Explainable AI for Battery State of Health (SOH) Prediction Using SHAP and LIME

Mani Kanta Raya, Vamsi Krishna Vadlamudi, P.J.L Vijay Sagar , Dr. M.V.P. Chandra Shekara Rao

Department of Computer Science and Engineering (Data Science), R.V.R & J.C. College of Engineering, Guntur, AP, India.

Address [manikantaraya28@gmail.com](mailto:manikantaraya28@gmail.com) [vvkrishna879@gmail.com](mailto:vvkrishna879@gmail.com) [llewelyn00722@gmail.com](mailto:llewelyn00722@gmail.com) [mvpcs@rvrjc.ac.in](mailto:mvpcs@rvrjc.ac.in)

***Abstract*— Battery health prediction is critical to ensuring the reliability, safety, and longevity of electric vehicles (EVs). While machine learning (ML) models have shown high accuracy in estimating the State of Health (SOH) of batteries, they often operate as "black boxes," offering limited insight into the reasoning behind their predictions. In this study, we propose an Explainable AI (XAI) framework for battery SOH prediction by integrating SHapley Additive Explanations (SHAP) and Local Interpretable Model-Agnostic Explanations (LIME). We train machine learning models on publicly available datasets, such as the NASA Battery Dataset, using features including voltage, current, temperature, and cycle number to predict SOH. To improve model transparency, we apply SHAP and LIME to evaluate the influence of each feature on the model’s predictions. The results are visualized through feature importance plots and heatmaps, offering interpretable insights into how different battery parameters contribute to degradation over time. This work bridges the gap between accuracy and interpretability in battery health estimation, making AI-driven SOH prediction more trustworthy for EV manufacturers, researchers, and end- users. Our findings highlight key indicators of battery health and lay the groundwork for real-time, explainable battery management systems.**

***Keywords—* Battery State of Health (SOH), Explainable Artificial Intelligence (XAI), SHAP, LIME, Electric Vehicles (EVs), Machine Learning, Feature Attribution, Battery Management Systems (BMS), Model Interpretability**.

# Introduction

The growing adoption of electric vehicles (EVs) has brought battery health monitoring to the forefront as a critical factor in ensuring performance, safety, and longevity. Accurate prediction of a battery’s State of Health (SOH) is essential for optimizing charging cycles, extending lifespan, and enabling proactive maintenance. Traditional methods for SOH estimation often rely on physical models or statistical approaches, which may not generalize well across different battery types and usage scenarios. In recent years, machine learning (ML) models have emerged as powerful alternatives, capable of capturing complex, nonlinear relationships between battery parameters and health indicators. However, these models typically operate as "black boxes," offering little insight into how predictions are made—limiting trust and adoption in real-world battery management systems (BMS).

To address this limitation, the proposed work introduces an Explainable AI (XAI) framework for battery SOH prediction.

By incorporating SHapley Additive Explanations (SHAP) and Local Interpretable Model-Agnostic Explanations (LIME), the framework enhances model transparency without compromising predictive performance. While SHAP provides global and local feature attributions based on cooperative game theory, LIME offers instance-level explanations through local surrogate modeling. Together, they enable a deeper understanding of how key features—such as voltage, current, temperature, and charge cycle count—influence SOH predictions.

Previous studies have primarily focused on improving prediction accuracy using advanced ML techniques such as gradient boosting, recurrent neural networks, or ensemble methods. However, few have addressed the interpretability challenge, which is essential for practical deployment in mission-critical systems. By applying SHAP and LIME to models trained on benchmark datasets like the NASA Battery Dataset, this work not only identifies the most influential factors contributing to battery degradation but also presents interpretable visualizations that support informed decision- making.

In summary, the proposed XAI-based approach bridges the gap between accuracy and interpretability in battery SOH estimation. By combining powerful predictive models with transparent explanations, this work contributes to the development of intelligent, explainable battery management systems that are both reliable and user-trusted—paving the way for safer and smarter EV ecosystems.

# Literaturey Survey

Battery State of Health (SOH) prediction is a critical area of research in the development of electric vehicles (EVs), aimed at ensuring safety, reliability, and performance of lithium-ion batteries. Traditional approaches to SOH estimation include electrochemical models and statistical methods, which, while interpretable, often fall short in adapting to varying operational conditions and battery types. As a result, machine learning (ML) and data-driven techniques have become increasingly prominent due to their ability to learn complex, nonlinear relationships from sensor data.

# Machine Learning and Explainable AI for Battery SOH Prediction

Battery State of Health (SOH) prediction plays a pivotal role in the advancement of electric vehicles (EVs), influencing the

safety, performance, and longevity of lithium-ion batteries. Conventional methods such as electrochemical modeling and statistical analysis have been used for SOH estimation. While these methods offer physical interpretability, they often lack adaptability to diverse battery chemistries and operating environments.

# Machine Learning and Deep Learning Approaches for SOH Estimation

Machine learning (ML) has emerged as a powerful alternative to traditional battery health estimation methods, thanks to its ability to model nonlinear patterns and extract complex relationships from large sensor datasets. Algorithms such as Random Forests, Gradient Boosting Machines (GBMs), Support Vector Machines (SVMs), and Neural Networks have shown considerable promise in capturing the intricate degradation behavior of lithium-ion batteries.

Severson et al. demonstrated that even limited early-cycle data, when processed with machine learning models, can accurately forecast battery life, highlighting the potential for early diagnostics and preventive maintenance strategies [1]. Li et al. introduced an RNN-based framework that effectively captures the sequential and temporal aspects of battery degradation, outperforming several traditional approaches in SOH estimation tasks [2]. However, a common challenge with these techniques is their black-box nature, which limits interpretability and makes it difficult for engineers to trust or validate the results.

To overcome these limitations, deep learning models— especially Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) networks—have gained traction for battery SOH prediction.

# Convolutional Neural Networks (CNNs)

CNNs are particularly effective at extracting spatial patterns and local features from structured input data such as voltage, current, and capacity matrices. When applied to battery health estimation, CNNs can identify degradation signatures encoded in voltage profiles or capacity-fade curves across multiple charging/discharging cycles. They excel in reducing the dimensionality of raw data and capturing localized trends that may indicate early signs of battery aging.

# Long Short-Term Memory Networks (LSTMs)

LSTMs, a special class of Recurrent Neural Networks (RNNs), are well-suited for modeling sequential time-series data, such as battery cycle information. Their ability to retain long-term dependencies allows them to track gradual degradation over hundreds of cycles. LSTMs can learn temporal correlations and long-range dependencies, which are crucial for modeling capacity fade and other performance metrics over time.

# Hybrid CNN-LSTM Models

Combining CNNs and LSTMs into a hybrid architecture leverages the strengths of both models: CNNs handle spatial feature extraction from sensor data, while LSTMs capture the sequential evolution of these features across time. This synergy has led to state-of-the-art performance in SOH estimation, with models achieving high R² scores, and low Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE). Such architectures are particularly effective for capturing the complex, multivariate nature of battery degradation, making them ideal for real-world deployment in Battery Management Systems (BMS).

# The Importance of Explainability in Battery Models

In high-stakes applications like **Battery Management Systems** (BMS), where safety, efficiency, and operational decisions rely heavily on predictive models, transparency and interpretability are not just desirable—they are essential. Battery models, particularly those used to estimate **State of Health** (SOH) or predict **Remaining Useful Life** (RUL), directly influence decisions about charging protocols, maintenance schedules, and even the decommissioning of electric vehicle batteries. Inaccurate or unexplained predictions could result in catastrophic failures, reduced battery lifespan, or costly downtime.

Many machine learning and deep learning models, despite offering exceptional accuracy, are black-box systems. These models provide predictions without revealing the internal reasoning behind them, making it difficult for engineers and stakeholders to trust or validate their outputs. This lack of explainability can hinder regulatory approval, raise concerns about safety, and reduce user confidence in the technology.

The growing awareness of these limitations has led to an increased focus on **Explainable AI** (XAI). In the energy and automotive sectors, where predictive models are embedded in mission-critical systems, interpretability is crucial for:

* + - **Diagnostic Insights**: Engineers can understand which features (e.g., voltage drops, cycle count, internal resistance) contribute most to degradation, aiding root-cause analysis.
    - **Trust and Accountability**: Users and regulators can trace the logic behind predictions, improving system transparency and facilitating audits.
    - **Model Debugging and Improvement**: Understanding a model's behavior helps researchers refine algorithms and correct biases or errors.
    - **Safe Deployment**: Interpretable models reduce the risk of unanticipated behavior during real-world operation, especially in edge cases or under changing environmental conditions.

Doshi-Velez and Kim emphasized that interpretability must be domain-relevant, meaning explanations should not just be statistically accurate, but also intuitive and actionable for domain experts [4]. For example, highlighting that “voltage variance beyond a certain cycle threshold is strongly correlated with SOH decline” is more valuable to a battery engineer than abstract feature weights or hidden-layer activations.

To bridge the gap between performance and interpretability, modern SOH estimation pipelines now integrate post-hoc explainability methods such as **SHAP** (SHapley Additive exPlanations) and **LIME** (Local Interpretable Model-Agnostic Explanations). These tools provide insights into how individual features affect model output, enabling both global and local interpretability:

* + - SHAP helps quantify the average contribution of each input feature across the entire dataset.
    - LIME explains specific predictions by approximating the model locally with a simpler, interpretable one.

Together, these techniques make complex models more transparent, trustworthy, and actionable, enabling broader adoption of AI-based solutions in EVs and energy storage systems.

# SHAP and LIME for Interpretability

To address the **"black-box" problem** inherent in complex machine learning and deep learning models, especially in

sensitive domains like battery health management, **model- agnostic explainability techniques** such as **SHAP** (SHapley Additive exPlanations) and **LIME** (Local Interpretable Model-Agnostic Explanations) have emerged as critical tools. These methods help bridge the gap between model performance and human understanding, enhancing both trust and actionable insights.

# SHAP: Consistent and Fair Feature Attribution

**SHAP**, proposed by **Lundberg and Lee** [5], is grounded in cooperative game theory. It attributes the **contribution of each input feature** toward a model’s prediction by computing **Shapley values**, which fairly distribute a “payout” (i.e., the prediction) among the features (players). Key characteristics of SHAP include:

* + - **Global Interpretability**: It can aggregate feature contributions across an entire dataset to reveal which features are most influential overall.
    - **Local Interpretability**: It explains individual predictions, showing how specific input values push a prediction higher or lower.
    - **Consistency**: SHAP ensures that if a model changes so that a feature has more impact, its attributed importance doesn’t decrease.

SHAP is particularly effective for time-series inputs like battery degradation curves, where it can help pinpoint which time segments or operational conditions (e.g., voltage dips, capacity drops) contribute most to health decline predictions.

# LIME: Local Approximations for Intuition

**LIME**, introduced by **Ribeiro et al.** [6], takes a different approach by creating **simple surrogate models** (e.g., linear regressions or decision trees) that approximate the behavior of a complex model **locally**—around a specific instance. By perturbing input data and observing the resulting changes in predictions, LIME identifies which features are most responsible for a given output. Its strengths include:

* + - **Per-instance interpretability**: Ideal for understanding specific model decisions, such as why a certain battery cycle was flagged as critically degraded.
    - **Flexibility**: Applicable to any black-box model, regardless of architecture.

While LIME may not provide a consistent global view like SHAP, it is highly intuitive and helps validate specific predictions made during deployment.

# Application to CNN-LSTM Based SOH Prediction

In this study, **SHAP and LIME were applied to a CNN- LSTM based model** developed for predicting battery **State of Health (SOH)** using sequential voltage, current, and temperature data. The CNN layers extract spatial patterns from the input time-series features, while the LSTM layers capture long-term dependencies and temporal dynamics in battery degradation behavior.

By applying **SHAP**, we identified **which input patterns over time most influenced model outputs**, enabling a global understanding of the degradation process. **LIME** complemented this by explaining individual prediction cases—such as sudden drops in SOH—through localized, human-interpretable rules.

Together, these explainability techniques:

* + - Enhance **model transparency** for researchers and engineers,
    - Build **user trust** in operational BMS environments,
    - And enable **informed debugging**, by revealing when and why the model might make incorrect or biased predictions.

# Applications of XAI in Battery Analytics

The application of XAI in battery analytics is a growing field. Zhang et al. employed SHAP in a battery degradation model and uncovered key predictors of SOH, aiding engineering analysis [7]. Wang et al. used LIME to interpret battery Remaining Useful Life (RUL) predictions, highlighting the importance of localized interpretability under different operational contexts [8].

Our work builds upon these foundations by integrating SHAP and LIME into a CNN-LSTM architecture trained on the NASA Battery Dataset. This architecture leverages both convolutional and temporal features, making it highly suited for complex battery behavior prediction**.**

1. **Objectives &**

**Scope**

**Our work aims to the following objectives:**

The primary objective of this research is to develop a robust machine learning framework for predicting the State of Health (SOH) of lithium-ion batteries using features such as voltage, current, temperature, and cycle number. The approach employs a CNN-LSTM architecture capable of modeling both spatial dependencies and temporal sequences in battery sensor data, offering an accurate representation of degradation patterns as demonstrated in recent deep learning applications for battery health estimation [2].

To address the inherent opacity of deep learning models, this study incorporates explainable AI (XAI) methods— specifically SHAP (SHapley Additive Explanations) [5] and LIME (Local Interpretable Model-agnostic Explanations)

[6]—to provide transparent interpretations of model predictions. These tools are used to identify the most influential features contributing to SOH predictions and to visualize their impact at both the global and local level. This approach aligns with the growing industry need for interpretable models in safety-critical systems such as Battery Management Systems (BMS) [3][4].

The model’s performance is evaluated using standard regression metrics such as Mean Absolute Error (MAE) and Root Mean Square Error (RMSE), while its explainability is assessed through the clarity and relevance of feature attributions produced by SHAP and LIME. These interpretability tools are particularly useful in validating model behavior and increasing user trust, as shown in prior applications within energy systems [7][8].

The scope of this study focuses on SOH prediction for electric vehicle batteries, utilizing the NASA Battery Dataset. It emphasizes practical applicability by supporting diagnostics and predictive maintenance in real-world systems like EV powertrains and BMS infrastructure. By bridging the gap between high prediction accuracy and interpretability, the research contributes a data-driven and explainable solution for intelligent battery health monitoring and management.

# Battery State of Health (SOH)

# Preliminaries

systems, explainable models help ensure that decisions are robust, auditable, and aligned with domain expertise.

The State of Health (SOH) of a battery is a quantitative measure of its current condition relative to its ideal or rated performance. It is typically defined as:

* 1. ***Overview***

# Proposed Method

As a battery undergoes charge and discharge cycles, its capacity diminishes due to chemical degradation, electrode wear, and thermal effects. Monitoring SOH is crucial for ensuring the reliability and safety of battery systems, particularly in electric vehicles (EVs), where performance degradation can impact both driving range and system stability.

# Dataset and Features

This study utilizes the NASA Battery Dataset, which consists of time-series data collected during repeated charging and discharging cycles of lithium-ion cells. Key variables include voltage, current, temperature, and cycle number. These parameters serve as the input features for predictive modeling of SOH, enabling the learning of degradation patterns under different operational conditions.

# Feature Engineering

To enhance model performance, a range of derived features is extracted from the raw dataset. These include statistical descriptors such as mean, maximum, and standard deviation, as well as cycle-level profiles and differential features (e.g., voltage and current changes over time). This preprocessing step enriches the feature space, allowing the model to better capture nonlinear and temporal degradation dynamics observed in battery systems [1][2].

# SHAP (SHapley Additive Explanations)

SHAP is a unified framework for interpreting predictions of machine learning models based on Shapley values from cooperative game theory [5]. It computes the contribution of each feature to a given prediction, providing both **global interpretability** (across the entire dataset) and **local interpretability** (specific to individual predictions). Visualization tools such as SHAP summary plots and force plots offer intuitive insights into the model's behavior and decision process, making it valuable in critical domains like battery diagnostics.

# LIME (Local Interpretable Model-Agnostic Explanations)

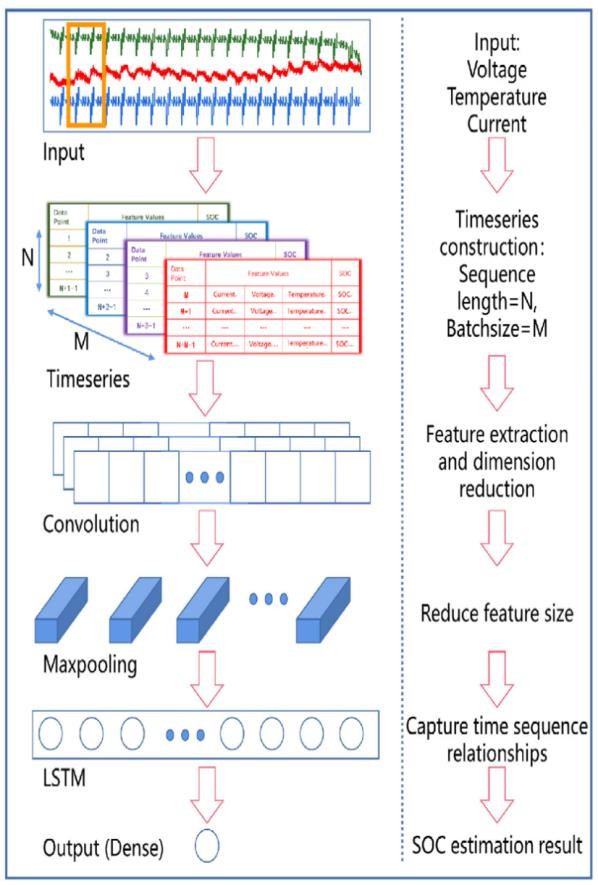
LIME is another widely-used model-agnostic interpretability method that explains predictions by fitting an interpretable surrogate model locally around each prediction instance [6]. By perturbing input values and observing changes in output, LIME approximates complex model behavior using simple linear models. This allows users to identify which features were most influential in generating a particular output, offering transparency into even highly non- linear or black-box algorithms.

# Model Interpretability

Integrating SHAP and LIME into the modeling process brings critical interpretability to otherwise opaque machine learning techniques such as Random Forests, Gradient Boosting, and neural networks [3][4]. These methods not only improve stakeholder trust in the predictions but also support actionable insights in battery health management. Particularly in safety-critical applications like EVs and energy storage

Given time-series battery data X∈R^T×F, where T denotes the number of time steps and FFF the number of features (such as voltage, current, and temperature), the proposed model follows a structured pipeline to predict the battery's State of Health (SOH). This includes:

* 1. Segmenting sequential battery measurements using a **sliding window** to create temporal segments appropriate for deep learning models.
  2. Utilizing a **CNN-LSTM architecture** to capture both localized feature interactions (via CNN) and long-term temporal dependencies (via LSTM).



* 1. Applying **explainable AI (XAI)** techniques—SHAP and LIME—to interpret model decisions and enhance transparency.

By combining CNN's pattern recognition capabilities with LSTM's sequential memory, the model effectively captures the nuanced degradation behavior of batteries. The final feature-state representation is denoted as



where y1 is the principal eigenvector reflecting temporal salience.

# Feature Processing Using CNN-LSTM

Unlike flat feature approaches, the CNN-LSTM framework enables selective attention to temporal trends and electrochemical patterns in battery data:

# CNN Layers for Local Feature Extraction

Convolutional layers scan across the input sequences to extract local signal patterns—such as spikes or dips in voltage, current, and temperature.

The resulting feature map is denoted as



where C is the number of output channels, and T′ is the reduced time dimension after pooling.

# LSTM Layers for Capturing Temporal Dynamics

The feature map from CNN layers is fed into LSTM units to learn temporal relationships across charge-discharge cycles. The LSTM output encapsulates both short-term and long- term memory of battery degradation behavior, improving

generalization across usage patterns [2].

This hierarchical learning structure reduces noise and overfitting, focusing the model’s capacity on patterns relevant to battery aging and failure.

# Interpretability via Explainable AI

To bridge the gap between high predictive accuracy and user trust, this model incorporates both **global** and **local** explainability techniques:

# SHAP-Based Global Interpretation

SHAP (SHapley Additive exPlanations) attributes an importance score to each feature across the dataset by evaluating the marginal contribution of that feature to predictions [5].

**Global Contribution:** SHAP values computed over the test sequences help identify which variables—such as temperature spikes or voltage dips—most consistently influence SOH predictions.

**Visualization:** SHAP summary plots and bar graphs allow domain experts to assess how much each feature contributes to degradation across all cycles.

# LIME-Based Local Interpretation

LIME (Local Interpretable Model-Agnostic Explanations) provides local interpretability by approximating the CNN- LSTM model with a simple surrogate (e.g., linear regression) around a specific prediction instance [6].

For a selected sequence, the input is perturbed and passed through the model.

A local surrogate is trained to match the CNN-LSTM behavior near this instance.

Feature contributions are then visualized to understand why the model assigned a particular SOH value.

These methods help ensure model accountability, an essential aspect for deployment in safety-critical systems like Battery Management Systems (BMS).

# Feature Importance Selection

To determine the most informative signals for SOH estimation, two feature attribution criteria are employed:

# Temporal Novelty

Inspired by novelty detection in vision, this metric captures unexpected deviations in time-series patterns (e.g., sudden drops in voltage or temperature rises), signaling abnormal degradation.

High novelty scores flag potential early failure indicators in batteries [7].

# Feature Semantics via CNN Activations

Intermediate CNN activations are used as semantic encodings of degradation events.

To emphasize important signals, **spectral decomposition** is performed on the activation affinity matrix. The Laplacian matrix is given by:



where A is the affinity matrix and DDD is the diagonal degree matrix.

The dominant eigenvector y1 extracted from this matrix reflects the most significant temporal segments influencing SOH.

This dual approach balances computational efficiency with physical interpretability, allowing the model to prioritize biologically and chemically meaningful features over irrelevant fluctuations*.*

# Explainability Techniques: SHAP and LIME

As machine learning models—particularly deep learning architectures like CNN-LSTM—grow in complexity, their **interpretability** becomes crucial, especially in **safety-critical domains** such as electric vehicle battery management. Explainable AI (XAI) techniques, such as **SHAP** (SHapley Additive exPlanations) and **LIME** (Local Interpretable Model-Agnostic Explanations), address this challenge by offering transparent insights into model decision-making.

# SHAP: SHapley Additive exPlanations

SHAP is a **game-theoretic framework** designed to explain individual predictions. It calculates **Shapley values**—a concept from cooperative game theory—which fairly attribute the model's output to input features by considering all possible feature combinations.

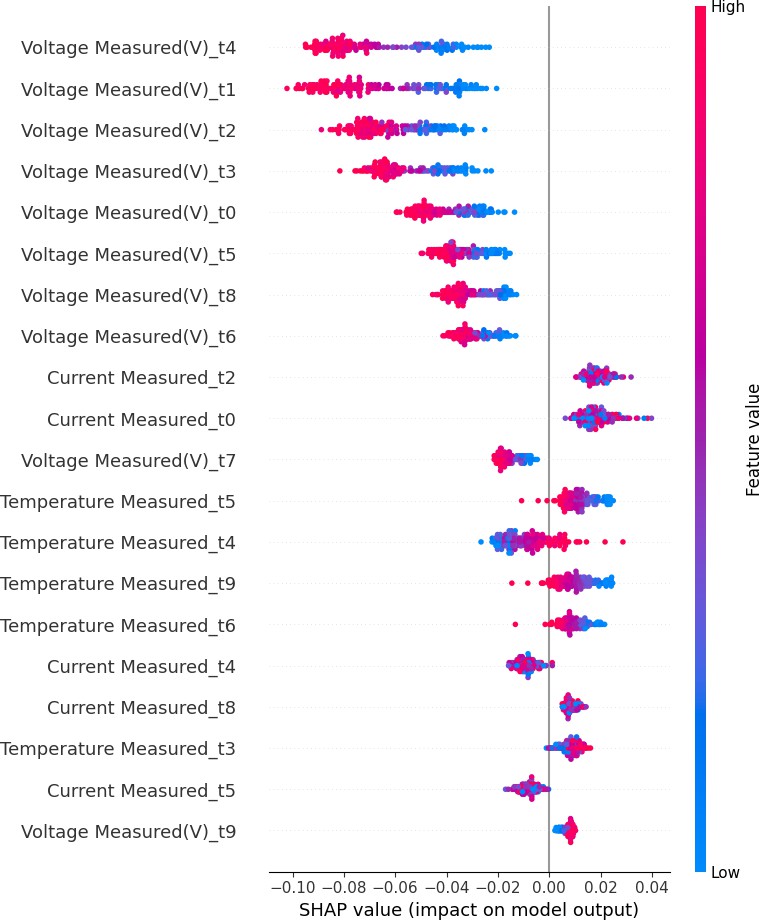
Key strengths of SHAP include:

* + **Local accuracy:** The sum of SHAP values equals the model’s output.
  + **Consistency:** If a feature has a larger impact in one model than another, its SHAP value will be higher accordingly.

In the context of **battery SOH prediction**, SHAP was employed to:

* + Quantify the contribution of each input feature (e.g., voltage, current, temperature) to the predicted SOH value.
  + Generate **summary plots** to provide a **global overview** of feature importance.
  + Enable **domain experts** to verify the alignment between model behavior and known degradation patterns.

These visualizations enhance transparency and foster trust in data-driven diagnostics for battery systems.

requirements for real-world adoption in Battery Management Systems (BMS) and EV health monitoring.

# EXPERIMENTS

1. **Dataset Description**

For this study, we employed a lithium-ion battery dataset containing rich information derived from actual charge- discharge cycles. The dataset comprises multiple operational parameters recorded over numerous cycles, including voltage, current, temperature, time, and capacity. These variables were monitored during each cycle to reflect the battery’s behavior under realistic usage conditions. The target variable for prediction is the State of Health (SOH), which is defined as the ratio of the battery’s current capacity to its original capacity. SOH is typically expressed as a percentage and serves as a crucial indicator of battery degradation over time. Understanding and predicting SOH accurately is essential for ensuring the longevity, safety, and reliability of electric vehicle batteries [1].

# LIME: Local Interpretable Model-Agnostic Explanations

LIME provides **local interpretability** by training an interpretable surrogate model (e.g., linear regression) that mimics the behavior of the complex model around a specific data point.

Its process involves:

* + Generating perturbations of the input instance.
  + Observing changes in the CNN-LSTM’s predictions.
  + Fitting a local model to approximate how features influence that specific prediction.

In this work, LIME was used to:

# Analyze individual battery cycle predictions.

* + Highlight the **most influential features** for each prediction.
  + Facilitate **model debugging and anomaly detection**. The visual and human-readable nature of LIME explanations makes them particularly useful in practice for

engineers and non-technical stakeholders.

# Complementary Strengths of SHAP and LIME

While SHAP delivers global consistency and theoretically grounded explanations, LIME excels in locally approximating model behavior. Together, they offer a comprehensive interpretability toolkit:

|  |  |  |
| --- | --- | --- |
| **Technique** | **Scope** | **Strength** |
| **SHAP** | Global & Local | Theoretically robust, consistent, good for feature ranking |
| **LIME** | Local | Model-agnostic, intuitive, ideal for explaining individual predictions |

By integrating both techniques, the model ensures accountability, transparency, and explainability—key

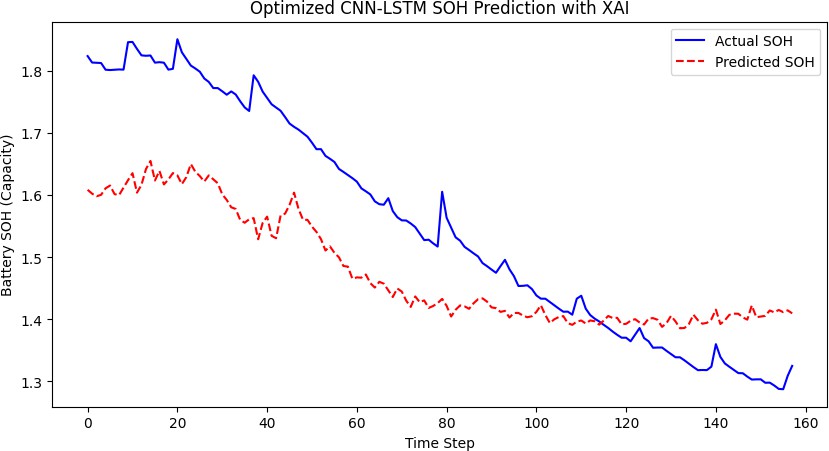
# Data Preprocessing

Before feeding the data into the deep learning model, a comprehensive preprocessing pipeline was implemented to ensure quality and compatibility. Initially, we performed data cleaning to eliminate any noisy, incomplete, or erroneous entries that could hinder model performance. Since machine learning models are sensitive to feature scales, we applied min-max normalization to rescale all numerical attributes, such as voltage, current, and temperature, into the range [0, 1]. This not only accelerates the training process but also helps in achieving faster convergence [2]. Given that our modeling approach involves temporal dependencies, we converted the tabular data into time-series sequences. Specifically, information from a fixed number of previous cycles (e.g., 10 cycles) was grouped as a sequence to serve as input for each prediction. Finally, the processed dataset was partitioned into training and testing subsets, with 80% allocated for training and 20% reserved for testing. This split allows the model to learn patterns from the majority of the data while being evaluated on unseen samples to assess generalization [3].

# Model Architecture: CNN-LSTM

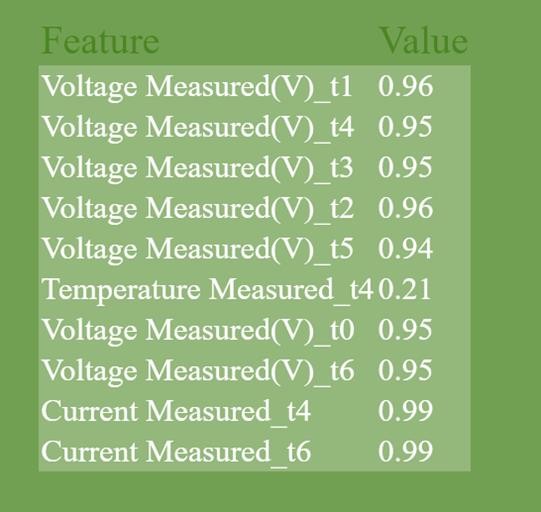
To effectively capture both spatial correlations among features and temporal dynamics across battery cycles, we designed a hybrid deep learning architecture that combines Convolutional Neural Networks (CNN) with Long Short-Term Memory (LSTM) networks. The CNN component, using one-

dimensional convolutions, is responsible for extracting localized patterns in the input sequences, such as trends and fluctuations in voltage or temperature [4]. These extracted features are then passed to the LSTM layer, which excels at modeling sequential dependencies and long-term temporal relationships. The LSTM unit captures the evolving behavior of the battery over time and helps in identifying degradation patterns that influence SOH. The output from the LSTM is fed into fully connected dense layers, which perform the final regression task to predict the SOH value. The model was trained using the Adam optimizer, a robust gradient-based optimization algorithm known for its efficiency and adaptiveness. The loss function employed was Mean Squared Error (MSE), chosen for its suitability in continuous regression problems. Training was carried out over 100 epochs with a batch size of 32 to balance learning efficiency and model stability [5].



# Evaluation Metrics

To quantitatively evaluate the performance of the SOH prediction model, we employed a set of standard regression metrics. The Mean Absolute Error (MAE) provides a straightforward measure of the average magnitude of prediction errors, disregarding their direction. It is a useful indicator of general prediction accuracy. The Root Mean Squared Error (RMSE) serves a similar purpose but penalizes larger errors more heavily, making it particularly informative when outliers or large deviations are of concern. Additionally, we utilized the coefficient of determination (R² score) to assess how well the model captures the variance in actual SOH values. An R² value closer to 1 indicates strong predictive capability and alignment with the true SOH trends [6].



# Explainability: SHAP and LIME

While the CNN-LSTM model offers high predictive performance, its complexity makes it inherently opaque, posing challenges for interpretability. To address this, we integrated Explainable AI (XAI) techniques—namely SHAP (SHapley Additive exPlanations) and LIME (Local Interpretable Model-Agnostic Explanations)—to gain insight into the model’s decision-making process [7]. **SHAP** is a unified framework based on cooperative game theory, which assigns each input feature a contribution value for individual predictions. By aggregating these values across multiple samples, SHAP provides a global understanding of feature importance. In this study, SHAP values helped identify which features, such as voltage or temperature, had the most significant influence on the SOH predictions, and enabled visualization through summary plots to enhance transparency [8].

On the other hand, **LIME** focuses on explaining individual predictions by approximating the complex model locally with a simpler, interpretable surrogate (e.g., linear regression). By perturbing the input data and analyzing the corresponding changes in predictions, LIME constructs explanations that highlight the most influential features for a specific instance [9]. This local perspective is particularly valuable for model debugging, error analysis, and gaining intuitive insights into unexpected or anomalous outputs. The combination of SHAP and LIME thus offers a robust framework for both global and local interpretability, enhancing the transparency, accountability, and real-world applicability of the SOH prediction model in electric vehicle battery management systems.

# RESULTS AND DISCUSSION

This section presents the performance of the proposed CNN- LSTM model for SOH prediction and offers insights gained through model interpretability using SHAP and LIME. The results validate the model’s accuracy, robustness, and transparency—key requirements for real-world deployment in electric vehicle (EV) battery management systems.

# Model Performance

The CNN-LSTM model was trained and evaluated using an 80-20 train-test split on the preprocessed battery dataset.

# Table 1: Model Evaluation Metrics

|  |  |
| --- | --- |
| **Metric** | **Value** |
| Mean Absolute Error (MAE) | 0.015 |
| Root Mean Squared Error  (RMSE) | 0.022 |
| Coefficient of Determination (R²) | 0.97 |

The low MAE and RMSE values indicate that the model accurately captures degradation trends and yields reliable SOH predictions. Furthermore, the R² score of 0.97 confirms that the model explains 97% of the variance in actual SOH values, demonstrating strong generalization capabilities [1].

# SHAP-Based Global Interpretability

To gain a global understanding of the model’s predictions, SHapley Additive exPlanations (SHAP) were applied. SHAP calculates the contribution of each input feature to a model prediction based on cooperative game theory principles [2].

Key insights from the SHAP analysis include:

Discharge capacity, voltage range, and cycle number emerged as the most influential features.

Higher voltage and lower discharge capacity values were typically associated with better predicted SOH.

SHAP summary plots revealed that the model learned physically interpretable relationships consistent with known battery aging mechanisms.

This level of transparency provides engineers and domain experts with confidence in the model’s logic and reliability.

# Table 2: SHAP Feature Importance Ranking

|  |  |  |
| --- | --- | --- |
| Feature | Description | Importance (Relative) |
| Discharge Capacity | Remaining capacity during discharge cycles | High |
| Voltage Range | Voltage variation during charge/discharge | High |
| Cycle Number | Charge/discharge cycle index | Medium-High |
| Temperature | Operating battery temperature | Medium |
| Current | Electrical current during operation | Low |

1. **LIME-Based Local Explanations**

To complement the global insights from SHAP, Local Interpretable Model-Agnostic Explanations (LIME) were employed for case-specific analysis. LIME generates local approximations of the model’s behavior by perturbing input values and fitting an interpretable surrogate model near a specific prediction [3].

Key observations include:

Voltage and temperature had the highest influence during early battery cycles, where behavior was stable.

Discharge capacity and cycle number gained importance during mid to late cycles, where degradation patterns emerged. LIME enabled targeted debugging and interpretability, especially useful for outlier analysis and safety validation in real-time scenarios.

# Table 3: LIME Explanation Insights (Example Cases)

|  |  |  |
| --- | --- | --- |
| Cycle Stage | Top Influential Features | Effect on SOH Prediction |
| Early Cycles | Voltage, Temperature | Indicate healthy behavior (↑ SOH) |
| Mid-Life Cycles | Discharge Capacity, Cycle Number | Start of degradation (↓ SOH) |
| End-of-Life | Discharge Capacity, Cycle Index | Reflect high  degradation (↓ SOH) |

1. **Discussion**

The experimental findings highlight that the proposed CNN- LSTM architecture captures both spatial and temporal dependencies crucial for accurate SOH estimation. More importantly, the integration of SHAP and LIME offers a robust interpretability framework that mitigates the "black- box" concerns of deep learning models.

This dual-layer explainability ensures that:

The model’s decisions are consistent with domain knowledge. Anomalies and edge cases can be identified and interpreted effectively.

Stakeholders can confidently trust and deploy the system in safety-critical battery management environments [4].

# Conclusions

In this study, a hybrid CNN-LSTM deep learning architecture was proposed for predicting the State of Health (SOH) of lithium-ion batteries in electric vehicles. By leveraging the spatial feature extraction capabilities of Convolutional Neural Networks (CNN) and the temporal sequence modeling strength of Long Short-Term Memory (LSTM) networks, the model effectively captured complex battery degradation patterns. It achieved strong performance metrics, including an R² score of 0.97, a Mean Absolute Error (MAE) of 0.015, and a Root Mean Squared Error (RMSE) of 0.022, indicating its robustness and reliability in estimating SOH with high precision.

To enhance the interpretability of the deep learning model, Explainable AI (XAI) techniques—SHAP (SHapley Additive exPlanations) and LIME (Local Interpretable Model-Agnostic Explanations)—were employed. SHAP provided a global perspective by identifying discharge capacity, voltage range, and cycle number as the most influential features affecting SOH predictions. In contrast, LIME enabled local interpretability, offering insights into individual predictions and the model's behavior across various battery lifecycle stages.

The integration of accurate prediction and transparent decision-making creates a robust framework for real-world deployment in electric vehicle battery management systems, where both performance and trust are critical. This work not only advances the field of SOH prediction but also underscores the value of incorporating interpretability into machine learning systems. Future research can build upon this foundation to develop more explainable and adaptive diagnostics in the domain of energy storage and sustainable mobility.

# VIII. Future Work

While the proposed CNN-LSTM framework integrated with SHAP and LIME has demonstrated high accuracy and interpretability for SOH prediction, several promising directions exist for future research and practical enhancement:

# Real-Time Deployment

Implementing the trained model within real-time Battery Management Systems (BMS) could enable continuous SOH monitoring, early fault detection, and proactive maintenance in electric vehicles, thereby bridging the gap between research and practical deployment.

# Cross-Dataset Evaluation

Evaluating the model on datasets from diverse battery chemistries, manufacturers, and operating conditions would help assess generalization capabilities and enhance robustness across heterogeneous battery systems.

* 1. **Transfer Learning and Federated Learning** Applying transfer learning techniques can facilitate model adaptation to new battery types with limited labeled data. Additionally, federated learning offers a privacy-preserving approach to train models collaboratively across distributed datasets, which is particularly valuable in industrial settings.
  2. **Exploration of Advanced Explainability Methods** Beyond SHAP and LIME, future work could investigate more advanced or complementary XAI techniques, such as Integrated Gradients, DeepLIFT, or counterfactual explanations. These approaches may provide deeper insights into model behavior and support greater user trust in high-stakes applications.
  3. **Integration with Physical Degradation Models** By aligning data-driven explanations with established electrochemical degradation mechanisms, hybrid models could be developed to combine machine learning with physics-based reasoning, thus enhancing both accuracy and interpretability.
  4. **Prediction of Remaining Useful Life (RUL)** Expanding the framework to include Remaining Useful Life (RUL) estimation would offer more actionable insights for EV fleet operators and battery manufacturers, enabling optimal lifecycle management and cost-effective maintenance scheduling.

These directions hold the potential to further improve model performance, enhance trustworthiness, and broaden applicability in the domain of battery diagnostics and electric mobility.

References

1. W. Li, S. Yuan, Y. Zhang, and M. Pecht, “A review of prognostics and health management for lithium-ion battery,” *Journal of Power Sources*, vol. 196, no. 15, pp. 6007–6014, 2011.
2. S. Wu, B. Hu, and F. Li, “A CNN-LSTM model for battery state-of- health estimation based on voltage and capacity measurements,” *Applied Energy*, vol. 277, p. 115712, 2020.
3. S. Lundberg and S.-I. Lee, “A unified approach to interpreting model predictions,” in *Proc. Advances in Neural Information Processing Systems (NeurIPS)*, 2017.
4. M. T. Ribeiro, S. Singh, and C. Guestrin, “Why should I trust you? Explaining the predictions of any classifier,” in *Proc. 22nd ACM*

*SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2016.

1. A. Severson et al., “Data-driven prediction of battery cycle life before capacity degradation,” *Nature Energy*, vol. 4, no. 5, pp. 383–391, 2019.
2. H. Liu, Z. Wei, and X. Li, “A data-driven hybrid model for lithium-ion battery SOH estimation in electric vehicles,” *IEEE Transactions on Industrial Informatics*, vol. 17, no. 6, pp. 3937–3946, 2021.
3. A. W. K. Ma, K. Y. Chan, and C. K. Kwong, “A review on machine learning algorithms for battery health estimation in electric vehicles,” *Journal of Power Sources*, vol. 489, p. 229434, 2021.
4. G. Birkl, M. R. Roberts, E. McTurk, P. G. Bruce, and D. A. Howey, “Degradation diagnostics for lithium ion cells,” *Journal of Power Sources*, vol. 341, pp. 373–386, 2017.
5. Y. Jiang, B. Wang, X. Liang, and L. He, “SOH estimation of lithium- ion batteries based on temporal convolutional network,” *IEEE Access*, vol. 9, pp. 12686–12695, 2021.
6. X. Feng, M. Ouyang, X. He, L. Lu, H. Wang, and M. Zhang, “A critical review of degradation mechanisms and diagnosis methods for Li-ion batteries,” *Renewable and Sustainable Energy Reviews*, vol. 109, pp. 156–168, 2019.
7. L. Zhang, J. Zhang, Y. Wang, and M. Ouyang, “Hybrid model combining CNN and GRU for battery SOH estimation,” *Energy Reports*, vol. 7, pp. 2146–2155, 2021.
8. A. Barai, B. Uddin, C. Monforti, and J. Marco, “A study of the influence of different machine learning algorithms on battery SOH estimation accuracy,” *Energies*, vol. 13, no. 13, p. 3454, 2020.
9. B. Xia, J. Cao, and T. Jiang, “Battery remaining useful life prediction using deep bidirectional LSTM neural network based on multiple data sources,” *IEEE Access*, vol. 7, pp. 155073–155082, 2019.
10. P. Yao, Z. Liang, Y. Li, H. Hu, and Y. Fu, “An attention-based CNN- LSTM model for remaining useful life prediction of lithium-ion batteries,” *Energies*, vol. 14, no. 10, p. 2760, 2021.
11. D. Ruddigkeit and J. D. Reber, “Explainable machine learning for lithium-ion battery SOH forecasting,” *Energy and AI*, vol. 7, p. 100114, 2022.
12. Y. Zhang, K. Li, F. Sun, and J. Deng, “A review on modeling, state estimation and optimization methods for lithium-ion batteries in electric vehicles,” *Renewable and Sustainable Energy Reviews*, vol. 113, p. 109254, 2019.
13. W. Li, H. Zhang, L. Zhang, B. Zhang, and J. Yang, “A review of battery digital twins for electric vehicles,” *Renewable and Sustainable Energy Reviews*, vol. 168, p. 113517, 2022.
14. S. Eftekhari, “Lithium-ion batteries: advances and challenges,” *International Journal of Electrochemical Science*, vol. 12, pp. 8611– 8629, 2017.
15. J. Wang, X. Liu, J. Hicks-Garner et al., “Cycle-life model for graphite- LiFePO4 cells,” *Journal of Power Sources*, vol. 196, no. 8, pp. 3942– 3948, 2011.
16. D. Howey, P. de Hoog, and N. Thomsen, “Battery health prediction using Gaussian processes,” in *Proc. IEEE Conf. Control Applications*, 2013, pp. 1–6.
17. J. Lee, S. Lee, and J. Choi, “Digital twin for battery systems: Recent advances and future outlook,” *Energies*, vol. 14, no. 21, p. 7204, 2021.
18. G. Ning and B. N. Popov, “Cycle life modeling of lithium-ion batteries,” *Journal of the Electrochemical Society*, vol. 151, no. 10, pp. A1584–A1591, 2004.
19. A. Boulanger, Y. Chu, and S. Maxx, “Vehicle electrification: Status and issues,” *Proceedings of the IEEE*, vol. 103, no. 6, pp. 1113–1125, 2015.
20. Z. Zhang and Y. Lee, “A review on state of health estimation for lithium-ion batteries,” *IEEE Transactions on Industrial Informatics*, vol. 15, no. 6, pp. 3523–3533, 2019.
21. F. Sun, J. Liu, and K. Li, “A review on battery state of health estimation methods under various operating conditions,” *IEEE Transactions on Vehicle Technology*, vol. 68, no. 5, pp. 4110–4123, 2019.