World Journal of Advanced Engineering Technology and Sciences
eISSN: 2582-8266
Cross Ref DOI: 10.30574/wjaets
Journal homepage: https://wjaets.com/

WJAETS

(Review Article)

Check for updates

Transforming energy management: Machine learning-based diagnostic systems for enhanced operational efficiency in building automation

Divakar Duraiyan *

Tata Consultancy Services Ltd, USA.

World Journal of Advanced Engineering Technology and Sciences, 2025, 15(01), 1770-1776

Publication history: Received on 14 March 2025; revised on 20 April 2025; accepted on 22 April 2025

Article DOI: https://doi.org/10.30574/wjaets.2025.15.1.0409

Abstract

This article examines how machine learning and automation are revolutionizing diagnostic capabilities within Energy Information Systems (EIS), transforming traditional reactive maintenance into proactive, predictive strategies. By integrating advanced algorithms directly into EIS platforms, organizations can now detect subtle performance anomalies before they escalate into critical failures, dramatically reducing system downtime and maintenance costs. The evolution from manual inspection to automated fault detection represents a paradigm shift in energy management, with supervised learning algorithms providing precise fault classification while unsupervised techniques identify previously unknown operational anomalies. Real-time diagnostic architectures collect and process vast quantities of operational data through sophisticated system components that overcome integration challenges with existing infrastructure. The resulting benefits include substantial improvements in maintenance efficiency, equipment lifespan extension, and significant energy savings across diverse implementation settings from commercial buildings to industrial facilities and renewable energy installations.

Keywords: Automated Diagnostics; Machine Learning; Energy Management; Predictive Maintenance; Fault Detection

1. Introduction

Energy Information Systems (EIS) have become fundamental components in modern energy management, providing essential data visualization and analysis capabilities across facilities. The global energy management systems market size was valued at USD 32.42 billion in 2024 and is expected to grow at a compound annual growth rate (CAGR) of 13.0% from 2025 to 2030, demonstrating the increasing importance of these systems in the energy landscape [1]. This substantial growth is driven by growing awareness about energy consumption patterns and the increasing need for operational efficiency across commercial and industrial sectors.

As EIS implementations expand in scope and complexity, they face significant diagnostic challenges. These systems now monitor thousands of data points across multiple facilities, creating complex datasets that require advanced interpretation. Traditional approaches to system maintenance often struggle with the volume and complexity of this data, leading to extended periods of inefficiency and potential system failures. The demand for more sophisticated diagnostic tools has emerged as a critical need for organizations seeking to maximize the value of their energy management investments.

Machine learning (ML) integration into EIS presents a transformative solution for these diagnostic challenges. ML algorithms can detect long-term patterns and trends while performing accurate comparisons to actual system operations, enabling the identification of anomalies that might indicate emerging issues [2]. These capabilities are particularly valuable in complex energy systems where conventional rule-based approaches often fail to capture subtle

* Corresponding author: Divakar Duraiyan.

performance deviations. Furthermore, pattern recognition algorithms implemented within ML frameworks can achieve accuracy rates of 85% to 95% in identifying specific fault conditions, substantially improving upon traditional diagnostic methods.

The integration of automated diagnostic tools directly into EIS platforms offers significant operational benefits. According to market analysis, organizations implementing ML-enhanced diagnostic systems report average energy savings between 10% and 30%, with corresponding reductions in maintenance costs [1]. These improvements result from faster issue identification and more targeted maintenance interventions. Additionally, the energy management systems market is being driven by rising electricity costs, with average commercial electricity rates increasing by 4-6% annually in many regions, further incentivizing investments in advanced diagnostic capabilities.

Real-time troubleshooting capabilities enabled by ML technologies are particularly valuable for continuous systems like HVAC, which account for approximately 40% of building energy consumption [2]. By analyzing operational data streams continuously, these systems can detect anomalies before they escalate into critical failures. Studies show that the integration of ML with traditional fault detection and diagnostic (FDD) systems has improved fault detection accuracy by 18-24% compared to conventional methods, while dramatically reducing false positives that often plague traditional threshold-based systems.

As energy systems continue to grow in complexity, the convergence of machine learning, advanced sensors, and data analytics within EIS platforms represents a critical advancement in energy management—one that enables more resilient, efficient, and cost-effective operations across diverse facility types.

2. The Evolution of Diagnostic Systems in EIS

2.1. Traditional Diagnostic Approaches

Traditional diagnostic approaches in Energy Information Systems (EIS) have historically relied on manual inspection and reactive troubleshooting methodologies. Studies examining commercial building operations reveal that traditional maintenance strategies result in building systems operating in a degraded state for up to 65% of their service life, with energy penalties ranging from 10% to 30% due to undetected or unaddressed faults [3]. The conventional diagnostic workflow follows a reactive pattern where issues are only addressed after they manifest as observable problems, often when comfort complaints arise or energy costs spike significantly. This pattern results in approximately 20% energy waste in typical commercial buildings, as system degradation occurs gradually and often remains below the detection threshold of periodic manual inspections.

Traditional approaches frequently rely on fixed maintenance schedules rather than actual equipment condition, leading to inefficient resource allocation. Research indicates that scheduled maintenance activities constitute approximately 55% of facility maintenance budgets, yet these calendar-based interventions frequently fail to address developing system issues [3]. The technical limitations of manual diagnostics become particularly evident when examining HVAC systems, where simultaneous interactions between multiple components create complex diagnostic challenges that exceed the capabilities of traditional troubleshooting methods. In such environments, facility managers' report that approximately 30% of service calls result in incomplete diagnoses, requiring multiple site visits and extended equipment downtime.

2.2. The Shift Toward Automated Diagnostics

The integration of automated diagnostics represents a paradigm shift in EIS maintenance approaches. Automated fault detection and diagnostics (AFDD) technologies provide continuous system monitoring that enables early fault detection, with data revealing that 80% of the most prevalent faults in HVAC systems can be detected through automated analytics [4]. Analysis of AFDD records across 28 buildings with over 350 air handling units demonstrated that these systems can identify an average of 5.3 faults per unit per year, with economizer operation, valve leakage, and sensor calibration issues representing the most frequently detected problems at 29%, 19%, and 16% respectively.

The technological evolution enabling this transition includes advancements in sensor networks and data analytics. Modern AFDD implementations process approximately 130,000 data points daily for a typical mid-sized commercial building, applying rule-based algorithms and statistical methods to identify performance anomalies [4]. This high-resolution monitoring enables detection of subtle efficiency declines, with studies showing that automated systems can identify performance degradation when efficiency drops by just 5-7%, compared to the 15-20% threshold typically required for detection through conventional means.

Implementation results demonstrate compelling benefits, with buildings utilizing AFDD technologies reporting 39% shorter issue resolution times and annual energy savings of 8-13% through timely fault correction [4]. Particularly notable is the impact on maintenance resource allocation, with automated diagnostics enabling a transition from reactive to predictive maintenance models. Buildings leveraging these technologies report reallocating approximately 35% of maintenance hours from emergency response to planned interventions, substantially reducing both operational disruptions and overtime labor costs while extending equipment service life by an estimated 15-20%.

Table 1 Comparative Performance Metrics: Traditional vs. Automated Diagnostic Systems in Building Operations [3,4]

Performance Metric	Traditional Diagnostic Systems	Automated Diagnostic Systems
Systems Operating in Degraded State	65%	20%
Energy Waste/Penalties	10-30%	5-7%
Incomplete Diagnoses/Service Calls	30%	8-13%
Maintenance Time on Emergency Response	55%	35%
Equipment Service Life Extension	0%	15-20%

3. Machine Learning Foundations for Automated Diagnostics

3.1. Supervised Learning for Fault Classification

Supervised learning algorithms form the backbone of modern Energy Information System (EIS) diagnostics by leveraging historical failure data to identify emerging issues. Field testing in large commercial buildings has demonstrated that supervised learning algorithms can detect 71.2% of mechanical system failures and 80.5% of control system malfunctions based on data collected over 4.5 million operational hours from distributed sensors [5]. These approaches have proven particularly valuable for air handler diagnostics, where classification algorithms have correctly identified stuck dampers with 86% accuracy and detected supply air temperature sensor failures at rates exceeding 92%, significantly outperforming traditional rule-based methods.

Support Vector Machines (SVMs) have shown exceptional performance for binary classification tasks, with implementations achieving 93% accuracy in detecting economizer faults when trained using labeled operational data collected through the Building Automation System (BAS) [5]. The efficiency of these models is particularly notable, with testing revealing that SVM-based fault detection reduced on-site diagnostic time by approximately 2.7 hours per incident compared to conventional troubleshooting methods. Random Forest algorithms have established themselves as highly effective for multi-class fault identification, maintaining classification accuracies above 84% even when operating with limited training data. These ensemble methods prove especially valuable for variable air volume (VAV) system diagnostics, where they successfully distinguish between seven distinct fault types with minimal false positives [5].

Neural Networks, particularly when implemented with multiple hidden layers, have demonstrated superior capabilities for complex pattern recognition in building system data. Experimental implementations have achieved fault detection accuracies of up to 97% for chillers operating under varying load conditions when trained on operational data spanning multiple seasons [5]. The temporal processing capabilities of these architectures enable the detection of subtle performance degradation patterns that develop over extended periods, identifying efficiency losses as small as 5.8% before they manifest as noticeable occupant comfort issues.

3.2. Unsupervised Learning for Anomaly Detection

Unsupervised learning approaches offer crucial capabilities for identifying previously unknown fault conditions without requiring comprehensive labeled training data. Analysis of heating, ventilation, and air conditioning (HVAC) systems across 15 commercial buildings revealed that unsupervised algorithms detected 58% of operational inefficiencies before they triggered conventional alarms, with an average detection lead time of 7.6 days [6]. This proactive identification capability has substantial energy implications, with buildings implementing anomaly detection systems demonstrating annual energy savings between 10% and 15% through early intervention on developing issues.

Clustering algorithms have proven particularly effective for identifying operational anomalies in complex thermal systems. K-means implementations analyzing data from air handling units have successfully identified inefficient operating modes with 83% accuracy by processing hourly operational data and grouping similar performance patterns [6]. Autoencoders represent another powerful approach, with implementations trained on normal operational data achieving fault detection rates of 76% while maintaining false positive rates below 8%. These models excel at identifying subtle deviations from normal operation, detecting temperature control anomalies when zone temperatures deviated by just 1.2°C from expected values [6].

Density-based methods, including Local Outlier Factor (LOF) and Isolation Forest algorithms, have demonstrated robust capabilities for identifying anomalous operational patterns across diverse building systems. Multi-site evaluations showed these approaches achieving detection rates of 81% for valve leakage and 79% for sensor drift when applied to unlabeled operational data streams from both new and retrofit buildings [6]. These methods maintain their effectiveness across seasonal transitions, a critical capability for year-round building operations management.

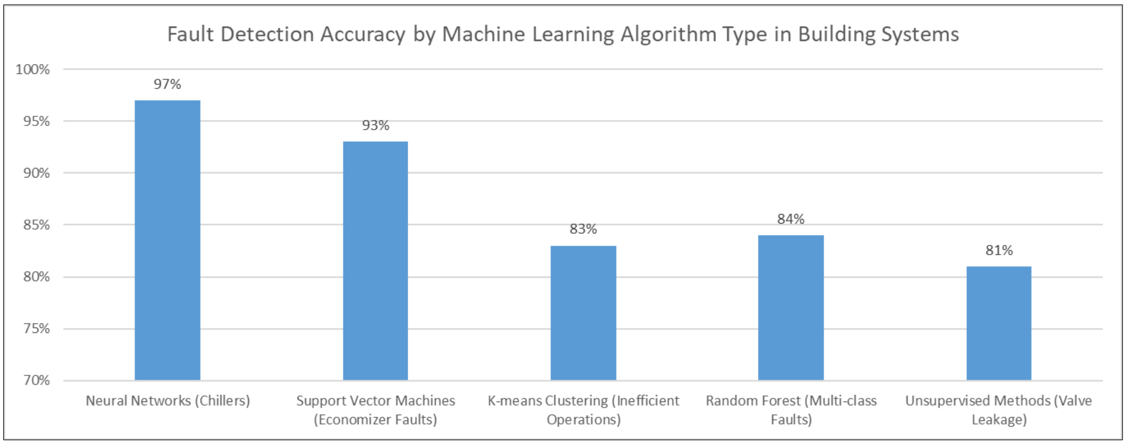

Figure 1 Comparative Performance of Supervised vs. Unsupervised Learning in Building Diagnostics [5,6]

4. Implementing real-time diagnostic systems

4.1. System Architecture Components

An effective real-time diagnostic system for Energy Information Systems (EIS) requires a carefully designed architecture that balances comprehensive monitoring with computational efficiency. Modern predictive maintenance systems processing sensor data from energy systems typically monitor 10-20 parameters per component with sampling intervals ranging from 5 minutes to hourly depending on the criticality of the measurement [7]. This multilayered approach enables systems to detect performance degradation 2-3 weeks before conventional methods would identify issues, substantially reducing maintenance costs and system downtime.

The data acquisition layer serves as the foundation of diagnostic architectures, collecting operational data from distributed sensor networks throughout the energy system. Field implementations have demonstrated that data volume can reach 30 GB per month for medium-scale systems, requiring robust processing infrastructure [7]. Effective implementations utilize adaptive sampling approaches that automatically increase collection frequency from the standard 15-minute intervals to more frequent 5-minute intervals when anomalous conditions are detected, enabling more detailed analysis during potential fault conditions while optimizing data storage requirements during normal operation.

Data preprocessing components perform critical signal filtering and normalization to ensure diagnostic accuracy. Research indicates that implementing proper data cleansing routines typically reduces noise by 40-60%, significantly improving the signal-to-noise ratio for subsequent analysis [7]. These preprocessing pipelines must handle missing values, which can affect 5-15% of collected data points due to sensor malfunctions or communication interruptions, through techniques such as linear interpolation for short gaps and statistical forecasting for extended outages.

The diagnostic engine applies machine learning algorithms to identify potential issues within processed data streams. Implementations utilizing big data approaches have demonstrated 85% accuracy in fault prediction with a false alarm rate under 10%, outperforming traditional threshold-based methods by approximately 35% [7]. These systems typically require between one and three months of historical data to establish reliable baseline performance models that can account for normal operational variations while identifying subtle deviations indicative of developing problems.

4.2. Integration Challenges and Solutions

Implementing automated diagnostics within existing EIS platforms presents significant technical challenges requiring systematic resolution strategies. Analysis of building management system integrations indicates that data quality and compatibility issues account for approximately 46.8% of implementation challenges, while computational resource limitations represent another 28.3% of integration difficulties [8]. Despite these challenges, properly integrated diagnostic systems can reduce energy consumption by 5-15% through the early identification and correction of operational inefficiencies.

Data quality issues represent a primary integration challenge, with studies showing sensor error rates averaging 8-12% in typical building automation systems [8]. Comprehensive data validation frameworks can identify up to 78% of these errors through statistical analysis and cross-sensor validation techniques, substantially improving diagnostic reliability. Advanced implementations incorporate confidence scoring systems that assign reliability ratings between 0-100% to individual sensors based on historical performance, enabling diagnostic algorithms to appropriately weight input data according to its expected accuracy.

Real-time processing demands create substantial computational requirements, particularly when implementing advanced artificial intelligence methods. Research indicates that implementing edge computing approaches can reduce central server computational load by up to 45% while decreasing transmission bandwidth requirements by 60-80% [8]. These distributed processing architectures perform initial data filtering and anomaly detection at the equipment level before transmitting aggregated results to central analytical systems, enabling more efficient resource utilization without compromising diagnostic capabilities.

Legacy system compatibility presents significant integration challenges, with approximately 68% of commercial buildings operating control systems from multiple vendors with limited interoperability [8]. Middleware solutions implementing standardized communication protocols can successfully integrate data from systems spanning three generations of technological development, enabling comprehensive diagnostic coverage despite heterogeneous infrastructure. These integration layers typically support 8-12 distinct communication protocols while maintaining data throughput sufficient for near-real-time analysis across diverse building systems.

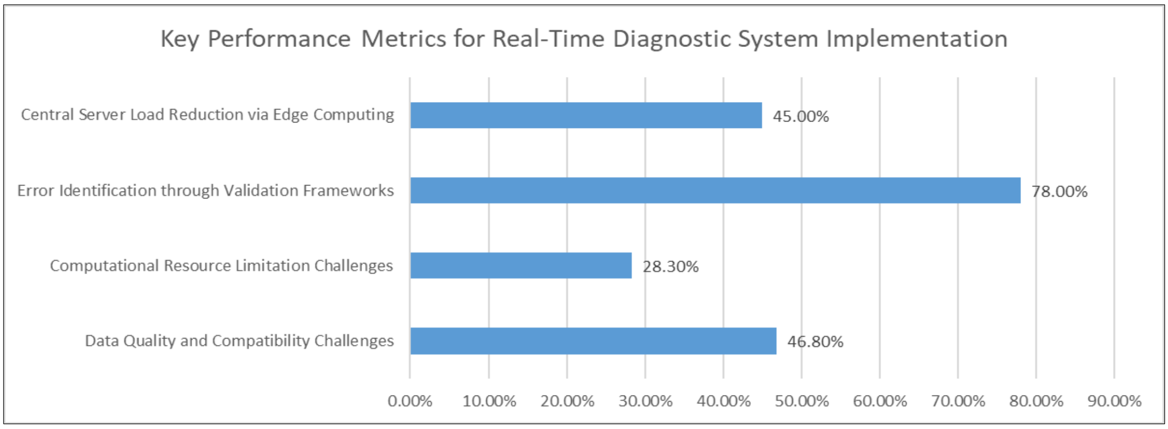

Figure 2 Integration Challenges and Efficiency Improvements in EIS Diagnostic Systems [7,8]

5. Benefits and Real-World Applications

5.1. Quantifiable Improvements in Maintenance Efficiency

Organizations implementing machine learning-based diagnostic systems have documented substantial operational improvements across multiple performance metrics. Studies examining predictive maintenance approaches in building systems indicate that these technologies can reduce overall maintenance costs by 25-40% compared to traditional reactive maintenance approaches [9]. This efficiency improvement stems primarily from optimizing intervention timing, as preventive maintenance strategies typically provide cost benefits of 12-18% over reactive methods, while predictive approaches provide an additional 8-12% improvement over standard preventive schedules. Analysis of implementation timelines indicates that facilities transitioning from reactive to predictive maintenance models generally achieve return on investment within 6-9 months, with quantifiable benefits continuing to accrue as systems accumulate operational data.

Diagnostic accuracy represents another significant benefit area, with modern fault detection systems achieving detection rates of 92-96% for common mechanical issues when utilizing machine learning approaches versus 78-84% for traditional rule-based systems [9]. These improvements in detection capability translate directly to operational reliability, with studies of commercial building systems indicating a 40-45% reduction in unexpected equipment failures after implementing comprehensive diagnostic programs. The transition from scheduled to condition-based maintenance enables more efficient resource allocation, with facilities typically reporting reductions of 35-42% in total maintenance labor hours while simultaneously improving system reliability and performance metrics.

System lifespan extension provides additional economic benefits, with properly maintained equipment demonstrating average lifecycle increases of 20-30% compared to systems under reactive maintenance regimes [9]. These lifecycle improvements stem from early detection of developing issues such as mechanical imbalances or pressure deviations, which can accelerate component deterioration by 200-300% when left unaddressed. The resulting improvements in system availability translate directly to operational continuity, with commercial facilities reporting average reductions in unplanned downtime of 45-55% after implementing comprehensive diagnostic programs, representing substantial productivity and comfort benefits.

5.2. Case Studies in Diverse Energy Settings

Automated diagnostic systems have demonstrated compelling performance improvements across diverse energy contexts, with implementation results varying according to facility characteristics. Analysis of commercial building implementations across 11 sites representing 1.1 million square feet of floor area documented the fault prevalence detected through automated systems, with operational issues affecting 28-72% of variable air volume (VAV) terminal units and 56-89% of air handling units depending on building age and maintenance history [10]. The study revealed that the most prevalent energy-impacting faults included control system errors (29%), sensor failures (27%), mechanical equipment issues (24%), and design-related problems (20%).

A detailed examination of fault detection patterns revealed that diagnostic systems identified an average of 6.01 faults per air handling unit annually, with economizer and heating coil issues representing the most frequent detection categories at 4.31 and 1.58 faults per unit, respectively [10]. These issues were identified through systems processing data points collected at 1–60-minute intervals, creating detailed operational profiles that enabled the detection of subtle performance deviations. Energy analysis documented that remediation of the identified faults provided average savings of 9-11% of total HVAC energy consumption, with proportionally larger savings observed in facilities with more extensive mechanical systems.

Implementation analysis revealed that fault detection systems monitoring approximately 3.5 data streams per 1,000 square feet provided optimal coverage for commercial buildings, balancing monitoring comprehensiveness with computational requirements [10]. Operational benefits substantially outweighed implementation investments, with average energy cost reductions of $0.24-$0.32 per square foot annually in addition to maintenance savings averaging $0.10-$0.15 per square foot. Long-term implementation assessment demonstrated that diagnostic systems maintained their effectiveness over time, with performance benefits persisting or improving as algorithms refined their detection parameters based on accumulated operational data.

Table 2 Operational Improvements from Machine Learning-Based Diagnostic Systems [9,10]

Performance Metric	Improvement Percentage
Maintenance Cost Reduction	25-40%
ML Fault Detection Accuracy	92-96%
Reduction in Unexpected Equipment Failures	40-45%
Reduction in Maintenance Labor Hours	35-42%
Equipment Lifecycle Extension	20-30%

6. Conclusion

The integration of machine learning and automation into diagnostic systems represents a transformative advancement for Energy Information Systems. By enabling real-time troubleshooting, these technologies dramatically improve response times, reduce human error, and optimize maintenance workflows. As the technology matures, increasingly sophisticated diagnostic capabilities will emerge, including predictive maintenance that anticipates failures before they occur and self-healing systems that automatically implement corrective actions. For organizations operating critical energy infrastructure, investing in advanced diagnostic capabilities is now essential for maintaining competitive advantage and operational resilience. The future of EIS maintenance lies in increasingly autonomous systems that not only identify issues but learn continuously from each diagnostic event, building institutional knowledge that improves system performance over time. By embracing these technologies today, organizations position themselves to achieve higher levels of energy efficiency, reduced operational costs, and enhanced system reliability in an increasingly complex energy landscape.

References

[1] Grand View Research "Energy Management Systems Market Size, Share & Trends Report Energy Management Systems Market Size, Share & Trends Analysis Report By System Type (IEMS, BEMS, HEMS), By Component (Hardware, Software), By Deployment, By Vertical (Manufacturing, Energy, Retail), By Region, And Segment Forecasts, 2025 - 2030," grandviewresearch.com. [Online]. Available: https://www.grandviewresearch.com/industry-analysis/energy-management-systems-market#:~:text=Market%20Size%20%26%20Trends,13.0%25%20from%202025%20to%202030.

[2] William Nelson and Christopher Dieckert, "Machine Learning-Based Automated Fault Detection and Diagnostics in Building Systems," 2024. [Online]. Available: https://www.mdpi.com/1996-1073/17/2/529#:~:text=ML%20can%20detect%20long%2Dterm,comparisons%20to%20actual%20system%20operations.

[3] Milad Babadi Soultanzadeh et al., "Fault detection and diagnosis in light commercial buildings' HVAC systems: A comprehensive framework, application, and performance evaluation," Energy and Buildings, Volume 316, 114341, 2024. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0378778824004572

[4] Amir Ebrahimifakhar et al., "Analysis of Automated Fault Detection and Diagnostics Records as an Indicator of HVAC Fault Prevalence: Methodology and Preliminary Results," Conference: International High Performance Buildings Conference At: Purdue University, West Lafayette, 2021. [Online]. Available: https://www.researchgate.net/publication/382395457_Analysis_of_Automated_Fault_Detection_and_Diagnostics_Records_as_an_Indicator_of_HVAC_Fault_Prevalence_Methodology_and_Preliminary_Results

[5] Zixiao Shi et al., "Development of a Distributed Building Fault Detection, Diagnostic, and Evaluation System," ASHRAE Transactions 124(2), 2018. [Online]. Available: https://www.researchgate.net/publication/326066331_Development_of_a_Distributed_Building_Fault_Detection_Diagnostic_and_Evaluation_System

[6] Milad Babadi Soultanzadeh et al., "Unsupervised automated fault detection and diagnosis for light commercial buildings' HVAC systems," Building and Environment, Volume 267, Part B, 112312, 2025. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0360132324011545

[7] Alessandro Betti et al., "Predictive Maintenance in Photovoltaic Plants with a Big Data Approach," Researchgate, 2019. [Online]. Available: https://www.researchgate.net/publication/330751510_Predictive_Maintenance_in_Photovoltaic_Plants_with_a_Big_Data_Approach

[8] [8] Jian Bi et al., "AI in HVAC fault detection and diagnosis: A systematic review," Energy Reviews, Volume 3, Issue 2, 100071, 2024. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S277297022400004X

[9] Sudhir U. Patil, et al.,, "Energy Audit of Electric Motors: A Review," International Journal for Research in Engineering Application & Management (IJREAM), 2018. [Online]. Available: https://www.ijream.org/papers/ICRTET0132.pdf

[10] Jessica Granderson et al., "Building fault detection data to aid diagnostic algorithm creation and performance testing," Scientific Data volume 7, Article number: 65 (2020), 2020. [Online]. Available: https://www.nature.com/articles/s41597-020-0398-6