2026 Volume: 02 Issue: 11 www.ijrpa.com ISSN 2456-9995 Research Article

International Journal Research Publication Analysis

Page: 01-12

AGENTIC AI-BASED PREDICTIVE MAINTENANCE FOR ELECTRIC VEHICLES USING IOT SENSOR DATA

***Geethapriya G. H., Raghu G., Jeevitha M., Kishan G., Rachana M. K.**

Department of Computer Science and Engineering Jyothy Institute of Technology Bangalore, India.

Article Received: 25 September 2026
Article Revised: 15 October 2026
Published on: 04 November 2026

***Corresponding Author: Geethapriya G. H.**
Department of Computer Science and Engineering Jyothy Institute of Technology Bangalore, India.

DOI: https://doi-doi.org/101555/ijrpa.7288

ABSTRACT

The global transition toward eco-friendly transportation has positioned the reliability of Electric Vehicle (EV) propulsion systems as a key industrial concern. Sudden motor or battery malfunctions during high-speed transit present severe safety hazards. Traditional maintenance models, predominantly reactive or rigidly scheduled, often fail to address incipient faults effectively, resulting in avoidable downtime and high ownership costs.

This systematic review analyzes 16 high-impact research papers and technical reports to evaluate the transition from traditional condition monitoring to intelligent maintenance frameworks powered by Agentic AI and the Internet of Things (IoT). Existing reviews often lack coverage of recent developments (2024–2026) in Agentic Multi-Agent Systems (MAS) integrated with full-stack environments such as the MERN stack for autonomous fleet management.

Our analysis identifies a modular agentic architecture—balancing real-time edge diagnostics using ESP32 microcontrollers and cloud-based Prognostics and Health Management (PHM)—as the most viable solution for modern EV reliability. Results indicate that AI-driven models achieve fault classification accuracies of 85%–95%, potentially reducing unplanned outages by 35%.

INDEX TERMS: Electric Vehicles, Predictive Maintenance, IoT, Agentic AI, Multi-Agent Systems, Fault Detection, MERN Stack.

I. INTRODUCTION

The transition from internal combustion engines to EV platforms represents a definitive shift

in automotive engineering, positioning the electric motor and battery system as the "kinetic core" of the vehicle [1]. Unlike conventional engines, EV drivetrains undergo rigorous thermal cycling and road induced mechanical shocks that can lead to catastrophic failure if not continuously monitored [10]. Equipment reliability has become a major concern for manufacturers across industries; for instance, unplanned downtime in industrial plants is estimated to cost between $10,000 and $250,000 per hour, totaling over $50 billion annually [8]. Despite the significance of this problem, a 2017 study revealed that 70% of companies lack complete awareness of when their equipment is due for maintenance.

Predictive Maintenance (PdM), a keystone of Industry 4.0, provides a transformative solution by utilizing real-time sensor data and advanced analytics to forecast failures before they occur [4], [9]. This approach extends asset lifespan and optimizes technical performance through timely interventions. The motivation for an agentic AI approach arises from the need for autonomous decision-making in complex, non-linear environments where human intervention is slow or costly [16].

Agentic AI involves specialized software entities—agents—that continuously monitor sensors, detect anomalies, predict failures, and even schedule service operations independently [2].

This survey investigates how the integration of low-cost IoT nodes, such as the ESP32, with agentic frameworks can democratize advanced maintenance technologies [12], [15]. The remainder of this paper is organized as follows: Section II presents the survey methodology; Section III discusses the taxonomy of maintenance technologies; Section IV provides a comparative study; Section V analyzes the discussion and technical selection; Section VI outlines open research challenges; and Section VII concludes the study.

II. SURVEY METHODOLOGY

To ensure a rigorous, transparent, and reproducible synthesis of the state of the art in EV maintenance, this study follows the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines [3]. The methodology ensures that the selected literature represents high-quality research in AI-driven diagnostics and IoT integration.

A. Search Strategy

Systematic searches were conducted across multiple recognized scientific databases, including IEEE Xplore, ScienceDirect, Scopus, SpringerLink, and Google Scholar [2]. These databases provide extensive coverage of artificial intelligence, fault detection, and power

systems. The search period was restricted to publications from 2020 to 2026 to ensure the inclusion of recent advancements in Edge AI and Agentic frameworks [3], [8].

B. Inclusion and Exclusion Criteria

This review focuses on peer-reviewed journal articles, conference papers, and technical reports. Inclusion criteria required a primary focus on:

- AI applications in EV fault detection (motor, battery, or power electronics) [1], [13].
- IoT-enabled real-time monitoring using edge or hybrid nodes [14], [15].
- Prognostics and Health Management (PHM) using Agentic or Multi-Agent Systems [4], [16].

Conversely, articles focusing on unrelated domains, such as general finance or conventional internal combustion engine
maintenance, or those lacking practical validation and methodological rigor, were excluded.

C. **KEYWORDS:** Primary search strings utilized Boolean operators for specificity:
("Predictive Maintenance" OR "PHM") AND ("Electric Vehicles" OR "EV") AND ("Agentic AI" OR "MAS") AND "IoT"

D. Study Selection and Filtering Process

The selection process was executed in four distinct phases to minimize bias while ensuring technical relevance. In the Identification phase, the initial keyword search yielded over 3,400 records across the selected databases. After removing duplicates and non-English publications, 2,150 records remained.

In the Screening phase, titles and abstracts were independently reviewed against the inclusion criteria. This phase eliminated papers that addressed general IoT applications without a predictive maintenance component or those focusing exclusively on software-as-a-service (SaaS) without hardware integration details. This reduced the dataset to 420 articles.

In the Eligibility phase, the remaining papers were analyzed for technical depth. Special attention was given to the hardware–software interface, particularly implementations involving microcontrollers such as ESP32 or ARM Cortex M series [5], [12]. Papers lacking clear performance metrics (e.g., accuracy, precision, or Remaining Useful Life (RUL) error bounds) were excluded. This resulted in 180 high-quality articles.

Finally, 16 core sources representing the state of the art in Agentic AI and MERN-integrated architectures were selected for synthesis.

E. Quality Assessment

To further refine the selection, each core paper was evaluated using a Weighted Quality Assessment (WQA) score. Three primary dimensions were assessed on a scale of 1 to 5: 1) Technological Innovation: Whether the study presents a novel AI architecture or IoT configuration [6].

2) Empirical Validation: Whether results are based on real world datasets or high-fidelity simulations such as Digital Twins [7].

3) Scalability: Whether the solution is feasible for largescale EV fleet deployment [13].

Only studies with an average score of 4.0 or higher were prioritized for the comparative analysis in Section IV.

F. Data Extraction and Synthesis

A structured coding framework was employed to extract critical parameters, including metadata (author, year, publisher), technical specifications (AI models such as Random Forest, LSTM, and CNN), and key output metrics (Remaining Useful Life, Mean Absolute Error, and anomaly detection performance) [4], [10].

The synthesis also emphasized identifying hardware constraints and communication protocols (e.g., MQTT, HTTP, Web Sockets) used in IoT architectures [5], [14]. This systematic approach enables a detailed comparison of how Agentic AI addresses the limitations of traditional centralized cloud based models.

III. Taxonomy of Available Technologies

The evolution of EV maintenance is characterised by a transition from static, human-dependent monitoring to dynamic, autonomous orchestration [5]. The literature identifies four primary technological paradigms that define the current state of the art:

A. Traditional Sensing and Model-Based Paradigms

1) Signal-Based and Model-Based Diagnostics: These classical methods represent the foundation of Condition-Based Maintenance (CBM) [15]. Signal-based techniques focus on time-domain or frequency-domain analysis of physical symptoms, such as Fast Fourier Transform (FFT) analysis of motor vibrations or current signature analysis [14], [15]. Although computationally efficient, they depend heavily on expert defined thresholds, which struggle to handle the non-linear thermal and mechanical transients typical of EV propulsion systems [10].

Model-based approaches utilise mathematical representations (e.g., Kalman Filters or State-

Space models) to provide analytical redundancy. However, the high complexity involved in modelling battery electrochemical degradation or multiphase inverter switching makes these models difficult to generalise across different EV makes and models.

B. Data-Driven Intelligence Paradigms

1) Resource-Heavy Cloud Models: With the advent of Big Data, research has shifted toward centralised Deep Learning (DL) architectures [4], [8]. These systems leverage high performance cloud clusters to train and execute models, including Convolutional Neural Networks (CNNs) for spatial feature extraction and Gated Recurrent Units (GRUs) or LSTMs for temporal State-of-Health (SoH) forecasting [4].

By aggregating telemetry from large vehicle fleets, cloud models can recognise subtle, fleet-wide degradation trends that single-vehicle models may miss. However, the "Cloud Bottleneck"— characterised by high data transmission costs and latencies often exceeding 500 ms — prevents these models from participating in mission-critical, on-road safety decisions.

2) Lightweight Edge Computing Models: To mitigate cloud latency, the literature emphasises Edge Intelligence [12]. Deployed on low-power microcontrollers such as the ESP32 or ARM Cortex-M, these models focus on "on-device" inference [5], [14]. They utilise lightweight algorithms such as Isolation Forest, One-Class SVM, or TinyML-optimised neural networks [14].

These models act as the first line of defence, offering real-time (¡100 m s) anomaly detection for events such as motor stalls or battery thermal runaway. However, their limited memory (typically ¡4 MB) restricts their diagnostic depth, often resulting in a binary "Healthy/Unhealthy" output rather than detailed Remaining Useful Life (RUL) estimations [14].

C. The Emerging Agentic Orchestration Paradigm

1) Agentic Multi-Agent Systems (MAS): The most recent advancement (2024–2026) involves the integration of Agentic AI with the Industrial Internet of Things (IIoT) [2], [16]. Unlike traditional AI systems that rely on human prompts, Agentic AI operates as a self-sufficient entity that perceives its environment and autonomously pursues complex goals.

In an EV context, this paradigm is typically orchestrated through a MERN (MongoDB, Express, React, Node.js) stack to manage the lifecycle of a fault [13]:

- Data Acquisition Agent: Polls CAN Bus and MEMS sensors via MQTT protocols, ensuring data integrity at the hardware level [12].

- Diagnostic Agent (Edge-based): Performs rapid anomaly detection directly on the vehicle [14].
- Prognostic Agent (Cloud-based): Analyses historical trends to forecast RUL with high precision using ensemble learning techniques [6].
- Orchestration & Scheduling Agent: Interfaces with a central MongoDB database to autonomously check workshop availability and schedule maintenance through a React-based dashboard, effectively closing the loop from "Detection" to "Action" [13].

The transition across these paradigms reflects a shift in the "Intelligence Bottleneck" from data collection to independent decision-making. While traditional models focus on identifying faults, Agentic frameworks extend this capability by determining the appropriate response, scheduling interventions, and optimising overall fleet uptime.

IV. COMPARISON STUDY

The table below summarises the performance metrics and limitations of the principal AI and IoT technologies discussed in the reviewed literature. This cross-study comparison provides a quantitative foundation for selecting an optimal architecture for EV predictive maintenance.

A. Performance Synthesis and Trade-offs

The analysis indicates that supervised learning methods, such as Random Forest (RF) and XGBoost, consistently achieve the highest accuracy (up to 100% in controlled environments such as [11]) when historical failure data is available [1], [6]. However, these metrics can be misleading in realworld EV scenarios. The high accuracy of these models is often tied to specific motor profiles, leading to significant *overfitting*, where a model trained on one fleet fails to generalise to different battery chemistries or motor configurations.

TABLE I: COMPARATIVE STUDY OF PDM TECHNOLOGIES.

Ref	Technology / Framework	Metric	Major Limitations
[1]	XGBoost Classifier	99% Acc.	Poor generalisation between vehicles
[6]	ANN (50 neurons)	99.9% R-val	Requires large labelled faulty data
[11]	Random Forest (RF)	100% Acc.	High memory usage for large ensembles
[4]	CNN-LSTM Hybrid	92% Prec.	Black-box opacity and high latency
[14]	Isolation Forest	73% Acc.	Higher false-negative rate in noisy data
[12]	LR-HGBC-CNN	94% Acc.	Complex multi-stage implementation
[15]	MCSA Threshold	RT Trend	No prognostic RUL forecasting

Furthermore, computational cost remains a critical differentiator. While the ANN architecture in [6] achieves a near perfect R-value, its requirement for multiple neurons and deep layers

makes it unsuitable for energy-constrained edge devices such as the ESP32. In contrast, the Isolation Forest approach in [14], despite its lower accuracy (73%), has a significantly smaller memory footprint—nearly 10 times lower than CNNLSTM hybrids—enabling true on-device monitoring without cloud dependency.

B. The Dilemma of Data Shortage

A significant finding from the surveyed literature is the crucial role of unsupervised and semi-supervised learning methods [2], [10]. In the EV domain, propulsion systems are engineered for high reliability; therefore, real-world catastrophic failure data is extremely scarce. This results in a *data imbalance* problem.

While deep learning models such as CNN-LSTM [4] are highly effective at detecting temporal degradation patterns, they require large volumes of labelled faulty data, which are often proprietary or unavailable to researchers.

Consequently, there is a growing trend toward hybrid approaches, such as LR-HGBC-CNN [12], which combine the interpretability of Logistic Regression (LR) with the feature extraction capabilities of CNNs. However, these multi-stage architectures introduce significant deployment complexity, requiring sophisticated orchestration (e.g., MERN stack) to manage model interactions across edge and cloud environments.

C. Prognostic Depth vs. Real-Time Response

Finally, the comparative study highlights a clear tradeoff between prognostic depth and real-time responsiveness. Traditional Motor Current Signature Analysis (MCSA) threshold methods [15] provide instantaneous alerts but lack the predictive capability required to estimate Remaining Useful Life (RUL).

In contrast, agentic frameworks discussed in Section III aim to address this limitation by decoupling the *Detection* metric (accuracy) from the *Action* metric (response time) [16]. By offloading RUL prediction to cloud-based systems while maintaining lightweight anomaly detection at the edge, researchers are achieving accuracy levels exceeding 95% without compromising the low-latency requirements of high-speed EV operation [12], [13].

This comparative analysis confirms that no single algorithm serves as a “silver bullet” [4]. The most effective solution is a heterogeneous Multi-Agent System (MAS), where different algorithms are deployed strategically based on their strengths and limitations.

V. Discussion and Technical Selection *A.*

Critical Evaluation of Present Paradigms

The comparative analysis presented in Section IV reveals a fundamental dichotomy in Electric Vehicle (EV) maintenance research: the trade-off between prognostic depth and operational agility [2], [4]. Category 2 (cloud-based deep learning) models, while achieving superior Remaining Useful Life (RUL) prediction accuracy, introduce a "communication bottleneck" [4]. In high-speed transit scenarios, latency exceeding 250 ms—compounded by network jitter during cellular handovers—can be the difference between a controlled motor shutdown and a catastrophic inverter failure.

Conversely, Category 3 (edge-only models) prioritise local safety but suffer from "intelligence isolation" [14]. These models can detect anomalies through threshold-based alerts on devices such as the ESP32, but they lack the computational context required to distinguish between transient load spikes and incipient bearing faults [5]. Furthermore, they do not possess the higher-level decision-making capability necessary to initiate external workflows, such as inventory management or service logistics, thereby leaving the maintenance loop incomplete.

B. The Supported Agentic AI and MERN Framework

This survey identifies a decentralized *Agentic AI architecture integrated with the MERN stack* as the most robust solution for modern EV fleet management [2], [13]. This selection is based on its ability to provide multi-layered resilience through functional decomposition:

1) Decoupled Latency via Edge Agents: By deploying a Fault Detection Agent directly on the ESP32 using C++/Arduino-based inference, the system ensures that safety-critical current signature analysis is performed at the source [12], [14]. This reduces decision-making latency to sub-100 ms, meeting the requirements for real-time propulsion safety.

2) Autonomous Workflow via MERN Orchestration: The integration of a MERN (MongoDB, Express, React, Node.js) stack transforms the system from a passive monitoring tool into an active operational platform [13]. While the Prediction Agent performs computationally intensive LSTM-based forecasting in the cloud [4], the Scheduling Agent leverages the Express/Node.js backend to autonomously interact with service databases. This ensures that maintenance is not only predicted but also executed—effectively closing the loop from anomaly detection to physical repair without human intervention.

3) Trust through Explainable AI (XAI): A key advantage of the proposed framework is its support for Explainable AI (XAI) [3], [9]. Traditional black-box models often face resistance

from industry stakeholders and regulatory authorities. The Agentic framework enables the integration of techniques such as LIME (Local Interpretable Model-agnostic Explanations) and SHAP (Shapley Additive Explanations). By visualizing feature contributions—such as radiator temperature or battery humidity—on a React-based dashboard, the system provides transparent and interpretable justifications for autonomous maintenance decisions.

C. Regulatory and Industrial Alignment

The selection of this hybrid architecture is further justified by its alignment with international standards [13]. The modular design of Agentic Multi-Agent Systems (MAS) supports compliance with ISO 13374 (Machinery Diagnostics) and ISO 26262 (Functional Safety for Road Vehicles) by ensuring a clear and traceable audit trail of agent interactions [10].

Furthermore, the use of a centralised MongoDB database enables the implementation of robust data governance and privacy mechanisms, supporting compliance with regulations such as the General Data Protection Regulation (GDPR) for vehicle telematics data.

Ultimately, the Agentic approach addresses the "semantic gap" between low-level sensor data and high-level industrial decision-making [16]. By distributing intelligence strategically—ensuring safety-critical decisions at the edge and long term planning in the cloud—this framework establishes a scalable blueprint for next-generation EV fleets with minimal unplanned downtime.

VI. OPEN RESEARCH CHALLENGES

Despite significant advancements in integrating Agentic AI [12] with IIoT frameworks, several formidable barriers persist that hinder the universal adoption of autonomous maintenance in the Electric Vehicle (EV) industry [2], [13]. Addressing these challenges is critical for transitioning from theoretical prototypes to production-ready fleet management systems.

A. Multimodal Data Heterogeneity and Sensor Fusion

EV data streams originate from highly heterogeneous sources, ranging from the Battery Management System (BMS) and CAN bus telemetry to high-frequency MEMS vibration sensors [10], [11]. These streams operate at varying sampling frequencies—kilohertz for vibration versus hertz for thermal data—complicating synchronization for multimodal sensor fusion. Future research should explore Temporal Alignment Networks and Graph Neural Networks (GNNs) to effectively model structural and temporal dependencies among these disparate data sources, ensuring that Agentic decision-making is grounded in a cohesive

digital representation of the vehicle state [3], [16].

B. Extreme Dataset Imbalance and the Role of Generative AI

Because modern EVs are engineered for high reliability, real-world catastrophic failure data is extremely scarce, leading to a long-tail distribution problem [6], [7]. This scarcity of faulty labels significantly degrades the performance of supervised learning models. To address this challenge, future research should focus on Physics-Informed Digital Twins and Generative Adversarial Networks (GANs) to synthesize high fidelity fault data [4]. By simulating rare failure modes—such as internal battery dendrite growth or inter-turn short circuits in motors—researchers can train Agentic AI systems in virtual environments prior to real-world deployment.

C. Cybersecurity, Trust, and Agentic Governance

The interconnected nature of IoT-enabled EV systems introduces a broad attack surface for adversarial threats [12]. Malicious actors may inject false telemetry to trigger unnecessary maintenance actions or suppress critical safety alerts. Robust User and Entity Behavior Analytics (UEBA) are required to monitor agent interactions in real time. Additionally, implementing blockchain-based decentralized ledgers for agent logs can provide immutable audit trails, ensuring that autonomous decisions remain verifiable and tamper-proof. This approach supports compliance with emerging automotive cybersecurity standards such as ISO/SAE 21434 [2].

D. Edge-to-Cloud Scalability and 5G Integration

Scaling an Agentic framework from a single EV to a large fleet introduces significant challenges in managing telemetry data volume [5], [8]. Traditional 4G/LTE networks often suffer from latency, jitter, and congestion, which can hinder real time MERN stack operations. The integration of 5G network slicing is essential to provide dedicated, low-latency communication channels for safety-critical data. Furthermore, Federated Learning (FL) offers a promising solution by enabling vehicles to train models locally and share only model updates, thereby reducing bandwidth usage and preserving data privacy [2].

E. Standardization and Interoperability

A major limitation in current systems is the lack of standardized communication protocols for AI agents in the automotive ecosystem [13]. For effective deployment, agents developed by different manufacturers must interoperate seamlessly—for example, a Fault Detection Agent communicating with a third party Scheduling Agent. Developing Semantic Web technologies and standardized ontologies for EV health data is essential for building an interoperable ecosystem in which multiple MERN-based platforms can collaborate efficiently [16].

The resolution of these challenges requires a transdisciplinary approach that integrates power electronics, distributed computing, and cybersecurity. As the industry progresses toward Industry 5.0, these research gaps represent the final frontier in achieving a fully autonomous, zero-unplanned downtime mobility infrastructure [8], [12], [16].

VII. CONCLUSION

Predictive maintenance has evolved from a niche industrial tool into a strategically vital capability for the Electric Vehicle (EV) industry [1], [9], [10]. This review demonstrates that integrating Agentic AI, IoT, and MERN stack technologies transforms maintenance from a reactive burden into an autonomous, value-adding service [2], [13].

By orchestrating specialized agents for data acquisition, fault detection, and proactive scheduling, the proposed architecture can reduce service restoration time by up to 60% and prevent approximately 45% of potential disruptions [4], [12].

As the industry moves toward Industry 5.0, synchronization between the physical vehicle's health and its virtual Digital Twin prognostic counterparts will define the future of zero unplanned-downtime mobility [8].

REFERENCES

1. S. Konkimalla, "AI-Based Predictive Maintenance for Electric Vehicles," *International Journal of Engineering Computer Science (IJECS)*, 2022.
2. S. Rana, "AI-Driven Fault Detection for Electric Vehicles," *International Conference on Smart Mobility and AI Systems*, 2025.
3. J. Costa *et al.*, "Systematic Review on AI Predictive Maintenance," *Journal of Artificial Intelligence Research*, 2026.
4. Zheng *et al.*, "Intelligent Maintenance Using Deep Learning," *IEEE Transactions on Industrial Informatics*, 2020.
5. R. Yadwad, "IoT-Based Fault Detection System," *International Journal of IoT Applications*, 2025.
6. N. A. Mohammed, "Machine Learning-Based Predictive Maintenance," *International Journal of Advanced Computing Systems*, 2023.
7. Ercevik, "Electric Vehicle Predictive Maintenance Framework," *Journal of Sustainable Transportation Systems*, 2024.
8. Emmanuel, "Artificial Intelligence in Industry 5.0," *International Journal of Smart Manufacturing*, 2025.

9. Patil, “Predictive Maintenance in Manufacturing Systems,” *IEEE Access*, 2024.
10. Gonzalez, “Electric Vehicle Fault Diagnosis Techniques,” *IEEE Transactions on Vehicular Technology*, 2021.
11. K. V., “Motor Predictive Maintenance Using Random Forest,” *International Journal of Electrical Systems*, 2022.
12. Chavan, “ESP32-Based Predictive Maintenance Architecture,” *International Conference on Embedded AI Systems*, 2026.
13. M. Rahman, “IoT-Based Predictive Maintenance for EV Systems,” *Journal of Intelligent Transportation Systems*, 2025.
14. P. Kolok *et al.*, “IoT Predictive Maintenance Framework,” *IEEE Internet of Things Journal*, 2025.
15. Shukla *et al.*, “Motor Monitoring and Predictive Analysis,” *International Journal of Motor Control Systems*, 2022.
16. Samatas *et al.*, “Artificial Intelligence and IoT Integration,” *IEEE Systems Journal*, 2021.