Decoded: Vehicle State of Health using AI/ML

Jason Waterman

UC Berkeley AI/ML Professional Certificate Capstone Project / 2025

GitHub Repository: www.github.com/watermj/lv-battery-soh

The Challenge

EVs are data-rich but insight-poor. The 12V battery remains the leading cause of roadside failures — despite vehicles broadcasting thousands of CAN/LIN signals that already reflect system health.

This project explores how those existing data streams can be harnessed through AI/ML-based inference to predict degradation before failure, without adding hardware cost.

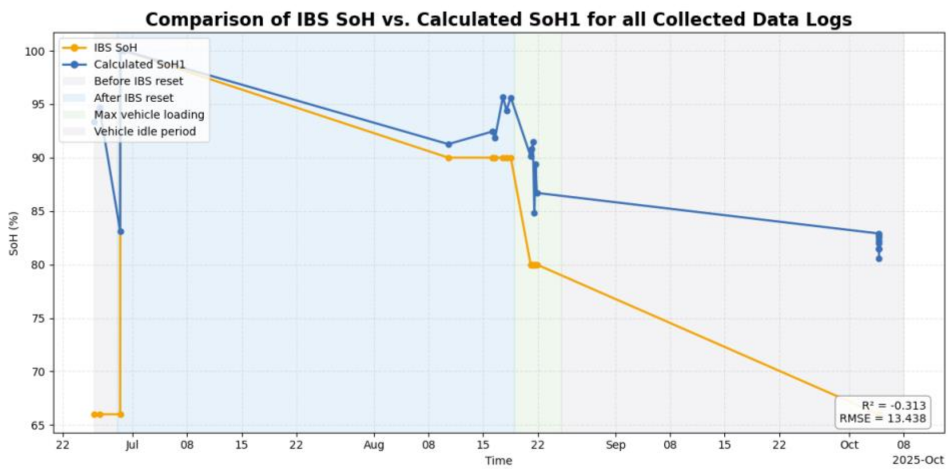


***Figure 1**: 3-month 12 V SOH log — IBS vs calculated SOH₁*

> *Calculated SOH₁ closely tracks IBS measurements across load, reset, and idle conditions, confirming feasibility of data-driven SOH inference*

The Approach

In my UC Berkeley project, I built an end-to-end **machine-learning pipeline** using in-vehicle data from a Fisker Ocean—voltage, current, temperature, and charging cycles—to model a 12V battery State of Health (SoH_1). Key steps included:

- IBS-based voltage, current, and temperature signals ("ground truth")
- Derived charge/discharge cycle metrics
- Regression and deep learning models (*Linear Regression → Decision Tree → Random Forest→ LSTM*)
- Automated feature discovery across 2,500+ vehicle network (CAN) signals to identify those co-varying with battery degradation

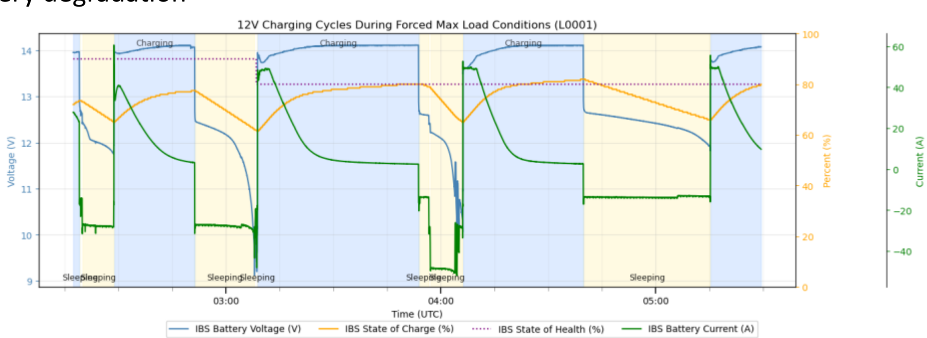
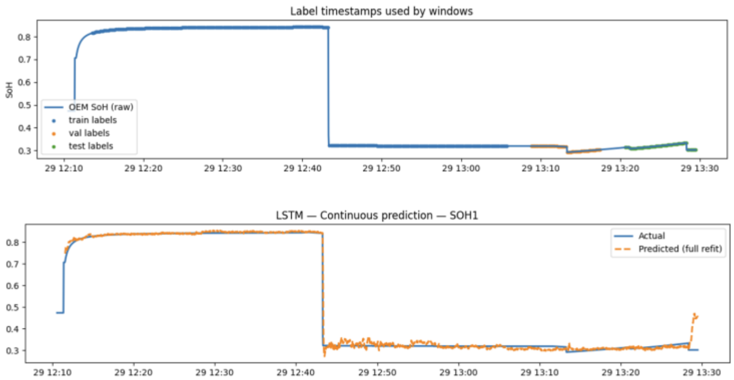


***Figure 2:** Typical 12V Charge/Discharge Cycle*

> *Charge / discharge cycle dynamics reveal recoverability dynamics which give insights into real-time 12V battery state-of-health*



Layer (type)	Output Shape	Param #	Connected to
input_layer_1 (InputLayer)	(None, 60, 747)	0	-
not_equal_1 (NotEqual)	(None, 60, 747)	0	input_layer_1[0]…
masking_1 (Masking)	(None, 60, 747)	0	input_layer_1[0]…
any_1 (Any)	(None, 60)	0	not_equal_1[0][0]
lstm_2 (LSTM)	(None, 60, 128)	448,512	masking_1[0][0], any_1[0][0]
dropout_2 (Dropout)	(None, 60, 128)	0	lstm_2[0][0]
lstm_3 (LSTM)	(None, 64)	49,408	dropout_2[0][0], any_1[0][0]
dropout_3 (Dropout)	(None, 64)	0	lstm_3[0][0]
dense_2 (Dense)	(None, 64)	4,160	dropout_3[0][0]
dense_3 (Dense)	(None, 1)	65	dense_2[0][0]

Figure 3: *LSTM Model architecture (right) and Predicted vs Actual SOH (left)*

> *LSTM achieved MAE = 0.0407 and RMSE = 0.0792 in predicting SOH trends, validating the feasibility of on-board predictive analytics*

The Opportunity

These initial results demonstrate that existing vehicle networks already contain reliable predictors of system health—including load- and sleep/wake-related signals—revealing a scalable path to **predictive vehicle intelligence**.

Operational Impact

- Early-warning diagnostics for 12V battery failures → *reduced roadside incidents and warranty costs*.
- Predictive maintenance scheduling → *lower service burden, higher fleet uptime*.
- Transferable framework for broader system health modeling (HV battery, DC/DC converters, HVAC, etc.)

Strategic Value Creation

- Integration with OTA and digital-twin ecosystems → *continuous learning from field data.*
- Foundation for monetizable Product AI features → *predictive health APIs, subscription diagnostics, fleet insights.*

What It Means for the Bottom Line

OEMs: Reduce warranty cost, enhance reliability, and readiness for predictive OTA updates.

Fleets: Real-time health dashboards and maintenance automation for uptime optimization.

Suppliers: Embed predictive diagnostics that create new service-based revenue, strengthen OEM partnerships, and unlock monetizable features in existing products.