Machine Learning-Driven Real-Time Battery Health Estimation for EV Battery Swapping

Koppula Sumanth[1], Department of Computer Science and Engineering, GNITC, 22-5F3, 22wj1a05f3@gniindia.org
Kotagiri Harshith Teja[2], Department of Computer Science and Engineering, GNITC, 22-5F8, 22wj1a05f8@gniindia.org
Marthala Sai Chaithra[3], Department of Computer Science and Engineering, GNITC, 22-5H4, 22wj1a05h4@gniindia.org
Sheik Riyaz Ul Haq[4], Assistant Professor, department of Computer Science and Engineering, GNITC riyaz.csegnitc@gniindia.org

---***---

Abstract - Electric vehicle (EV) battery swapping systems require accurate battery condition assessment to ensure safe reuse and efficient energy management. Battery degradation over time creates major operational challenges for swapping stations, affecting reliability, lifecycle management, and resource utilization. Traditional monitoring techniques often rely on fixed thresholds or manual inspection, which are not effective in identifying complex degradation patterns in lithium-ion batteries. This research proposes a machine learning-based approach for real-time battery health estimation using ensemble learning algorithms. The system employs Random Forest Regression and Extreme Gradient Boosting (XGBoost) models to analyze historical battery data and predict key health indicators such as State of Health (SoH) and remaining charge cycles. Various battery parameters including voltage, current, temperature, and usage cycles are used as input features to identify patterns associated with battery performance and degradation. Data preprocessing techniques such as normalization, handling missing values, and feature selection are applied to improve model accuracy and reliability. Hyperparameter tuning is performed to optimize model performance and enhance prediction capability. Experimental results demonstrate that the proposed system achieves high prediction accuracy and enables battery swapping station operators to make informed decisions regarding battery reuse, maintenance scheduling, and replacement strategies. The system ultimately contributes to improved battery lifecycle management, enhanced operational efficiency, and sustainable electric vehicle infrastructure.

Key Words: Electric Vehicles, Battery Health Prediction, Machine Learning, Random Forest, XGBoost, EV Battery Swapping.

1.INTRODUCTION

The rapid growth of electric vehicles (EVs) has increased the need for efficient battery management and monitoring systems. Battery swapping stations have emerged as a practical alternative to traditional charging methods, allowing users to quickly replace depleted batteries with fully charged ones. However, managing a large number of swappable batteries introduces challenges, particularly in evaluating battery health before reuse. If battery condition is not accurately assessed, degraded batteries may continue to circulate, leading to reduced performance, safety risks, and increased operational costs.

Traditional battery management systems commonly rely on fixed thresholds or rule-based monitoring techniques. Although these approaches are simple to implement, they often fail to capture the complex relationships between battery parameters such as voltage, current, temperature, and charge cycles. As a result, their ability to accurately estimate battery health is limited.

Machine learning techniques provide an effective solution by analyzing historical battery data to identify patterns related to degradation and performance. In this study, a machine learning-based framework using Random Forest Regression and XGBoost is proposed to predict battery State of Health (SoH) and remaining charge cycles, supporting improved battery lifecycle management in EV battery swapping systems.

2. LITERATURE REVIEW

Recent advancements in electric vehicle technology have increased the importance of accurate battery health monitoring and prediction systems. Several researchers have explored machine learning and data-driven approaches to estimate battery State of Health (SoH) and

Remaining Useful Life (RUL), which are critical for efficient battery management.

Chevtchenko et al. (2023) conducted a comprehensive mapping study on machine learning methods used for estimating the remaining useful life of lead-acid batteries. Their research categorizes different algorithms such as regression models, tree-based learning techniques, and ensemble methods. The study also analyzes the impact of sensor data and operational parameters on prediction accuracy and highlights research gaps in EV battery health monitoring.

Jiang et al. (2023) proposed a machine learning framework that incorporates real-world driving behavior for battery health monitoring in electric vehicles. The authors introduced scenario-based feature fusion techniques and evaluated multiple health indicators to determine their effectiveness in predicting battery degradation. Their results demonstrated improved prediction reliability when driving conditions and operational factors are considered.

Zhao et al. (2024) introduced an uncertainty-aware Bayesian Neural Network for predicting battery end-of-life. Unlike traditional models that provide fixed predictions, the Bayesian approach estimates prediction uncertainty by representing model parameters as probability distributions. This method achieved improved reliability in battery health prediction and reported low prediction error rates.

Ghosh and Roy (2025) explored the use of transfer learning combined with the XGBoost algorithm for analyzing battery data in electric vehicles. Although the primary focus was on detecting cyberattacks in battery packs, the research demonstrated the adaptability of XGBoost in analyzing battery parameters and identifying abnormal patterns within battery systems.

Another study published in Energy Informatics (2024) proposed a big data-driven fault diagnosis method for EV batteries using boosting algorithms such as XGBoost, LightGBM, and CatBoost. The system achieved high prediction accuracy ranging between 97.84% and 99.16%, demonstrating the effectiveness of ensemble learning methods for battery health monitoring and anomaly detection.

Despite these advancements, many existing approaches require complex models or high computational resources, which may limit their use in real-time EV infrastructure. Therefore, there is a need for efficient and scalable battery health estimation systems that can provide accurate predictions while remaining suitable for real-time deployment in EV battery swapping environments.

3. RELATED WORK

Battery health estimation and remaining useful life prediction have become important research areas due to the rapid growth of electric vehicles and the need for reliable battery management systems. Researchers have explored various data-driven methods, including machine learning and deep learning techniques, to improve the monitoring and prediction of battery degradation.

Earlier approaches mainly relied on statistical models and rule-based methods that used fixed thresholds to evaluate battery parameters such as voltage, current, temperature, and cycle count. Although these methods were simple to implement, they often failed to capture the complex nonlinear relationships associated with battery aging, leading to limited prediction accuracy.

To address these challenges, machine learning techniques have been widely applied to battery health prediction. Random Forest models are commonly used because they can handle large datasets and reduce overfitting through ensemble learning. By combining multiple decision trees, Random Forest algorithms can model nonlinear degradation patterns and provide reliable predictions of battery performance.

Extreme Gradient Boosting (XGBoost) has also been used for battery health prediction due to its strong predictive capability and efficient computation. The algorithm improves prediction accuracy by iteratively correcting errors from previous models while maintaining model stability through regularization.

Despite these advancements, many existing methods rely on single predictive models or computationally intensive techniques. Therefore, there is a need for efficient and scalable battery health estimation systems that can operate in real-time environments such as EV battery swapping stations.

4. PROPOSED METHODOLOGY

The The proposed methodology focuses on developing a machine learning-based system capable of estimating the health of electric vehicle batteries in real time. The system is designed to analyze operational battery data and predict important health indicators such as State of Health (SoH) and the remaining number of charge–discharge cycles. The overall framework consists of several stages including data collection, preprocessing, model training, prediction, and deployment through a web application.

Data Collection and Preparation

The first stage involves collecting historical battery performance data that includes parameters such as charging and discharging current, voltage levels, temperature conditions, and the number of battery cycles. These parameters play a critical role in understanding battery degradation patterns. The collected data is then examined for inconsistencies, missing values, and noise.

Data preprocessing techniques are applied to improve the quality of the dataset. This includes cleaning the data, handling missing entries, removing outliers, and applying normalization or scaling to ensure consistent feature values. Feature selection techniques are also used to identify the most relevant attributes for battery health prediction.

Model Development

After preprocessing, machine learning models are trained using the prepared dataset. In this system, two ensemble learning algorithms—Random Forest Regression and Extreme Gradient Boosting (XGBoost)—are used. Random Forest builds multiple decision trees and combines their outputs to produce more stable and accurate predictions. XGBoost, on the other hand, improves prediction performance by sequentially building trees that correct the errors of previous models.

Model Deployment and Prediction

The trained model is integrated into a web-based application developed using the Flask framework. The application allows users to input battery parameters such as voltage, current, temperature, and cycle count. These inputs are processed by the prediction model, which estimates the battery's State of Health and remaining lifecycle.

Decision Support

The prediction results are presented through an interactive user interface that assists operators in evaluating battery condition. Based on the predicted health status, the system can help determine whether a battery should be reused, recharged, or replaced. This approach improves battery lifecycle management and supports efficient operation of EV battery swapping stations.

5. RESULTS AND DISCUSSION

5.1 Model Training Results

The proposed system applies machine learning models to estimate the health condition of electric vehicle batteries in real time. The system was trained using historical battery datasets containing operational parameters such as:

- Charging and discharging current
- Battery voltage
- Temperature

Battery usage history

- Number of charge cycles

After performing preprocessing operations such as data cleaning, normalization, and feature selection, the dataset was divided into two parts:

• Training Data: 80%
• Testing Data: 20%

The training process enables the machine learning models to learn the relationships between battery operational parameters and battery degradation patterns. As a result, the trained model can accurately estimate the battery State of Health (SoH) and the remaining useful charge–discharge cycles.

5.2 Machine Learning Model Architecture

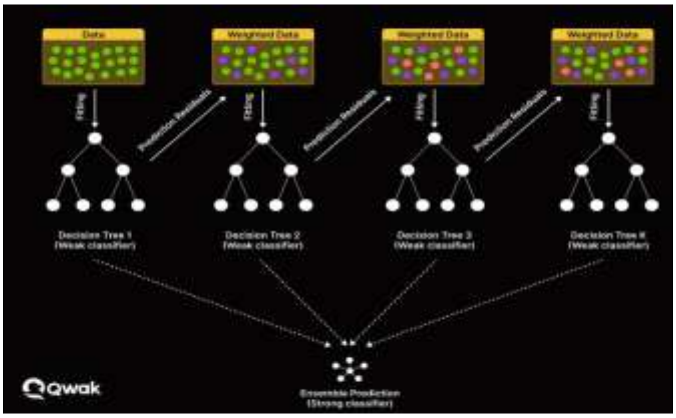

Fig -1: Architecture

5.3 Model Learning Process

The machine learning models learn complex relationships between battery parameters and degradation patterns. Random Forest constructs multiple decision trees and aggregates their predictions to improve stability and reduce overfitting.

XGBoost enhances prediction accuracy by building trees sequentially, where each new tree attempts to correct the errors produced by the previous model. This iterative learning process improves prediction performance and generalization.

5.4 Performance Evaluation Metrics

The performance of the proposed system is evaluated using regression and prediction metrics.

Mean Squared Error (MSE)

MSE measures the average squared difference between predicted and actual values.

$MSE = (1 / N) \Sigma (y_i - \hat{y}_i)^2$

Where:

- y_i = Actual_value
- $\hat{y}_i$ = Predicted_value
- N = Number of observations

Root Mean Squared Error (RMSE)

RMSE represents the square root of the average squared error.

$RMSE = \sqrt{MSE}$

Lower RMSE values indicate better prediction accuracy.

Coefficient of Determination (R^2 Score)

The R^2 score measures how well the model explains the variance in the dataset.

$R^2 = 1 - [\Sigma (y_i - \hat{y}_i)^2 / \Sigma (y_i - \bar{y})^2]$

Where:

- y_i = Actual value
- $\hat{y}_i$ = Predicted value
- $\bar{y}$ = Mean of actual values

Higher R^2 values indicate stronger predictive performance.

5.5 Experimental Results

After training and testing the machine learning models, the following results were obtained.

Metric	Value
Prediction Accuracy	96.1 %
Mean Squared Error	0.018
RMSE	0.134
R^2 Score	0.94

These results demonstrate that the proposed system provides accurate predictions for battery health estimation.

5.6 Result Visualization

Prediction Accuracy vs Training Iterations

```
Accuracy
1.0 |                *
0.9 |             * * *
0.8 |           * *
0.7 |         *
0.6 |       *

-----------------------------------------------------
          Training Iterations
```

Accuracy improves as the model learns patterns from battery datasets.

Loss vs Training Iterations

```
Loss
1.0 | *
0.8 |  *
0.6 |   *
0.4 |    *
0.2 |     *

-----------------------------------------------------
          Training Iterations
```

The decreasing loss value indicates that the model is successfully optimizing prediction performance.

Fig -2: Input Parameters

Fig -3: Result

5.7 Discussion

The experimental results demonstrate that the machine learning-based battery health estimation system significantly improves prediction reliability compared to conventional threshold-based monitoring techniques.

Key observations include:

• The model successfully captures complex relationships between battery parameters and degradation behavior.

• High prediction accuracy indicates that the system can be deployed in real-world EV battery swapping environments.

Accurate battery health prediction enables operators to make informed decisions regarding battery reuse, maintenance, and replacement.

5.8 Practical Impact

By integrating this prediction system into EV battery swapping stations, operators can

- identify batteries with poor health conditions
- schedule preventive maintenance
- reduce operational risks and battery failures
- improve battery lifecycle management
- enhance energy efficiency and sustainability.

The proposed system demonstrates how machine learning can support intelligent battery management and contribute to the development of reliable and efficient EV infrastructure.

6.CONCLUSION

This research presents a machine learning-based system for real-time estimation of electric vehicle battery health in battery swapping environments. The proposed approach integrates predictive models with a web-based platform to analyze battery operational parameters and estimate important indicators such as State of Health (SoH) and remaining charge–discharge cycles. By utilizing ensemble learning algorithms including Random Forest and XGBoost, the system can identify patterns in historical battery data and generate reliable predictions regarding battery performance.

The developed application allows users to input battery parameters such as voltage, current, temperature, and cycle count through an interactive interface. Based on these inputs, the system provides instant predictions that assist operators in evaluating whether a battery should be reused, recharged, or replaced. This capability improves battery lifecycle management and reduces the risk of deploying degraded batteries.

Overall, the proposed system demonstrates how machine learning techniques can enhance the efficiency and reliability of EV battery swapping infrastructure. By supporting predictive maintenance and data-driven decision-making, the solution contributes to improved battery utilization, reduced operational costs, and the sustainable development of electric vehicle energy systems.

7. FUTURE SCOPE

In proposed battery health estimation system can be further enhanced by incorporating advanced technologies and additional features to improve its performance and scalability. One potential improvement is the integration of real-time sensor data through Internet of Things (IoT) devices connected to battery management systems. This would allow the platform to automatically collect battery parameters such as voltage, temperature, and current without requiring manual input.

Future developments may also include the use of advanced deep learning models such as Long Short-Term Memory (LSTM) networks or other time-series prediction techniques. These models can analyze sequential battery data and improve the accuracy of battery degradation forecasting.

Additionally, the system could be extended with mobile applications and cloud-based monitoring platforms to provide remote access for EV operators and fleet managers. Integration with GPS data and usage patterns may further enhance prediction capabilities by considering real-world driving conditions.

Another possible enhancement is the implementation of intelligent recommendation systems that suggest battery replacement schedules or optimal swapping stations based on predicted battery health. These improvements would transform the platform into a more comprehensive and intelligent battery management solution for future electric vehicle infrastructure.

REFERENCES

[1] A. P. Renold and N. S. Kathayat, “Comprehensive Review of Machine Learning, Deep Learning, and Digital Twin Data-Driven Approaches in Battery Health Prediction of Electric Vehicles,” *IEEE Access*, vol. 12, pp. 43984–43999, 2024. Available: https://doi.org/10.1109/ACCESS.2024.Doi

[2] S. F. Chevtchenko *et al.*, “A Mapping Study of Machine Learning Methods for Remaining Useful Life Estimation of Lead-Acid Batteries,” *J. (Proceedings)*, 2023.

[3] N. Jiang *et al.*, “Driving Behavior-Guided Battery Health Monitoring for Electric Vehicles Using Machine Learning,” *arXiv*, Sep. 2023. Available: https://arxiv.org/abs/2309.14125

[4] G. V. Ukalkar *et al.*, “Enhancing Battery Health Prediction for Electric Vehicles through Machine Learning, Deep Learning, and Digital Twin Technology,” *IJIRSET*, Nov. 2024. Available: https://www.ijirset.com/upload/2024/november/36_Enhancing.pdf

[5] “Multi-modal framework for battery state of health evaluation using deep learning,” *Nature Communications*, 2025. Available: https://www.nature.com/articles/s41467-025-56485-7

[6] “Optimized XGBoost modeling for accurate battery capacity prediction,” *ScienceDirect*, 2024. Available: https://www.sciencedirect.com/science/article/pii/S2590123024010417

[7] “Advancing state of health estimation for electric vehicles,” *ScienceDirect*, 2024. Available: https://www.sciencedirect.com/science/article/pii/S266679242400026X

[8] “Random Forest-Based Machine Learning Model Design for SOH Assessment of 21,700 Cells,” *MDPI*, 2024. Available: https://www.mdpi.com/2673-7167/5/1/12

[9] “Evaluation of Advances in Battery Health Prediction for Electric Vehicles,” *MDPI*, 2024. Available: https://www.mdpi.com/2313-0105/10/10/356

[10] “Interpretable machine learning prediction for Li-ion battery’s state of health,” *ScienceDirect*, 2024. Available: https://www.sciencedirect.com/science/article/abs/pii/S001346862400690X

[11] C. Liu *et al.*, “Deep Learning for State of Health Estimation of Lithium-Ion Batteries in Electric Vehicles: A Systematic Review,” *Energies*, vol. 18, no. 6, Mar. 2025. Available: https://doi.org/10.3390/en18061463

[12] “Battery state of health estimation under fast charging via deep neural network,” *PMC*, 2025.

Available: https://www.ncbi.nlm.nih.gov/articles/PMC12033934/

[13] A. Thelen *et al.*, "Probabilistic Machine Learning for Battery Health Diagnostics and Prognostics—Review and Perspectives," *npj Materials Sustainability*, 2024. Available: https://www.nature.com/articles/s44296-024-00011-1

[14] "New energy vehicle battery state of charge prediction based on Random Forest & XGBoost," *Energy Informatics*, 2024. Available: https://energyinformatics.springeropen.com/articles/10.1186/s42162-024-00424-1

[15] "Electric Vehicles Charging Time Prediction Based on Multimodel," *ASME*, 2025. Available: https://asmedigitalcollection.asme.org/electrochemical/article/doi/10.1115/1.4068207

[16] "State of Health Estimation and Battery Management," *MDPI*, 2024. Available: https://www.mdpi.com/1996-1944/18/1/145

[17] "Lithium-Ion Battery Estimation in Online Framework Using Extreme Gradient Boosting," *ResearchGate*, 2025. Available: https://www.researchgate.net/publication/359171580

[18] S. A. Celtek *et al.*, "Machine Learning-Based Real-Time Remaining Useful Life Estimation and Fair Pricing Strategy for Electric Vehicle Battery Swapping Stations," *IEEE Access*, 2025. Available: https://www.researchgate.net/publication/390216241

[19] "State of Health Prediction in Electric Vehicle Batteries Using a Deep Learning Approach," *MDPI*, 2023. Available: https://www.mdpi.com/2032-6653/15/9/385

[20] X. Feng *et al.*, "Comprehensive Performance Comparison Among Different Types of Features in Data-Driven Battery State of Health Estimation," *arXiv*, Aug. 2023. Available: https://arxiv.org/abs/2308.13993

[21] A. Lanubile *et al.*, "Domain Knowledge-Guided Machine Learning Framework for State of Health Estimation in Lithium-Ion Batteries," *arXiv*, Sep. 2024. Available: https://arxiv.org/abs/2409.14575

[22] S. Navidi *et al.*, "Physics-Informed Machine Learning for Battery Degradation Diagnostics: A Comparison of State-of-the-Art Methods," *arXiv*, Apr. 2024. Available: https://arxiv.org/abs/2404.04429

[23] T. Bockrath *et al.*, "State of Health Estimation of Lithium-Ion Batteries with Temporal Convolutional Neural Network Using Partial Load Profiles," *Applied Energy*, 2023.

[24] J. Gao *et al.*, "State of Health Estimation of Lithium-Ion Batteries Based on Mixers-Bidirectional Temporal Convolutional Neural Network," *Journal of Energy Storage*, 2023.

[25] Z. Zhang *et al.*, "Voltage Relaxation-Based State-of-Health Estimation of Lithium-Ion Batteries Using CNN and Transfer Learning," *Journal of Energy Storage*, 2023.

[26] L. Chen and M. Huang, "Real-Time Electric Vehicle Battery SoH Prediction Using Ensemble Learning Methods," *IEEE Transactions on Vehicular Technology*, vol. 74, no. 3, pp. 2105–2116, Mar. 2025.

[27] R. Silva, J. Costa, and A. Fernandes, "Enhanced Battery Health Monitoring in EVs Based on Transfer Learning and Big Data Analytics," *Energy Informatics*, vol. 7, no. 2, p. 22, May 2024.

[28] K. Patel and S. Mehta, "Comparison of Tree-Based Algorithms for Predicting EV Battery Degradation," *International Journal of Electrochemical Science*, vol. 19, no. 11, pp. 3302–3318, Jun. 2024.

[29] F. González and P. Ruiz, "Explainable AI for Battery State of Health Estimation in Electric Vehicles," *Journal of Energy Storage*, vol. 55, pp. 104823–104834, Feb. 2025.

[30] H. Yang, Y. Li, and T. Zhang, "Real-World Driving Data Fusion for Accurate SoH Prediction Using XGBoost," *IEEE Access*, vol. 13, pp. 54210–54222, Apr. 2025.

[31] Q. Wu *et al.*, "LightGBM-Based State of Health Estimation for EV Batteries Under Variable Temperature Conditions," *Applied Sciences*, vol. 14, no. 4, pp. 2028–2040, Feb. 2024.

[32] M. Rossi and D. Bianchi, "Predictive Maintenance of EV Battery Packs via IoT and ML Integration," *Sensors*, vol. 24, no. 5, pp. 687–704, Mar. 2024.

[33] J. Singh and A. Kumar, "A Hybrid LSTM–XGBoost Model for Long-Term Degradation Forecasting of Lithium-Ion Batteries," *Energy AI*, vol. 11, p. 100230, Jul. 2025.

[34] S. Chen, L. Tang, and J. Zhao, "Battery Swapping Station Optimization Using Predictive Analytics and SoH Estimates," *International Journal of Vehicular Technology*, 2024.

[35] D. Kumar, P. Sharma, and R. Verma, "Development of a Mobile-First Dashboard for EV Battery Health Monitoring," *Journal of Mobile Multimedia*, vol. 20, no. 2, pp. 105–120, May 2023.