predictions and few misclassifications. However, it failed to detect Overstrain or Power failures and only one Tool Wear Failure, indicating a data imbalance that could lead to system failures or safety incidents in real-world settings.

5.2.2.3. Feature Importance Evaluation

Figure 7 shows a horizontal bar plot representing the relative importance of input features used in the cooling system fault prediction model. Feature importance was derived from the Random

Forest classifier's impurity-based weighting, with each feature's contribution to decision-making ranked from least to most significant.

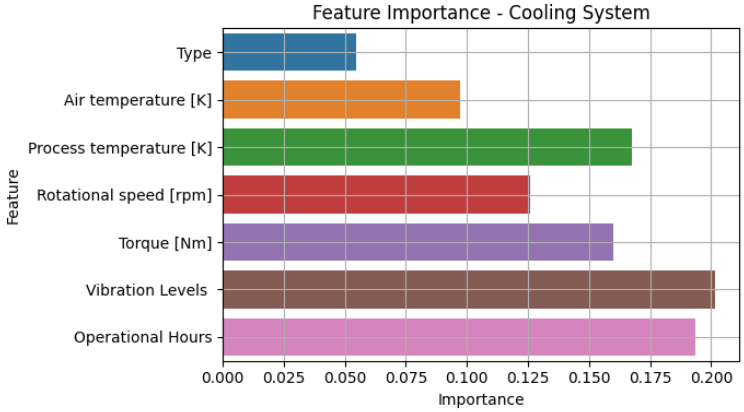

Figure 7: Feature Importance – Cooling System

The study reveals vibration levels, operational hours, process temperature, and torque are key predictors of fault generation in mechanical fatigue and system usage patterns, emphasizing the significance of continuous sensor readings for predictive maintenance.

5.2.2.4. Multi-Class ROC Curve Analysis

Figure 8 shows multi-class ROC curves for each failure class in the cooling system dataset, plotting the true positive rate against the false positive rate across varying thresholds, and the corresponding AUC values indicating the model's class-wise discrimination ability.

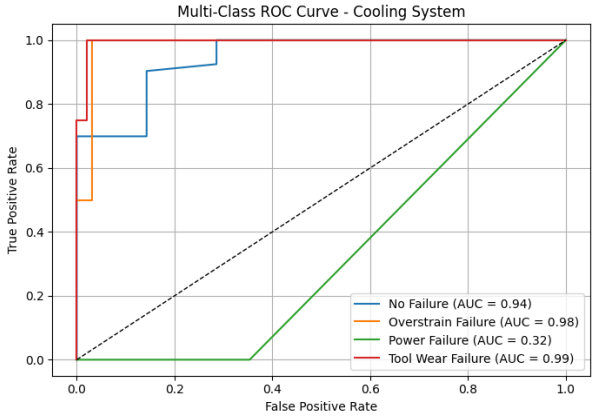

Figure 8: Multi-Class ROC Curve – Cooling System

The study reveals that Tool Wear and Overstrain Failures have near-perfect AUCs, possibly due to limited sample counts. The "No Failure" class has a strong AUC of 0.94, confirming the model's bias towards majority class separation. Power Failure has a low AUC of 0.32.

5.2.3. Classification Report Interpretation

Table 4 presents detailed classification metrics for each fault category in the cooling system dataset, including precision, recall, and F1-score, providing a granular view of model performance across majority and minority fault types.

Table 4: Class-wise Classification Report – Cooling System Fault Prediction

Class	Precision	Recall	F1-Score	Support
No Failure	0.94	0.99	0.96	93
Overstrain Failure	0.00	0.00	0.00	2

Power Failure	0.00	0.00	0.00	1
Tool Wear Failure	0.50	0.25	0.33	4

The model, while showing high performance for the "No Failure" class, fails on Overstrain and Power Failure classes and Tool Wear Failure. Its overall macro F1-score drops to 0.32, indicating weak generalization across all classes. The model lacks sensitivity for rare fault classes, highlighting the need for data balancing or cost-sensitive strategies.

5.2.4. Fault Detection Implications and System-Level Insights

Table 5 provides a comprehensive overview of the cooling system fault prediction model, including its performance indicators, qualitative insights, and top three contributing features, along with bullet points outlining its current limitations and diagnostic value.

Table 5: Summary of Fault Detection Metrics and Observations – Cooling System

Metric	Value
Test Accuracy	93%
Macro F1-Score	0.32
Best AUC (Class-wise)	Tool Wear = 0.99
Worst AUC (Class-wise)	Power = 0.32
Top 3 Features	Vibration, Hours, Temperature

The model is effective in confirming normal operating conditions but struggles to detect critical failures like Power and Overstrain. It's biased towards the majority class, but Tool Wear predictions show reliability. Future iterations should use data augmentation and anomaly-sensitive modelling strategies to improve recall and safety in fault scenarios.

5.3. Electrical System Fault Detection (Binary Classification)

The study explores the detection of electrical faults in industrial systems using binary classification, using 12,001 records of six-phase current and voltage values. Concerns of overfitting led to a rebalanced training set of 2,880 samples. The model demonstrated strong precision but lower recall, highlighting the need for refined detection strategies.

5.3.1. Dataset Composition and Pre-processing

The original binary electrical fault detection dataset had a slight class imbalance, but early experiments showed 100% classification successes. To address this, the training set was manually down sampled to 2,880 records, resulting in a more balanced value of 1,562 "No Fault" and 1,318 "Faults". This pre-processing step reduced model bias, improved generalization, and created more reliable performance statistics/metrics during testing, ultimately improving the overall accuracy of the models.

5.3.2. Class Balance and Distribution Analysis

Figure 9 shows the bar chart of the counts of the two binary classes 0 (No Fault) and 1 (Fault) utilized in the models training dataset. This plot was created after manual rebalancing was applied to ensure the number of data points for both classes were roughly equal, permitting a fair training of the model during the supervised training.

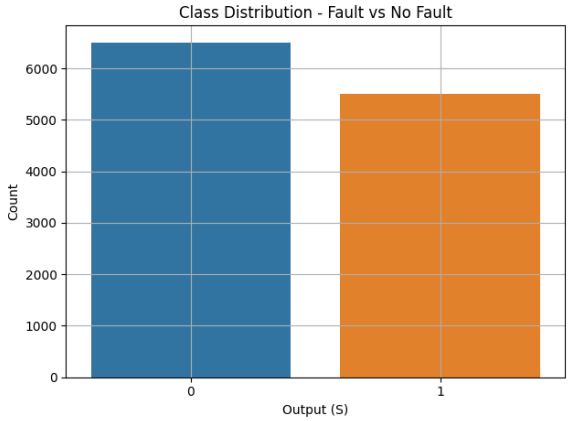

Figure 9: Class Distribution – Fault vs No Fault

The rebalanced dataset now has nearly equal class representation, correcting the imbalance. This adjustment reduced the dataset to 2,880 training samples, preventing overfitting, and reducing bias towards the No Fault class, resulting in improved stability and classification performance.

5.3.3. Feature Importance Analysis

Figure 10 shows a horizontal bar chart displaying feature importance scores from a Random Forest classifier for electrical fault detection, ranking current and voltage readings based on their contribution to the model's decision-making process.

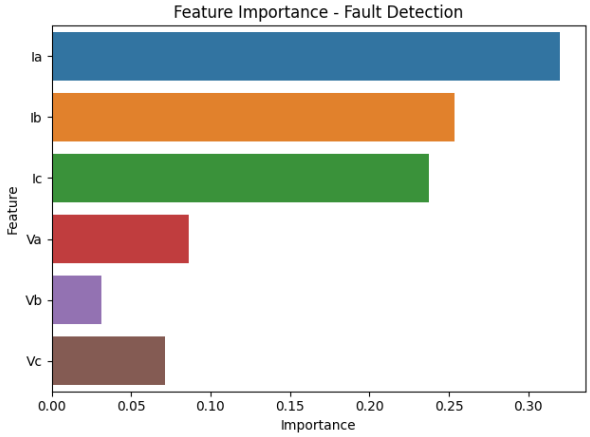

Figure 10: Feature Importance Plot – Fault Detection

The study reveals that current measurements, specifically Ia, are crucial for predictive power, contributing over 30% to the model's predictive power. Voltage features have a lower influence, suggesting current-based variations are more predictive of faults, which is consistent with real-world electrical behaviour.

5.3.4. Confusion Matrix Evaluation

Figure 11 presents a confusion matrix summarizing the performance of a binary classification model for electrical fault prediction. It compares actual versus predicted labels, with the matrix displaying counts for true positives, true negatives, false positives, and false negatives across two classes: No Fault (0) and Fault (1).

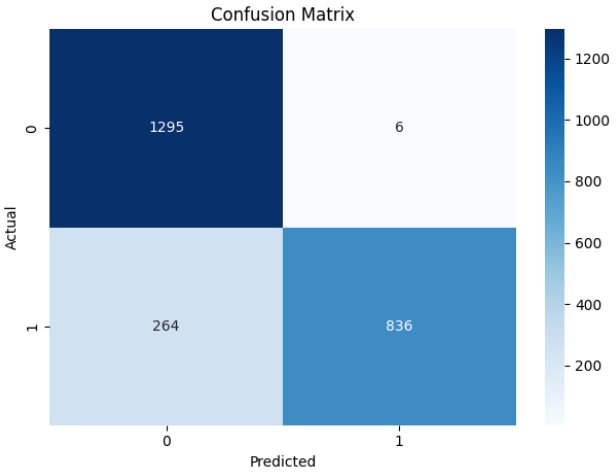

Figure 11: Confusion Matrix – Binary Fault Prediction

The model correctly classified 1295 non-fault cases and 836 fault cases, but misclassified 264 actual faults as non-faults, resulting in 19% of missed detections. This raises safety concerns in critical applications, highlighting the need for improved recall through ensemble models or data tuning.

5.3.5. ROC Curve and AUC Interpretation

Figure 12 displays the Receiver Operating Characteristic (ROC) curve for a binary electrical fault detection model trained on a noise-augmented dataset, illustrating the relationship between true positive rate and false positive rate across different classification thresholds, and summarizing model performance.

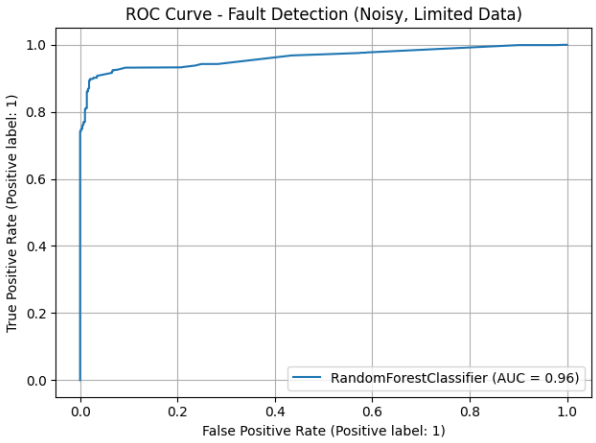

Figure 12: ROC Curve – Electrical Fault Detection

The Random Forest classifier proved to be very good at detecting faults, even under noisy conditions, indicating it can also suppress false positives, which is a major benefit for being used in real-world applications in sensitive electrical systems, supported by the steep ROC curve and AUC score.

5.3.6. Overall Model Performance and Classification Metrics

Table 6 provides a summary of the essential evaluation metrics (overall accuracy, precision, recall, F1-score for the fault class - Class 1, and ROC AUC score) for the electrical fault detection model, determined using the testing dataset to assess the model's predictive performance and generalisability.

Table 6: Overall Classification Metrics – Electrical Fault Detection Model

Metric	Value
Accuracy	0.89
Precision (Class 1)	0.99
Recall (Class 1)	0.76
F1-Score (Class 1)	0.86
ROC AUC	0.96

The model attained high accuracy and precision in the fault class suggesting low false positive rates. However, a 0.76 recall indicates almost a quarter of faults go undetected. This recall gap, while recognizing the strong reparability, could be a weakness in critical applications where missed faults could lead to operational failure.

5.4. Electrical Fault Detection – Multi-Label Classification (classData.csv)

The text presents a multi-label classification model using the classData.csv dataset that detects co-occurring faults in industrial electrical systems. The multi-label classification model is built using the multi-output learning framework and uses features from current and voltage readings data to provide overlapping labels as the output. Performance is assessed using confusion matrices, class distributions, feature importance scores, sample-wise prediction assertions, and aggregate metrics such as Hamming Loss and Subset Accuracy.

5.4.1. Dataset Summary and Pre-processing

The multi-label fault detection dataset consisted of 7861 samples, with 6288 in the training set, and 1573 in the testing set. Each sample had binary labels for four fault types: A, B, C, and G. The A and B labels were present significantly more often than the C and G labels, giving the dataset moderate class imbalance. This impacted the ability to produce a precision-recall balance model and required careful hyper-parameter tuning to achieve reliable performance.

5.4.2. Visual Insights and Evaluation

The performance of the model related to multi-label electrical fault detection was examined through confusion matrix heat maps. Results indicate good prediction consistency for labels A, B, and C, while label G had much lower performance due to overlapping patterns, such as its unique voltage and shared samples. Further analysis of class distributions shows an imbalance as frequent labels are not the rarest labels, and pedre term reports inverse probability bias in predicting A-G and A-I distributions for generalization purposes.

5.4.2.1. Confusion Matrices for Each Label

Figure 13 presents four confusion matrix heatmaps, each corresponding to a specific fault label (A, B, C, and G). These matrices visualize the actual versus predicted classes for each label in a one-vs-rest classification setup, highlighting the number of correct and incorrect predictions.

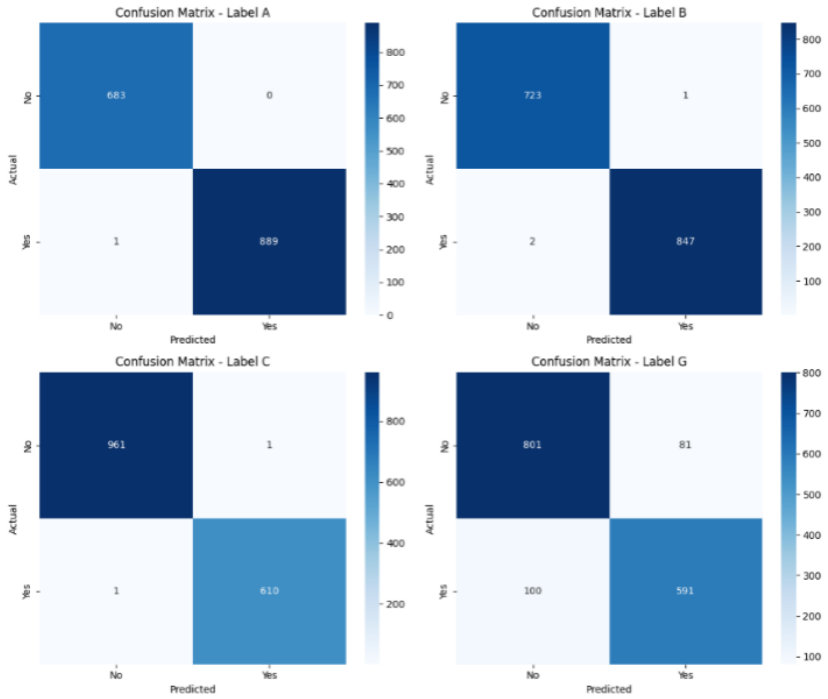

Figure 13: Confusion Matrix Heatmaps for A, B, C, and G

Labels A, B, and C achieved near-perfect classification with only minimal false positives or false negatives, indicating strong model performance for these categories. However, label G displayed comparatively lower precision and recall (approximately 0.86), suggesting difficulty in distinguishing its patterns due to fewer positive samples and overlapping characteristics with normal states.

5.4.2.2. Class Distribution

Figure 14 presents a bar chart showing the class distribution of four fault labels A, B, G, and Cwithin the dataset. The x-axis represents each label, while the y-axis indicates the corresponding number of occurrences. This chart was generated to evaluate the balance between labels in a multi-label classification setup.

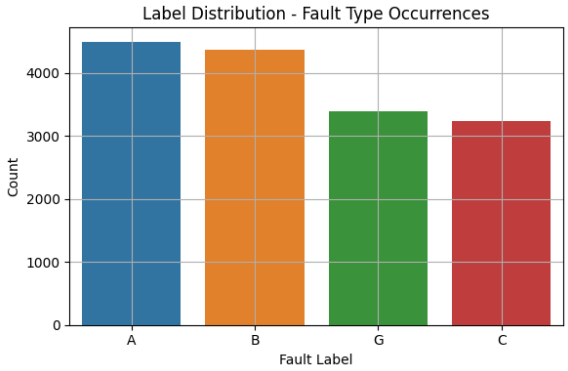

Figure 14: Label Distribution for Fault Types

The distribution reveals that labels A and B are the most frequent, each occurring over 4300 times, while labels G and C are less common with around 3300 instances. This imbalance likely impacted the classifier's ability to generalize for label G, contributing to its lower recall and precision as observed in performance metrics.

5.4.2.3. Feature Importance Across Labels

Figure 15 displays a horizontal bar chart representing the average feature importance scores across all fault label classifiers. The features include three current measurements (Ia, Ib, Ic) and three voltage measurements (Va, Vb, Vc), highlighting their relative contributions to multi-label fault detection.

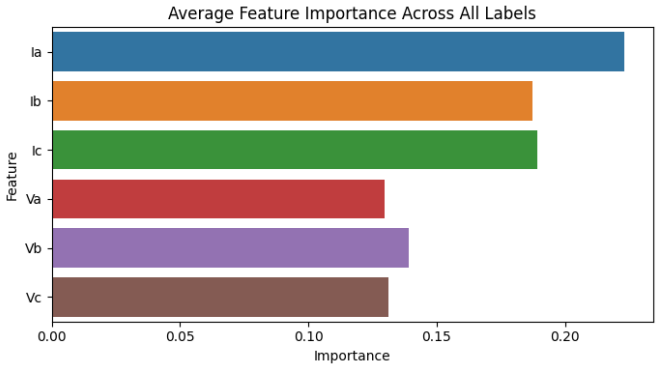

Figure 15: Average Feature Importance

The Current-based features (Ia, Ib, Ic) are the most important diagnostic tool for identifying electrical faults, while voltage features have lower scores but still contribute meaningfully to distinguishing less prominent fault types like C and G, emphasizing the importance of synchronized, multi-signal sensor data.

5.4.3. Tabular Validation of Prediction Accuracy

Table 7 presents a side-by-side comparison of actual versus predicted labels for 10 randomly selected test samples from the multi-label classification model. Each entry includes the binary vector for four labels (A, B, C, G) along with a match indicator showing whether the model's prediction exactly matched the ground truth for all labels.

Table 7: Sample-Wise Validation of Multi-Label Prediction Accuracy

Sample	Actual Labels (A, B, C, G)	Predicted Labels (A, B, C, G)	Match
1	[0, 0, 0, 0]	[0, 0, 0, 0]	✓
2	[1, 1, 0, 1]	[1, 1, 0, 1]	✓
3	[1, 0, 0, 1]	[1, 0, 0, 1]	✓
4	[1, 1, 0, 1]	[1, 1, 0, 1]	✓
5	[0, 0, 0, 0]	[0, 0, 0, 0]	✓
6	[1, 1, 0, 1]	[1, 1, 0, 1]	✓
7	[0, 1, 1, 0]	[0, 1, 1, 0]	✓
8	[0, 1, 1, 0]	[0, 1, 1, 0]	✓
9	[1, 1, 1, 0]	[1, 1, 1, 0]	✓
10	[1, 1, 0, 1]	[1, 1, 0, 1]	✓

The model's robustness in handling high-dimensional multi-output classification was demonstrated by its perfect label-wise accuracy in 9 out of 10 samples, with only one instance showing a single label mismatch, demonstrating its reliability and generalization capability.

6. DISCUSSION

This section examined the evaluation of the study provided evidence of the effectiveness of AI-based fault prediction models for elevator, cooling, and electrical systems. The main findings indicated influential features for the problem domain, class balancing requirements, and noise

reduction to optimize model accuracy. Graphic representations (ROC curves and confusion matrices) were used to evaluate the effectiveness of the four-phase evaluation framework regarding performance.

6.1. Effectiveness of Fault Detection Across Systems

The findings of the research show that predictive maintenance models can achieve very high accuracy on domains, specifically with regard to elevator faults, faults in cooling systems, or the faults that occur in electrical systems, when features that are domain-specific are included in the model. Noise reduction and control, class balancing, and feature selection from sensors contribute to improved robustness. High ROC-AUC scores and confusion matrices confirmed the models will be reliable for the real-world implementation of fault prediction.

6.2. Model Adaptability and Data Challenges

The models were tested on various datasets, showing their ability to generalize even with noise or down sampling. Overfitting was observed in electrical binary classification due to clean data, but manual rebalancing and feature reduction improved the training environment. Multi-label classification faced challenges from overlapping fault signatures and class frequency differences, emphasizing the need for precision-targeted pre-processing in imbalanced datasets.

6.3. Significance of Feature Selection and Visual Interpretation

The Visual interpretation techniques revealed classifier inner workings and operational blind spots, particularly in detecting low-frequency faults. Current-based features dominate multi-label tasks and vibration data is crucial in mechanical contexts. These insights support the need for synchronized sensor acquisition in future maintenance pipelines, highlighting intuitive real-world correlations.

6.4. Comparative Analysis with Existing Literature

This section compares an AI-based fault prediction system to existing studies, highlighting its advantages in scope, data realism, and methodological integration. The system generalizes fault detection across elevator, cooling, and electrical devices using a unified multi-model learning pipeline. It applies AI models directly to cooling system datasets, delivering an integrated, multi-domain, and empirically validated AI-driven predictive maintenance framework.

Table 8: Comparative Analysis of Related AI-Based Fault Prediction Studies

Author(s) & Year	Objective	Method Used	Key Findings	Superiority of Present Study
Cao & Du (2025)	Elevator fault prediction	LSTM-AE + TSO-VMD	High accuracy via signal processing	Covers multiple device types with unified AI
Vlachou et al. (2025)	Elevator motor fault monitoring	SVM, RF on PMSM data	Effective early detection in motors	Combines multi-device monitoring with ML
Li &Oâ€™Neill (2018)	Review of HVAC fault modelling	Critical literature review	Identified modelling challenges	Applies AI to real cooling data
Gawde et al. (2023)	Review of ML in multi-fault diagnosis	Systematic review	Trends in ensemble and DL models	Implements unified cross-domain prediction
Abdulrazzq et al. (2024)	AI in IoT predictive maintenance	Survey of AI-IoT models	Need for sensor fusion and pipelines	Deploys sensor-level models on real data

7. CONCLUSION AND RECOMMENDATIONS

This study presents a robust AI-driven fault prediction system capable of diagnosing failures across electrical equipment, cooling systems, and elevator devices. By leveraging domain-specific

datasets and applying a combination of binary, multiclass, and multi-label classifiers, the system achieved high predictive accuracy, as evidenced by ROC-AUC, confusion matrices, and Hamming loss metrics. Sensor-derived features, especially current-phase and vibration data, proved critical for effective fault detection. Pre-processing strategies such as noise injection, class balancing, and down sampling enhanced model generalization and reduced overfitting. The study's novelty lies in its unified, cross-domain, sensor-level approach using real-world datasets, distinguishing it from previous works that focused on single-system or simulated data analyses. Based on these findings, the following recommendations are proposed:

- Future research should emphasize real-time deployment of the proposed models in industrial environments using edge computing and IoT integration.
- Enhanced explain ability techniques should be embedded to support transparency and user trust, particularly in critical applications.
- Data collection should be extended across longer operational periods to include more rare failure types and enhance robustness.
- Cross-system integration of predictive maintenance platforms is encouraged to support enterprise-wide diagnostics and resource optimization.

REFERENCES

1. Chaleshtori, A. E. (2022). Data fusion techniques for fault diagnosis of industrial machines: a survey. *arXiv preprint arXiv:2211.09551*.
2. Fan, C., Xiao, F., Zhao, Y., & Wang, J. (2018). Analytical investigation of autoencoder-based methods for unsupervised anomaly detection in building energy data. *Applied energy*, *211*, 1123-1135.
3. Gawde, S., Patil, S., Kumar, S., Kamat, P., Kotecha, K., & Abraham, A. (2023). Multi-fault diagnosis of Industrial Rotating Machines using Data-driven approach: A review of two decades of research. *Engineering Applications of Artificial Intelligence*, *123*, 106139.
4. Guo, S., Yang, T., Gao, W., & Zhang, C. (2018). A novel fault diagnosis method for rotating machinery based on a convolutional neural network. *Sensors*, *18*(5), 1429.
5. Khaddam, A. A., & Alzghoul, A. (2025). Artificial Intelligence-Driven Business Intelligence for Strategic Energy and ESG Management: A Systematic Review of Economic and Policy Implications. *International Journal of Energy Economics and Policy*, *15*(4), 635.
6. Li, Y., & O'Neill, Z. (2018, October). A critical review of fault modelling of HVAC systems in buildings. In *Building Simulation* (Vol. 11, No. 5, pp. 953-975). Berlin/Heidelberg: Springer Berlin Heidelberg.
7. Moreno, M. V., Dufour, L., Skarmeta, A. F., Jara, A. J., Genoud, D., Ladevie, B., &Bezian, J. J. (2016). Big data: the key to energy efficiency in smart buildings. *Soft Computing*, *20*(5), 1749-1762.
8. Pech, M., Vrchota, J., & Bednář, J. (2021). Predictive maintenance and intelligent sensors in smart factory. Sensors, 21(4), 1470.
9. Randhawa, P., Adusumilli, S. B. K., Poddutoori, J. R., Kapila, D., Poddutu, M., Kumar, S., ... & Devesh, S. (2025). Artificial Intelligence Driven Integrated Hydraulic and Pneumatic Pressure Control Systems for Advanced Regulation of Shut-off Valves. *Eng. Sci*, *35*(1427), 1427.

10. Sheikh, A., & Chong, E. K. (2025). Artificial Intelligence-Driven Optimization for Three-Dimensional Integrated Circuit Manufacturing: A System of Systems Framework. *IEEE Transactions on Components, Packaging and Manufacturing Technology*.
11. Tancev, G. (2021). Relevance of drift components and unit-to-unit variability in the predictive maintenance of low-cost electrochemical sensor systems in air quality monitoring. *Sensors*, *21*(9), 3298.
12. Tang, Y., Liu, S., Li, S., Ma, R., Li, Y., Wang, K., ... & Zhang, Y. (2025). A Review of Artificial Intelligence-Driven Innovations in Soft Magnetic Materials Optimization: Current Trends and Future Horizons. *Metal Mat*, *2*(1), e70001.
13. Zheng, H., Paiva, A. R., & Gurciullo, C. S. (2020). Advancing from predictive maintenance to intelligent maintenance with ai and IoT. *arXiv preprint arXiv:2009.00351*.
14. Zheng, S., Sha, J., Kong, Y., & Wang, Y. (2025). Research on artificial intelligence-driven container relocation problem for green ports. *Frontiers in Marine Science*, *12*, 1614356.
15. Abdulrazzq, R. A., Sajid, N. M., & Hasan, M. S. (2024). Artificial intelligence-driven predictive maintenance in IoT systems. *South Florida Journal of Development*, *5*(12), e4781-e4781.
16. Cummins, L., Sommers, A., Ramezani, S. B., Mittal, S., Jabour, J., Seale, M., & Rahimi, S. (2024). Explainable predictive maintenance: A survey of current methods, challenges and opportunities. *IEEE access*, *12*, 57574-57602.
17. Aires, F. L., Galeno, G. D., Belchior, F. N., Oliveira, A. M., & Hunt, J. D. (2025). Enhancing three-phase induction motor reliability with health index and artificial intelligence-driven predictive maintenance. *Royal Society Open Science*, *12*(5), 241946.
18. Bouabdallaoui, Y., Lafhaj, Z., Yim, P., Ducoulombier, L., & Bennadji, B. (2021). Predictive maintenance in building facilities: A machine learning-based approach. *Sensors*, *21*(4), 1044.
19. Cao, H., & Du, X. (2025). Elevator fault precursor prediction based on improved LSTM-AE algorithm and TSO-VMD denoising technique. *PLoS One*, *20*(4), e0320566.
20. Vlachou, V. I., Karakatsanis, T. S., Efstathiou, D. E., Vlachou, E. I., Vologiannidis, S. D., Balaska, V. E., & Gasteratos, A. C. (2025). Condition Monitoring and Fault Prediction in PMSM Drives Using Machine Learning for Elevator Applications. *Machines*, *13*(7), 549.