Abstract

The transition toward sustainable mobility heavily relies on the widespread adoption of Electric Vehicles (EVs), whose performance is fundamentally dependent on efficient battery management systems (BMS). In traditional BMS approaches, challenges arise from the complex, non-linear behaviors of batteries, exacerbated by varying usage patterns, aging effects, and partial charge/discharge cycles. To address these challenges, this thesis proposes an AI-enabled digital twin framework that enhances battery life and performance in smart vehicles by holistically estimating the State of Charge (SOC), State of Energy (SOE), and State of Health (SOH) in real-time.

Building on prior research in both model-based and data-driven paradigms, the work systematically investigates and integrates advanced machine learning techniques to overcome the intrinsic variability and unpredictability of battery behavior. A key innovation lies in model-driven dataset generation and feature engineering, which ensure that data-driven models are both robust and resource-efficient. By leveraging domain knowledge of battery electrochemistry and degradation phenomena, the proposed methods extract salient features that capture subtle changes in internal battery states—even when subjected to partial discharge scenarios common in real-world EV operations.

Central to this thesis is the concept of a digital twin, which continuously synchronizes with the physical battery system through real-time sensing at the edge, capturing realistic driving profiles and operational conditions rather than relying solely on theoretical models and unrealistic scenarios like full discharge rates (SoC 100-0%). Cloud-based enable on-demand retraining and model updates. This dynamic dual-model approach ensures that SOC, SOE, and SOH estimations remain accurate and reliable throughout the battery lifecycle, seamlessly adapting to capacity fade and other aging factors. The architecture also facilitates a Pareto-based trade-off between computational overhead and estimation accuracy, guiding designers in deploying BMS solutions optimized for diverse hardware platforms.

The general framework as can be seen from figure 1, starts with training machine learning model (Lightgbm) on raw data after *segmenting it* into variable-length windows aligned to natural driving behavior. Feature engineering then extracts statistical and physics-informed indicators—e.g., dV/dt, discharge rate, mean voltage on those segmentations to estimate the SoH. Together with SoC/E dual model UpToDate model, this yields a training

corpus that spans chemistries (NMC, NCA, LFP) and ambient conditions (-10 °C – 45 °C) without exhausting laboratory resources.

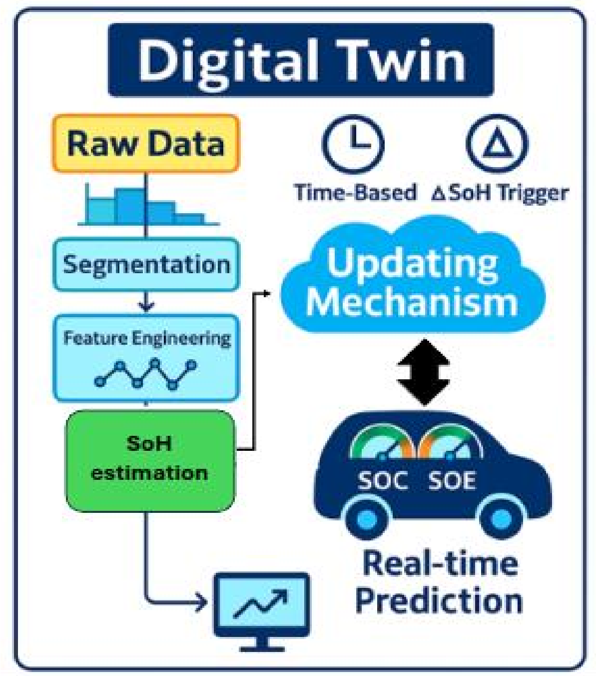

Figure 1: Methdology Overview

The proposed DT decreased the SOC/E error from 3.8 % (static modeling baseline) to 1.2 %, and SOH mean-absolute error to < 1.5%. Moreover, the thesis prototypes the edge tier on an SPC58 32-bit MCU and a RISC-V low-power core, demonstrating real-time execution (< 2 % CPU) and negligible memory inflation over legacy firmware. A four-month in-vehicle retrains the SoC/E model weekly, and shows that adaptive updates lengthen the "accurate-range" envelope by 12–15 % compared with a fixed outdated model.

By merging data-driven learning, battery-domain priors, and cloud-edge dual modeling, the proposed DT becomes a living, updating mechanism to each battery pack. Its life-cycle-aware also enables transparent residual-value scoring for second-life storage or recycling streams. In summary, this work provides a comprehensive framework that addresses critical gaps in battery state estimation—particularly those arising from dynamic real-world usage and aging. The methodologies, algorithms, and architectures proposed have broad implications for the future of smart, data-driven mobility and pave the way for more reliable, efficient, and sustainable EV ecosystems.