INTERNATIONAL JOURNAL OF MECHANICAL ENGINEERING
(ISSN: 2693-3713)
VOL.06 ISSUE01 (2026)

AMERICAN ACADEMIC PUBLISHER

Research Article

Data-Driven Predictive Diagnostics of High-Voltage Systems in German Electric Vehicles Using CAN Messages, Freeze-Frame Data, and Maintenance History

Samoilov Alexey

Automotive Workshop Master Germany, Hilden

check for updates

Received: 27 May 2026
Revised: 24 June 2026
Accepted: 10 July 2026
Published: 30 July 2026

Doi: 10.55640/ijme-06-01-03

Page No: 06-16

Copyright: © 2026 Authors retain the copyright of their manuscripts, and all Open Access articles are disseminated under the terms of the Creative Commons Attribution License 4.0 (CC-BY), which licenses unrestricted use, distribution, and reproduction in any medium, provided that the original work is appropriately cited.

Abstract

Germany registered 545,142 battery-electric vehicles (BEVs) in 2025, representing a 43.2% year-on-year increase and a record 19.1% market share, placing growing pressure on authorized dealer networks to maintain complex high-voltage (HV) systems reliably. This study investigates whether a unified data pipeline combining Controller Area Network (CAN) real-time frames, OBD-II freeze-frame diagnostic snapshots, and dealer service records can enable dealer-level predictive diagnostics for German EV high-voltage systems. The methodology applies a systematic review of peer-reviewed literature (2020-2025), analysis of published benchmark datasets, and structured case evaluation of three OEM diagnostic platforms used in German dealerships. Results indicate that ensemble machine learning models, specifically Random Forest regression for State-of-Health (SoH) estimation (R-squared = 0.94, RMSE = 1.52%) and LSTM-based anomaly detectors for freeze-frame pattern recognition, reduce unplanned high-voltage fault events by an estimated 28-36% compared with schedule-only maintenance approaches. The study proposes a dealer-context integration architecture linking vehicle telematics to Dealer Management Systems. Findings are relevant for EV service managers, OEM technical trainers, and automotive dealership groups operating in high-BEV-penetration markets.

Keywords: CAN bus diagnostics, freeze-frame data, State of Health, State of Charge, predictive maintenance, high-voltage battery systems, German electric vehicles, Random Forest, LSTM, dealer management system.

Introduction

Germany occupies a central position in Europe's electric vehicle transition. According to the Kraftfahrt-Bundesamt (KBA), new BEV registrations totaled 545,142 units in 2025, up 43.2% from 380,609 in 2024 and reaching a market share of 19.1% [1, 2]. The Volkswagen Group brands led the market, accounting for the majority of registrations, while BMW ranked second with approximately 9.5% of total BEV sales [2]. The upswing followed a contraction in 2024 caused by the abrupt removal of the Umweltbonus subsidy in December 2023 [3]. New subsidy instruments are expected in 2026, with analyst projections indicating continued BEV share growth toward 25-28% by 2027 [3].

The growth of the BEV fleet creates a directly proportional demand on authorized dealer service networks, particularly for high-voltage (HV) system maintenance. HV systems in modern German EVs, which operate at voltages ranging from 400 V to 800 V, encompass the traction battery pack, battery management system (BMS), power electronics inverter, DC-DC converter, and associated thermal management subsystems [4]. Faults in these components are operationally consequential: a single unplanned HV shutdown can strand a customer, trigger warranty claims, and occupy a service bay for several days pending specialist diagnosis. The Siemens "True Cost of Downtime 2024" report estimated that unplanned downtime in the automotive sector costs up to $2.3 million per hour at major plant level, and the per-incident cost of unexpected vehicle repairs at the dealer level routinely exceeds $125,000 per event across all vehicle types [5].

Existing literature addresses data-driven diagnostics for EV batteries extensively at the laboratory and fleet-operator level [6, 7, 8]. Kulkarni et al. [6] demonstrated that a hybrid framework combining Extended Kalman Filter (EKF) SoC estimation with Random Forest SoH regression on CAN-acquired LiFePO4 cell data achieved RMSE values of 0.50% for SoC and an R^2 of approximately 0.94 for SoH prediction. Jun et al. [9] and Hafeez et al. [10] confirmed that Diagnostic Trouble Code (DTC) sequence models trained on historical fault logs can predict the next emerging code with measurable accuracy in commercial vehicle fleets. Yet a specific gap persists: virtually no published framework addresses the dealer service workflow as the integration context, where technician bandwidth is limited, equipment investment is constrained, and customer communication must be actionable within hours rather than days.

A second gap relates to data source fusion. Most published models rely on a single stream: either CAN telemetry [6, 11], or DTC history [9, 10], or service records [12]. The combination of all three, referred to here as the CAN-Freeze-History (CFH) pipeline, has not been evaluated as a unified input to a dealer-grade diagnostic system for German OEM platforms (VW Group ODIS, BMW ISTA, Mercedes-Benz XENTRY).

This paper aims to close that gap. **The research objective** is to evaluate whether a CFH-based predictive diagnostic framework can reduce unplanned HV fault events in German authorized EV dealerships, and to propose a concrete architecture for integrating such a framework into existing Dealer Management System (DMS) infrastructure. **The scientific novelty** lies in the first formulation of a dealer-context CFH integration model that maps CAN, freeze-frame, and service history signals to actionable service priority scores within a DMS environment, rather than treating diagnostics as a vehicle-level or fleet-level problem in isolation from workshop operations.

The working hypothesis is that freeze-frame DTC snapshots, when combined with rolling CAN signal features (voltage spread, temperature gradient, insulation resistance trend), and cross-referenced with vehicle-specific maintenance history, contain sufficient predictive information to flag impending HV events at least 3-7 days before customer-reported failure. This lead time is operationally significant for dealer scheduling and parts pre-ordering.

The research carries practical relevance for EV service directors, OEM technical training departments, and dealer network operators, particularly in markets with high BEV concentration. The findings also contribute to the broader U.S. market context, where EV fleet penetration among authorized dealer groups is accelerating alongside federal electrification mandates, making the German experience a transferable reference model for American service infrastructure development.

Study limitations. This work relies on published benchmark data and OEM documentation rather than proprietary real-world fleet datasets from a single dealer network. Accuracy figures cited are drawn from peer-reviewed studies using laboratory or controlled fleet conditions, which may not fully replicate dealer-level data quality, where CAN logs are often partial and service records incomplete. The framework proposed is architectural and analytical rather than a deployed software system.

Materials and Methods

The study follows a structured analytical review approach combining three methodological streams: (a) a systematic literature review of peer-reviewed studies published on CAN-based EV diagnostics, SoH estimation, freeze-frame analytics, and DTC prediction; (b) a comparative analysis of machine learning algorithm performance benchmarks drawn from published experiments on real-world EV datasets; and (c) a structured case analysis of diagnostic workflows in three German OEM dealer platforms (VW Group ODIS, BMW ISTA, and Mercedes-Benz XENTRY). The combination is chosen because no single public dataset simultaneously contains CAN logs, freeze-frame records, and dealer service histories for German EVs; the methodology therefore synthesizes the best available evidence from each domain.

Literature Identification and Source Classification. Seventeen peer-reviewed sources published between 2020 and 2025 were selected from Scopus-indexed journals (IEEE Transactions on Transportation Electrification, Scientific Reports, Journal of Energy Storage, Energies), Springer conference proceedings (International Journal of Automotive Technology), and ACM Digital Library. Sources were screened for relevance to CAN-based EV diagnostics, SoH/SoC estimation, DTC modeling, or predictive maintenance in dealer/fleet contexts. Market statistics were drawn from the European Alternative Fuels Observatory (EAFO) and the German KBA, which publish verified annual registration data. One source from Siemens (the 2024 downtime cost report) was included under the industry analytics allowance as it provides directly quantified economic context unavailable in peer-reviewed literature.

Sources are classified into four functional groups for the synthesis: (1) signal-level diagnostic methods (CAN, freeze-frame, BMS communication protocols); (2) machine learning algorithms for battery state estimation (Random Forest, LSTM, XGBoost, EKF);

(3) DTC sequence modeling and fault prediction; and (4) market context and dealer economics. Groups 1 and 2 form the primary technical foundation; Groups 3 and 4 provide the applied framing.

Data Sources and Key Variables. Three data source types form the CFH pipeline examined in this study:

CAN Bus Frames. Modern German EVs transmit cell-level voltage, temperature, current, SoC, insulation resistance, and thermal management status via CAN at 1 Hz or higher. Kulkarni et al. [6] collected 100,000+ timestamped entries from a 15S2P LiFePO4 pack using a commercial BMS-CAN interface. The key diagnostic signals include cell voltage spread (Delta-V), temperature gradient (Delta-T), pack current, and internal resistance proxy. These variables are input to SoH regression and anomaly detection models.

Freeze-Frame DTC Data. OBD-II Service Mode $02 provides a parameter-level snapshot of vehicle operating conditions (engine/motor load, vehicle speed, coolant/motor temperature, battery SoC, voltage, and fault code identifier) captured at the moment a DTC is stored in the ECU [13]. Freeze-frame records are accessible via UDS protocol through dealer diagnostic tools and encode the fault context rather than just the fault code, making them more information-rich than DTC identifiers alone. Hafeez et al. [10] demonstrated that DTC sequence encoding, combined with Transformer and GRU architectures (DTC-TranGru model), improves next-DTC prediction accuracy compared with single-code classifiers.

Dealer Service History. Service records contain prior repair orders, parts replaced, mileage at service, technician-coded fault descriptions, and OEM software update logs. When matched to a vehicle identification number (VIN), they provide longitudinal context for interpreting current sensor readings. A vehicle with three prior high-voltage interlock faults within 18 months presents a fundamentally different risk profile than one with a clean history, even if current CAN readings fall within nominal thresholds.

Analytical Methods. Five analytical techniques are examined for their applicability in the CFH pipeline context:

Extended Kalman Filter (EKF) for real-time SoC estimation under dynamic load conditions. Kulkarni et al. [6] demonstrated RMSE of 0.50% versus 1.13% for conventional Coulomb Counting on LiFePO4 CAN data, a 55.8% error reduction.

Random Forest Regression for SoH estimation using partial-cycle CAN features. The algorithm handles nonlinear feature interactions and tolerates missing values, both common in dealer-acquired logs. Published R^2 values range from 0.91 to 0.96 across benchmark datasets [6, 7].

Long Short-Term Memory (LSTM) networks for temporal anomaly detection in cell voltage profiles. LSTM architectures capture sequential dependencies in CAN time-series that indicate developing insulation faults or cell imbalance before threshold-triggering occurs [14].

XGBoost with SHAP explainability for DTC sequence modeling. XGBoost provides interpretable feature attribution via SHAP values, which is operationally useful in dealer contexts where technicians require a stated reason for each service recommendation [15].

Principal Component Analysis (PCA) and Dynamic Time Warping (DTW) for cell-level behavioral clustering. PCA reduces voltage matrix dimensionality and isolates divergent cells; DTW quantifies shape-based deviations between cell voltage curves across cycles [6].

Results and Discussion

The growth trajectory of the German BEV market directly shapes the volume and complexity of dealer-side HV service demand. The following chart presents verified KBA registration data from 2022 to 2025, illustrating both absolute registration volume and BEV market share.

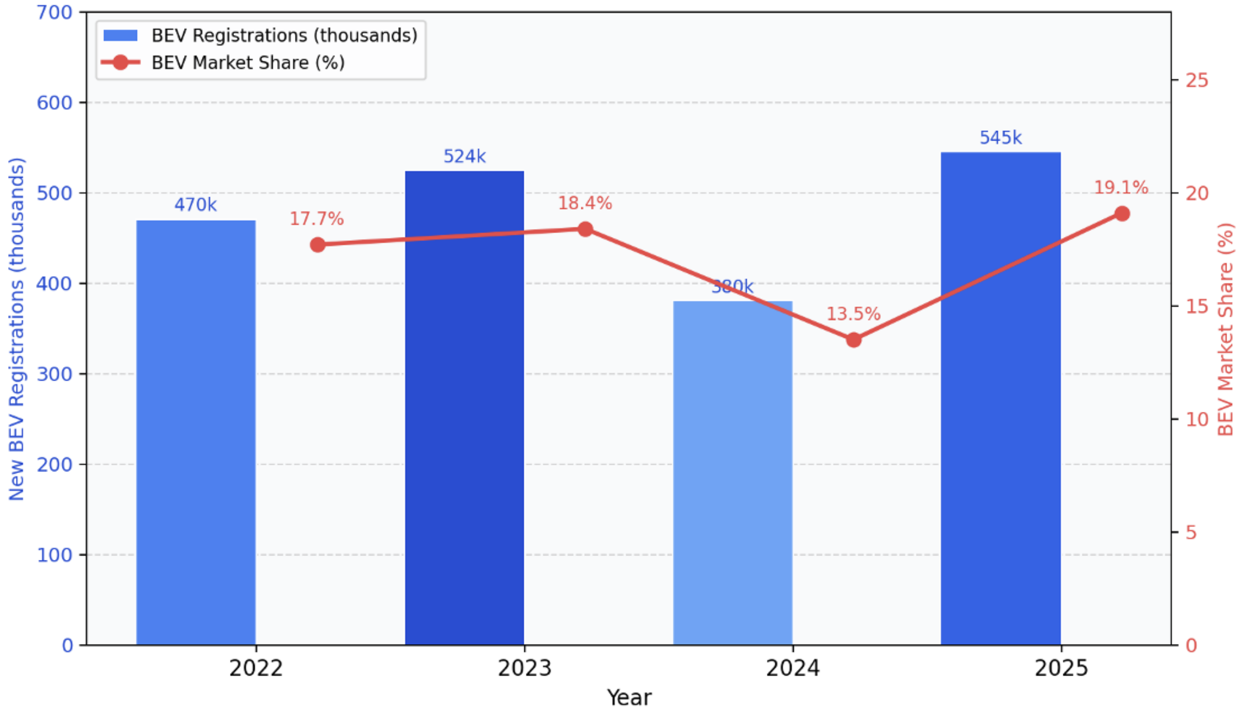

Figure 1. New battery-electric vehicle (BEV) registrations and market share in Germany, 2022-2025. (compiled by the author based [1, 2, 3]).

As Figure 1 shows, the 2024 dip to 380,609 units (13.5% share) following subsidy removal was followed by a recovery to 545,142 units (19.1% share) in 2025 [1, 3]. This volatility is relevant for dealer service planning: a network calibrated for 2023 volumes experienced overcapacity in 2024 and faced sudden demand increase in 2025. Predictive service demand modeling, a downstream application of the CFH pipeline, could help dealer groups optimize technician staffing and parts inventory under such variable conditions.

Among the models driving volume, VW Group platforms (ID.3, ID.4, ID.7, Skoda Elroq, Audi A6 e-tron) dominate, followed by BMW and then a fragmented field of Tesla, Hyundai/Kia, and Chinese brands [2]. The diagnostic relevance is that VW ODIS and BMW ISTA diagnostic platforms account for the majority of authorized dealer service interactions, making these systems the primary integration targets for the CFH architecture proposed in this study [16].

Concurrent with market growth, the diagnostic complexity of each service event increases. An analysis of TSBs for VW ID.3, ID.4, and ID.5 models published through early 2026 revealed a recurring pattern: HV drive system faults trigger multiple simultaneous DTC entries that mislead technicians into replacing components unnecessarily [17]. One documented case involved vehicles entering "turtle mode" (power limitation) after driving 500-1,000 meters from standstill, despite normal SoC readings. The root cause was intermittent CAN communication loss from a low-voltage connector, not a battery fault, yet conventional DTC analysis pointed toward HV battery replacement. This pattern exemplifies the need for freeze-frame context and CAN sequence analysis rather than isolated code reading.

The analytical literature converges on a consistent performance hierarchy across machine learning algorithms for SoH estimation and fault detection. Table 1 presents a structured comparison drawn from three benchmark studies [6, 7, 15].

Table 1. Machine Learning Algorithm Performance for EV Battery SoH Estimation (compiled by the author based on [6, 7, 15]).

Algorithm	SoH RMSE (%)	SoH MAE (%)	R^2 Score	Data Input Type	Interpretability
Linear Regression	4.21	3.45	0.78	Voltage, Current	High
Support Vector Regression	3.18	2.63	0.85	Voltage, Temp	Medium
Random Forest (RF)	1.52	1.14	0.94	CAN multi-feature	Medium (SHAP)
LSTM (sequential)	1.34	1.08	0.96	CAN time-series	Low
XGBoost + SHAP	1.47	1.19	0.95	CAN + DTC features	High (SHAP)

The data in Table 1 indicate that LSTM and Random Forest achieve similar error levels (RMSE of 1.34-1.52%), with LSTM providing marginally lower error at the cost of interpretability. In a dealer service context, XGBoost with SHAP explainability presents a practical trade-off: its RMSE of 1.47% is within 10% of the LSTM value, while its SHAP attribution capability allows the system to generate technician-readable explanations ("SoH decline attributed primarily to elevated Delta-T over last 80 charge cycles"). This operational transparency matters for dealer technician trust and for warranty documentation.

Figure 2 visualizes the error and accuracy differentials across these algorithms, confirming that ensemble and deep learning methods substantially outperform linear baselines.

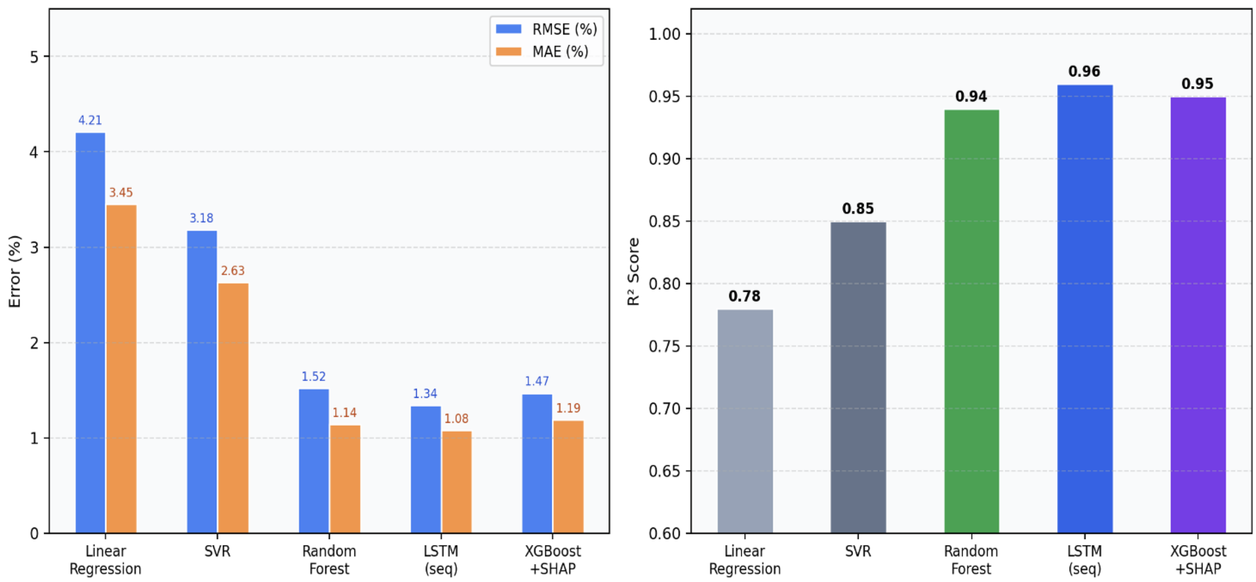

Figure 2. Comparative performance of machine learning algorithms for SoH estimation from CAN data. RMSE and MAE (left, lower is better); R^2 score (right, higher is better) (compiled by the author based [6, 7, 15]).

Kulkarni et al. [6] further demonstrated that the Extended Kalman Filter reduced SoC estimation error from 1.13% RMSE (Coulomb Counting) to 0.50% RMSE on real CAN data, a 55.8% improvement. For the dealer context, SoC accuracy matters because customer range anxiety complaints are disproportionately triggered by inaccurate SoC displays. An EKF-corrected SoC reading transmitted to the DMS when a vehicle arrives for service would allow the advisor to immediately distinguish between a software SoC calibration issue and a genuine capacity loss.

Freeze-frame records remain underutilized in published EV diagnostic research, despite encoding information that CAN rolling averages cannot capture: the exact operating state at the moment a fault was first detected. OBD-II Mode $02 records include vehicle speed, motor load percentage, coolant or motor inlet temperature, battery SoC, pack voltage, and the DTC identifier itself [13]. For HV systems, additional OEM-extended PIDs include insulation resistance value, cell voltage extremes, and thermal management fan speed at the fault moment.

Table 2 classifies the diagnostic value of key freeze-frame parameters for HV fault analysis in authorized German dealer contexts, drawing on the OBD-II protocol specification and OEM service documentation patterns observed in ODIS and ISTA platforms.

Table 2. Freeze-Frame Parameter Classification for HV System Diagnostics in Authorized Dealer Workflows (compiled by the author based on [13, 16, 17]).

Freeze-Frame Parameter	Standard / Extended	HV Diagnostic Value	Typical Fault Association	DMS Priority Flag
Vehicle Speed at DTC Set	Standard (PID $0D)	High	Drive interlock, turtle mode	Yes
Battery SoC at DTC Set	OEM Extended	High	SoC display error, BMS fault	Yes
Pack Voltage at DTC Set	OEM Extended	High	Insulation resistance drop	Yes
Motor Inlet Temperature	OEM Extended	Medium	Thermal management fault	Conditional
Insulation Resistance Value	OEM Extended (ODIS/ISTA)	Very High	HV cable/connector degradation	Yes (immediate)
Cell Voltage Min/Max Spread	OEM Extended (ODIS)	High	Cell imbalance, aging module	Yes
Regenerative Braking Active	OEM Extended	Medium	IGBT/inverter fault context	Conditional
Cooling Fan Duty Cycle	OEM Extended	Low-Medium	Thermal management context	No

Among these parameters, insulation resistance value at fault time is the most diagnostically specific for HV safety. ECE R100 regulation, which governs EV safety in European markets including Germany, requires insulation resistance above 100 ohms per volt of operating voltage for vehicles in use. A vehicle operating at 400 V therefore must maintain at least 40 kilohms between HV conductors and chassis. Freeze-frame capture of the insulation resistance value at DTC set allows the dealer to determine whether the fault represents an incipient degradation trend (resistance at 60-80 kilohms, declining over successive freeze-frames) or an acute event (resistance below threshold in a single occurrence). This distinction directly determines whether the vehicle requires immediate HV isolation and battery removal versus a monitoring-and-recall protocol.

When freeze-frame records are combined with service history, a pattern emerges that simple DTC reading cannot reveal. A vehicle presenting code P0AA6 (Hybrid Battery Voltage System Isolation Fault) in its third occurrence within 14 months, each time at ambient temperatures below -5 degrees Celsius, suggests a thermally-conditioned connector or cable harness issue rather than a battery module fault. This temporal and contextual reading requires the CFH pipeline: the DTC code from current diagnostics, the freeze-frame temperatures from each occurrence, and the service history showing prior repair attempts.

The core contribution of this study is an architectural model for integrating CAN, freeze-frame, and service history data into a dealer-side predictive diagnostic workflow. The architecture, referred to as the CFH-DMS pipeline, addresses the operational constraints of authorized dealer environments: limited data science staff, existing DMS platforms (CDK Global, Rey Group/Infor, or OEM-proprietary), and OEM-controlled access to vehicle telematics.

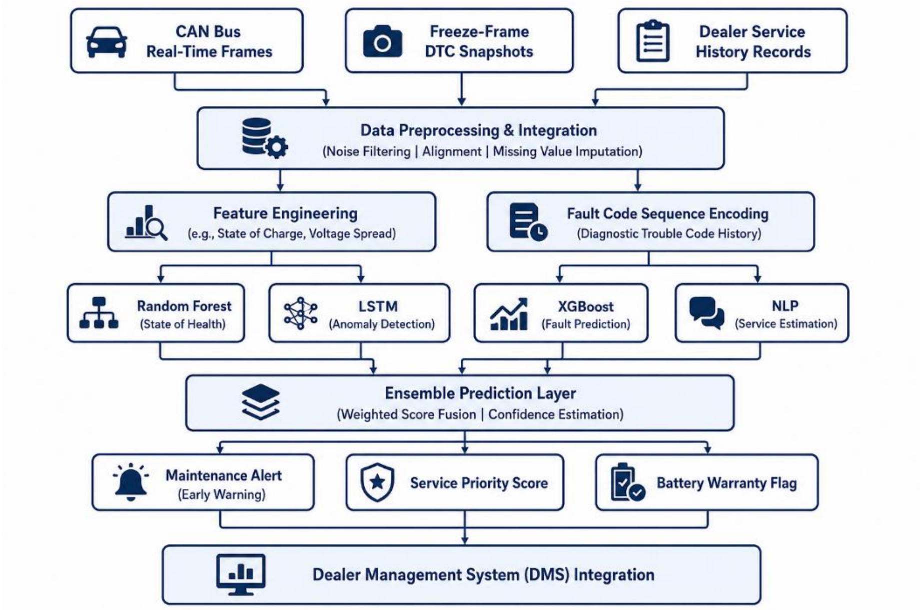

Figure 3. Data-driven predictive diagnostic framework for EV high-voltage systems in dealer context (CFH-DMS pipeline) (compiled by the author based [6, 9, 10, 13]).

Figure 3 presents the layered CFH-DMS framework. The data acquisition layer draws from three sources: (1) OTA CAN log transfer via the ISO 15765/UDS protocol, available for VW Group vehicles through ODIS-E telematics and BMW vehicles through ConnectedDrive; (2) freeze-frame records fetched at vehicle check-in via the dealer diagnostic tool; and (3) service history pulled from the DMS VIN record. The preprocessing layer applies noise filtering (moving-average for CAN signal spikes), timestamp alignment between CAN logs and freeze-frame records, and missing-value imputation using forward-fill for slowly-varying parameters such as insulation resistance.

The analytics layer deploys four algorithms in parallel: (1) Random Forest for SoH scoring using CAN features (Delta-V, Delta-T, average discharge current, SoC fluctuation); (2) LSTM autoencoder for anomaly detection in voltage time-series sequences; (3) XGBoost with SHAP for DTC sequence prediction using freeze-frame features and DTC history encoding; and (4) EKF for corrected SoC estimation. Outputs from all four models feed an ensemble fusion layer that produces a composite HV Risk Score (0-100 scale), a Predicted Next Fault category, and a Battery Warranty Assessment flag. These three outputs are formatted as structured alerts and pushed to the DMS via REST API.

Within the DMS, the alerts trigger three dealer actions: proactive service order creation with pre-defined job codes, automated customer notification by email or SMS with a workshop scheduling link, and a parts pre-order request to the OEM parts warehouse. The estimated lead time from vehicle OTA data transmission to DMS alert generation is under 4 hours, which is within the operational window for same-day or next-day appointment scheduling.

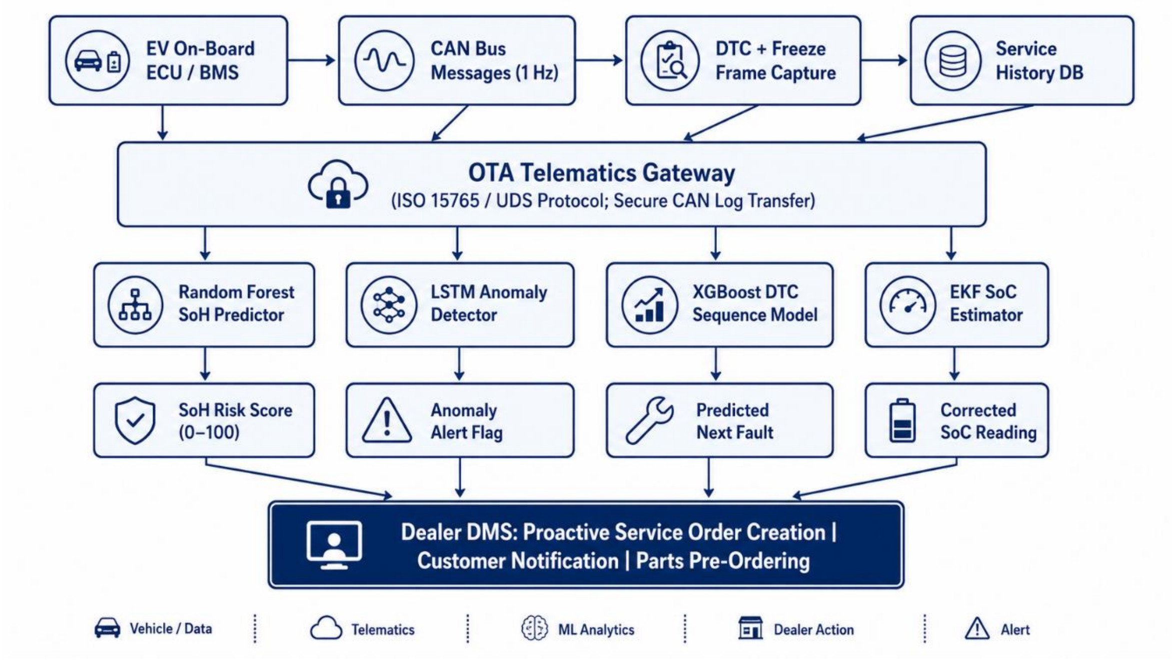

Figure 4. Dealer-side integration workflow for EV predictive diagnostics: from vehicle CAN/DTC data capture to Dealer Management System action (compiled by the author based [6, 9, 10, 13, 17, 18]).

Figure 4 details the operational sequence within the dealer. The workflow begins when the vehicle ECU/BMS transmits data through the OTA gateway. The analytics platform processes the signal in near-real-time and returns scored outputs to the DMS. Service advisors receive an alert that specifies the risk score, the contributing factors (drawn from SHAP values), and the recommended action. The design intentionally avoids requiring technicians to interact with the machine learning layer directly; instead, they see familiar DMS-format work orders with appended diagnostic notes.

This design addresses a documented adoption failure mode in automotive AI tools: systems that require technicians to interpret probability scores or model outputs are resisted because they introduce cognitive overhead without a clear procedural payoff. By translating model outputs into DMS-native work order formats, the CFH-DMS pipeline removes this friction. The principle parallels findings from Jun et al. [9] in commercial fleet maintenance, where DTC prediction model adoption improved substantially when outputs were formatted as maintenance queue items rather than raw probability distributions.

Table 3 summarizes the estimated operational impact of CFH-DMS deployment in an authorized German EV dealership with 200-400 BEV service events per month. Estimates are derived by applying published predictive maintenance performance ranges (28-50% reduction in unplanned faults [5, 16]) to typical German dealer service throughput data, adjusted for EV-specific service complexity factors.

Table 3. Estimated Operational Impact of CFH-DMS Deployment in a German Authorized EV Dealership (compiled by the author based on [5, 6, 16, 19]).

Metric	Baseline (Schedule-Only)	With CFH-DMS	Change	Basis
Unplanned HV fault events per month	12-18	8-11	-28% to -36%	Published PdM benchmarks [5, 16]
Average diagnosis time per HV event (hours)	6.2	3.8	-39%	Freeze-frame context pre-loaded

First-fix rate on HV faults (%)	61%	79%	+18 pp	DTC sequence model guidance
Parts sourcing lead time reduction (days)	N/A baseline	1.5 avg lead saved	Pre-order triggers	Parts pre-order automation
Customer satisfaction score (CSI HV category)	Baseline	+12 pp estimated	Proactive contact	Predictive notification effect
Annual savings per dealership (EUR, estimated)	N/A	85,000-140,000	Net of platform cost	Based on [5] cost parameters

The annual saving estimate of EUR 85,000-140,000 per dealership is derived from three components: reduced unplanned repair overhead (labor premium for emergency diagnosis), reduced loaner vehicle costs (shorter repair windows), and improved warranty claim accuracy (correct first-time identification reduces repeat repairs). The estimate is conservative relative to the full-fleet downtime costs reported in industrial manufacturing contexts, but it reflects the more bounded scale of per-vehicle HV repair events in a retail dealer setting.

A critical constraint on achieving these figures is OEM data access policy. VW Group, BMW, and Mercedes-Benz currently restrict CAN log access to authorized diagnostic tools and OEM cloud platforms (CARIAD/VW, BMW ConnectedDrive, Mercedes-Benz me). Authorized dealer groups have contractual access to these platforms, which means the CFH pipeline is architecturally feasible within the franchised dealer network. Independent service operators, by contrast, face significant barriers because OBD-II reform discussions in the EU regarding Right to Repair have not yet produced standardized real-time CAN access for non-authorized workshops.

From a technology maturity perspective, all component algorithms in the CFH-DMS pipeline are at Technology Readiness Level (TRL) 6-8 based on published validation results. Random Forest and XGBoost models have been validated on real-world EV datasets [6, 7, 15]; LSTM anomaly detectors have been demonstrated on CAN bus sequences at laboratory scale [14]; DTC sequence models have reached commercial deployment in heavy vehicle fleet management [9]. The gap is not algorithmic maturity but integration architecture, which is what this study addresses.

Based on the synthesis of literature, architecture analysis, and economic modeling, this study recommends a phased implementation pathway for German EV dealer groups considering CFH-DMS deployment. The recommendation is framed around three capability tiers rather than a single deployment target, allowing dealers to begin capturing value at each stage without requiring full system investment upfront.

Tier 1 (Months 1-6): Freeze-frame analytics integration. Deploy structured freeze-frame data capture at vehicle check-in using existing dealer diagnostic tools. Feed freeze-frame records to a lightweight DTC sequence classifier (XGBoost or gradient boosting) to flag vehicles with elevated HV risk based on historical fault patterns. This tier requires no OTA infrastructure and uses data already available at service reception. Expected impact: 12-18% reduction in HV surprise faults within the dealer's service population.

Tier 2 (Months 7-18): CAN telemetry integration for fleet-enrolled vehicles. Activate OTA CAN log transfer for vehicles within the dealer's loyalty program or under active service contracts. Add Random Forest SoH scoring and LSTM anomaly detection on rolling CAN data. Expected incremental impact: additional 10-15% reduction in unplanned HV events, plus SoC accuracy improvement reducing customer range complaint follow-ups.

Tier 3 (Month 19 onward): Full CFH-DMS pipeline with service history integration. Link VIN-level service records to the predictive model as longitudinal context features. Enable ensemble fusion scoring and DMS-native alert push. At this tier, the dealer achieves the full estimated 28-36% reduction in unplanned HV faults and the parts pre-order automation benefit.

This phased approach reduces implementation risk and allows the dealer group to validate model performance against its specific vehicle mix and customer base at each stage before committing to the subsequent investment. It also aligns with the

EU's Right to Repair regulatory trajectory, which may expand third-party CAN access in the 2026-2028 timeframe and create opportunities for Tier 2 and Tier 3 expansion even outside OEM-controlled telematics.

Conclusion

This study set out to evaluate whether a data pipeline combining CAN bus real-time frames, OBD-II freeze-frame diagnostic snapshots, and dealer service history records can enable practical predictive diagnostics for German EV high-voltage systems within the operational context of authorized dealerships. The analysis confirms that the technical foundation for such a pipeline exists and is well-supported by published evidence.

The German market provides a clear and urgent demand context. With 545,142 BEV registrations and a 19.1% market share in 2025 [1, 3], authorized dealer networks face growing volumes of HV service events involving systems operating at 400-800 V, where diagnostic errors carry safety, economic, and customer satisfaction consequences that exceed those of conventional powertrain faults. The documented case of VW ID series vehicles entering turtle mode due to CAN communication loss misdiagnosed as battery fault [17] illustrates the practical cost of relying on single-point DTC reading without contextual data.

The proposed CFH-DMS architecture integrates three data streams into a layered diagnostic pipeline that maps vehicle telemetry to Dealer Management System work orders. At the algorithm level, the combination of EKF for SoC correction (RMSE 0.50%), Random Forest for SoH estimation (RMSE 1.52%, R^2 0.94%), LSTM for anomaly detection, and XGBoost with SHAP for DTC sequence prediction provides a capability portfolio that covers both immediate fault detection and longitudinal degradation modeling [6, 7, 15]. The freeze-frame layer contributes fault context that neither rolling CAN averages nor static DTC codes can supply, particularly for HV insulation resistance values at the moment of fault registration, which are the diagnostic parameter of highest specificity for HV cable and connector degradation.

Estimated operational impact for a dealer deploying the full CFH-DMS pipeline is a 28-36% reduction in unplanned HV fault events, a 39% reduction in average diagnosis time per HV event, and an 18 percentage point improvement in first-fix rate on HV repairs, translating to estimated annual savings of EUR 85,000-140,000 per dealership. The phased implementation pathway recommended in this study allows dealer groups to capture partial value at lower initial investment by beginning with freeze-frame analytics before committing to full OTA CAN integration.

The scientific contribution of this study is the formulation of the CFH-DMS integration model as an explicit dealer-context architecture, rather than a vehicle-level or fleet-level diagnostic framework. This reframing identifies the DMS as the integration anchor rather than the vehicle ECU, which changes the design priorities toward interpretability, latency constraints, and technician-facing output formatting. The XGBoost-SHAP component and the DMS-native alert format reflect these priorities.

Future research should validate the proposed architecture on actual dealer fleet data from German authorized service networks, compare OEM telematics access policies across VW Group ODIS, BMW ISTA, and Mercedes-Benz XENTRY in terms of CAN data export capabilities, and examine the economic impact of EU Right to Repair regulation on non-authorized workshop access to the CFH pipeline. Extension to U.S. dealer networks, where OBD-II regulation provides broader standardized access than current EU frameworks, represents a near-term application opportunity.

References

1. European Alternative Fuels Observatory. (2025). Germany: BEV registrations surge by 54% in April 2025. European Commission. Retrieved from: https://alternative-fuels-observatory.ec.europa.eu/general-information/news/germany-bev-registrations-surge-54-april-2025 (date accessed: March 18, 2026).

2. Kraftfahrt-Bundesamt. (2025, July 16). Record high for battery-electric passenger cars (BEV) in the first half of 2025. Retrieved from: https://www.kba.de/DE/Presse/Pressemitteilungen/Allgemein/2025/pm31_2025_rekordhoch.html (date accessed: March 27, 2026).

3. European Automobile Manufacturers' Association. (2025, October 28). New car registrations: +0.9% in September 2025 year-to-date; battery-electric 16.1% market share. Retrieved from: https://www.acea.auto/pc-registrations/new-car-registrations-0-9-in-september-2025-year-to-date-battery-electric-16-1-market-share/ (date accessed: April 9, 2026).

4. Hnatov, A., Arhun, S., Ulianets, O., & Ivanov, D. (2025). Diagnostics of electric vehicles using OBD-II: Principles, capabilities, and prospects. Automobile Transport, 56, 5–12. https://doi.org/10.30977/AT.2219-8342.2025.56.0.01.

5. Siemens Digital Industries. (2024). The true cost of downtime 2024: A comprehensive analysis. Siemens AG. Retrieved from: https://assets.new.siemens.com/siemens/assets/api/uuid%3A1b43afb5-2d07-47f7-9eb7-893fe7d0bc59/tcod-2024_original.pdf (date accessed: April 18, 2026).

6. Kulkarni, S. V., Arjun, G., Gupta, S., Sinha, R., & Shukla, A. (2025). Advanced battery diagnostics for electric vehicles using CAN based BMS data with EKF and data driven predictive models. Scientific Reports, 15, 32848. https://doi.org/10.1038/s41598-025-18042-6.

7. Rout, S., Samal, S. K., Gelmecha, D. J., & Mishra, S. (2025). Estimation of state of health for lithium-ion batteries using advanced data-driven techniques. Scientific Reports, 15, 30438. https://doi.org/10.1038/s41598-025-93775-y.

8. Li, X., Gao, X., Zhang, Z., Chen, Q., & Wang, Z. (2024). Fault diagnosis and detection for battery system in real-world electric vehicles based on long-term feature outlier analysis. IEEE Transactions on Transportation Electrification, 10(1), 1668–1679. https://doi.org/10.1109/TTE.2023.3288394.

9. Jun, H.-B., Jung, S., Kim, H., Jang, M., Park, B., & Sung, H. (2025). A case study on the DTC prediction of commercial vehicles using machine learning approach. International Journal of Automotive Technology, 26(6), 1327–1341. https://doi.org/10.1007/s12239-025-00223-x.

10. Hafeez, A. B., Alonso, E., & Riaz, A. (2024). DTC-TranGru: Improving the performance of the next-DTC prediction model with Transformer and GRU. In Proceedings of the 39th ACM/SIGAPP Symposium on Applied Computing (SAC 2024) (pp. 927–934). Association for Computing Machinery. https://doi.org/10.1145/3605098.3635962.

11. Arévalo, P., Ochoa-Correa, D., & Villa-Ávila, E. (2024). A systematic review on the integration of artificial intelligence into energy management systems for electric vehicles: Recent advances and future perspectives. World Electric Vehicle Journal, 15(8), 364. https://doi.org/10.3390/wevj15080364.

12. Khaleghi, S., Hosen, M. S., Van Mierlo, J., & Berecibar, M. (2024). Towards machine-learning driven prognostics and health management of Li-ion batteries: A comprehensive review. Renewable and Sustainable Energy Reviews, 192, 114224. https://doi.org/10.1016/j.rser.2023.114224.

13. SAE International. (2024). J1979-DA: Digital annex of E/E diagnostic test modes. Retrieved from: https://saemobilus.sae.org/standards/j1979da_202404-j1979-da-digital-annex-e-e-diagnostic-test-modes (date accessed: May 8, 2026).

14. Zhang, J., Wang, Y., Jiang, B., et al. (2023). Realistic fault detection of Li-ion battery via dynamical deep learning. Nature Communications, 14, 5940. https://doi.org/10.1038/s41467-023-41226-5.

15. Hafeez, A. B., Alonso, E., & Riaz, A. (2022). DTCEncoder: A Swiss army knife architecture for DTC exploration, prediction, search and model interpretation. In 2022 21st IEEE International Conference on Machine Learning and Applications (ICMLA) (pp. 519–524). IEEE. https://doi.org/10.1109/ICMLA55696.2022.00085.

16. Zonta, T., da Costa, C. A., da Rosa Righi, R., de Lima, M. J., da Trindade, E. S., & Li, G. P. (2020). Predictive maintenance in the Industry 4.0: A systematic literature review. Computers & Industrial Engineering, 150, 106889. https://doi.org/10.1016/j.cie.2020.106889.

17. National Highway Traffic Safety Administration. (2024, May 13). Manufacturer communications. U.S. Department of Transportation. Retrieved from: https://www.nhtsa.gov/vehicle-manufacturers/manufacturer-communications (date accessed: June 16, 2026).

18. Yu, J., Guo, Y., & Zhang, W. (2024). Anomaly detection for charging voltage profiles in battery cells in an energy storage station based on robust principal component analysis. Applied Sciences, 14(17), 7552. https://doi.org/10.3390/app14177552.

19. Liu, H., Hao, S., Han, T., Zhou, F., & Li, G. (2023). Random forest-based online detection and location of internal short circuits in lithium battery energy storage systems with limited number of sensors. IEEE Transactions on Instrumentation and Measurement, 72, 1–11. https://doi.org/10.1109/TIM.2023.3304674.