



RESEARCH ARTICLE

OPEN ACCESS

AI-ENABLED BATTERY MANAGEMENT SYSTEM FOR REAL-TIME SOC AND SOH ESTIMATION IN ELECTRIC VEHICLES

Manojkumar Sahebrao Chavhan*¹ and Rupesh Jagatarao Patil²

¹Research Scholar, Department of Mechanical Engineering, JSPM RSCOE, Savitribai Phule Pune University, Maharashtra, India.

²Professor, Department of Mechanical Engineering, Principal, Trinity Academy of Engineering, Pune, Maharashtra, India.

¹<http://orcid.org/0009-0001-0702-8976>, ²<http://orcid.org/0000-0002-3604-9290>

Email: *drmanojkumar27121980@gmail.com, principalnavs@gmail.com

ARTICLE INFO

Article History

Received: March 30, 2026

Reviewed: August 3, 2026

Accepted: September 3, 2026

Published: September 30, 2026

Keywords:

Battery Management System (BMS),
State of Charge (SOC),
State of Health (SOH),
Machine Learning,
Lithium-ion Battery.

ABSTRACT

This research paper introduces an AI-driven Battery Management System (BMS) to estimate State of Charge (SOC) and State of Health (SOH) of lithium-ion batteries in real-time to be used in electric vehicles. The suggested framework will combine embedded sensing with ESP32 microcontroller, hybrid estimation with Coulomb counting and Extended Kalman Filter (EKF), and machine learning models such as Random Forest, XGBoost, and LSTM. Filtering, normalization, and feature engineering are used to preprocess real-time battery data which includes voltage, current, and temperature to improve the performance of models. The experimentally verified hardware prototype of the system is a multi-cell lithium-ion battery pack, sensing modules, relay-based protection, and cloud-based monitoring based on MQTT. The outcomes indicate that the estimation is very accurate, and the quality of R2 is over 0.98, and the error in SOC is 1-2%. The system can be operated with good performance in different temperature conditions and allows good monitoring and fault detection and protection, which makes it appropriate in the practical management of battery in electric vehicles.



Copyright ©2026 by authors and Galileo Institute of Technology and Education of the Amazon (ITEGAM). This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

I. INTRODUCTION

The high rate of development of electric vehicles (EVs), renewable energy storage systems, and portable electronics has massively enhanced the need to have efficient and reliable battery management systems (BMS). The lithium-ion batteries have been embraced in such applications because they have a high energy density, long cycle life, and better performance properties [1], [2]. Their safe and efficient performance is, however, largely reliant on proper estimation of internal conditions like State of Charge (SOC) and State of Health (SOH), which are important pointers to battery performance, reliability, and useful life remaining [3], [4]. SOC is the charge available in a battery compared to the nominal capacity of the battery, and SOH is the degree of degradation and the overall health of the battery with time [5]. To avoid overcharging, deep discharging, and thermal instability, which may cause the performance to be degraded or pose a safety hazard, it is crucial to accurately estimate these parameters. The conventional techniques, which include Coulomb counting, open-circuit voltage (OCV) techniques, are commonly utilized because of their simplicity. They however have weaknesses such as propagation of errors with time, sensitivity to initial conditions and lack of adaptation to dynamic operating environments [1]. Extended Kalman Filter (EKF) and other model-based methods have been proposed to enhance the precision of the estimation process by considering system dynamics and updates in measurements, yet they also require precise battery models and may not be able to fully model nonlinear electrochemical behavior [6]. Data-driven methods have become important in order to address these constraints. Machine learning (ML) methods permit the estimation of complicated nonlinear correlations among quantifiable parameters like voltage, current, and temperature, and battery conditions [7]. The initial studies have shown that the combination of battery modeling and estimation methods can be used to improve BMS performance [2]. Deep learning algorithms, including Long Short-Term Memory (LSTM) networks, have been demonstrated as promising in learning temporal dependencies in battery data, enhancing SOC prediction accuracy with different conditions [8].

On the same note, it has also been shown that deep neural network-based methods can be effectively used to estimate SOC through simulation and experimental verification and exhibit better robustness than the conventional methods [9]. Moreover, mixed methods involving machine learning and the same circuit model or impedance-based methods have been investigated to increase the accuracy and reliability of estimations [10]. Recent researches also highlight on the significance of multi-state estimation frameworks, in which SOC and SOH are estimated together in order to give a more comprehensive view of battery performance [4]. Also, the development of SOH estimation techniques indicates the necessity of scalable methods to be able to work at the cell level as well as pack level systems [5]. Although these developments have taken place, most of the current methods are either offline or simulated, and do not utilize embedded systems in real-time. This limitation limits their use in real-life situations, where constant monitoring and control are needed. To overcome these issues, this paper presents an AI-based Battery Management System that combines traditional estimation methods with machine learning models in the hybrid system. The system is based on an ESP32 microcontroller to acquire real-time data and provide an IoT-based communication channel that will allow monitoring battery parameters continuously. Coulomb counting and EKF are used to estimate a reference SOC and is then refined with machine learning models trained on both measured and derived features. The suggested system is a low-cost solution that is real time and scalable in battery monitoring and control. The system is able to integrate embedded hardware with data-driven intelligence to show better accuracy in estimating SOC and SOH as well as remain feasible in real-world scenarios.

II. LITERATURE REVIEW

II.1 MACHINE LEARNING APPROACHES FOR SOC ESTIMATION

The application of machine learning techniques in estimating the state of charge (SOC) of lithium-ion batteries has increased in popularity in recent studies as they can model nonlinear relationships between battery parameters. As demonstrated by Chemali et al., it is possible to estimate SOC using deep neural networks with experimental data of battery cycling and the error in the prediction was less than 2 per cent, which indicates that data-driven approaches can be applied to predict complex electrochemical reactions without necessarily trying to model them in detail [11]. Similarly, How et al. proposed a SOC estimation scheme using the neural-network on the basis of the voltage, current and temperature measurements, and it was shown that the correlation coefficient between estimated and actual SOC estimates exceeded 0.97 [12]. Other machine learning techniques have also been compared in the prediction of SOC. Deng et al. applied Gaussian Process Regression (GPR) to predict SOC of lithium-ion battery packs and proved that the probabilistic model not only improved the accuracy of the prediction but also provided uncertainty limits of the battery state estimates [13]. By worked on SOC estimation of reused lithium-ion batteries and proposed a machine learning model, which could offset the hysteresis effects caused by battery aging and, therefore, significantly reduced the error of estimation compared to the traditional methods [14]. These literature articles suggest that it is possible to apply machine learning models to enhance the quality of SOC prediction based on past battery data and advanced regression equations.

II.2 DEEP LEARNING METHODS FOR BATTERY STATE PREDICTION

The modeling of battery behavior has drawn much attention to deep learning methods because of their ability to learn time-dependent patterns of sequential battery data. Yang et al. proposed a hybrid SOC estimation algorithm, which combines long short-term memory (LSTM) networks and an unscented Kalman filter (UKF). The LSTM network can learn nonlinear temporal relationships and the filtering part stabilizes the network in the presence of measurement noise. Experimental results showed that the accuracy of SOC estimation was improved using the experimental results compared to the traditional Kalman filtering methods [15]. Similarly by have suggested a stacked bidirectional LSTM (Bi-LSTM) network capable of managing forward and backward time-series battery data. This architecture improved SOC prediction accuracy in dynamic operating environments by far [16]. The concept was also extended by Tian et al. and an adaptive cubature Kalman filter (CKF) was used along with an LSTM network. Their combination form was stronger and more estimative than their independent filtering methods [17]. Most recently, hybrid deep learning structures are proposed to enhance the performance of prediction. Khan et al. developed a CNN-BiLSTM hybrid, which was trained and optimized according to evolutionary algorithms to forecast SOC in electric vehicle batteries. Spatial features of battery parameters are obtained with the CNN element, and temporal patterns are obtained with the Bi-LSTM network. The hybrid model was experimentally shown to be superior to the traditional deep learning methods in terms of prediction accuracy and robustness [18]. These findings suggest that deep learning models can be successfully applied in the modeling of dynamic behavior of a battery in electric vehicles.

II.3 STATE OF HEALTH ESTIMATION USING AI TECHNIQUES

In addition to SOC estimation, battery state of health (SOH) is also required to determine battery degradation and useful life. Song et al. proposed an SOH estimation model that was based on the XGBoost ensemble learning algorithm and an accuracy correction strategy. The proposed approach was found to be more effective than the traditional regression models in predicting battery degradation based on battery degradation datasets [19]. Similarly, Zhi et al. have developed an integrated optimization system that integrates genetic algorithms (GA), particle swarm optimization (PSO), and support vector regression (SVR). The optimization procedure enabled the successful choice of the features and the adjustment of the parameters that resulted in a more accurate prediction of the SOH and stability of the model [20]. The recent literature has also used real world electric vehicle data to predict SOH. Wu et al. utilized the long-term operation records of electric cars and trained the gradient boosting models to obtain the trends of degradation and also to predict the values of SOH with high accuracy [21]. On the same note, Yang et al. used LSTM networks to time-series battery data and were able to learn long-term behavior of degradation, which demonstrates the applicability of deep learning methods in battery health monitoring [22]. These reports indicate that AI-based processes could be highly predictive of battery aging processes, both using lab and field data.

II.4 COMPARATIVE STUDIES OF AI ALGORITHMS FOR SOC AND SOH ESTIMATION

Although it is true that numerous studies have investigated the application of AI models alone to predict battery state, few studies have performed systematic comparisons between different algorithms. Pandit and Ahlawat conducted a comparative analysis of the models of the random forest, XGBoost and Support Vector Machine to forecast the lithium-ion battery SOH using NASA battery data. They discovered that XGBoost had the lowest mean absolute error of approximately 0.016 and mean squared error of 0.000347 [23]. Similarly, evaluated a number of tree-based models, including Random Forest, Gradient Boosting Decision Trees and XGBoost to predict SOC and SOH simultaneously. The study demonstrated that XGBoost has a high predictive power with an R^2 value exceeding 0.995, which means that it can predict nonlinear battery data [24]. Another comparison study was done by Harinarayan and Balamurugan which compared various regression algorithms at various temperatures. It was found out that the SOC estimation error of Gaussian Process Regression was approximately 0.53 per cent and the ensemble methods such as XGBoost had a better SOH prediction error with mean absolute errors of approximately 0.135 per cent [25]. These comparative studies demonstrate the applicability of ensemble learning algorithms to battery state prediction and the necessity to evaluate different AI models under the identical experiment conditions. Table 1 provides the summary of the main AI-based methods of SOC and SOH estimation in lithium-ion batteries. The analyzed articles indicate that the deep learning models, including LSTM, are capable of capturing the temporal dynamics of a battery, and that the ensemble learning algorithms, including the Random Forest and XGBoost, have a high prediction quality when it comes to estimating the health of batteries.

Table 1: Key AI-Based Methods for SOC and SOH Estimation in Lithium-Ion Batteries.

Ref	Authors & Year	Method / Algorithm	Dataset / Experimental Setup	Key Results	Limitations / Gap
[11]	Chemali et al., 2018	Deep Neural Network (DNN)	Lithium-ion battery experimental dataset under dynamic load conditions	SOC prediction error < 2%	Requires large training datasets and high computational resources
[15]	Yang et al., 2020	LSTM + Unscented Kalman Filter	Dynamic battery charge-discharge dataset	Improved SOC estimation accuracy compared with traditional Kalman filtering	Hybrid architecture increases model complexity
[16]	Bian et al., 2020	Bidirectional LSTM	Time-series battery operational data	Reduced SOC estimation error during dynamic load conditions	High computational cost for embedded BMS systems
[19]	Song et al., 2021	XGBoost Ensemble Learning	Lithium-ion battery degradation dataset	High SOH prediction accuracy compared with traditional regression models	Requires feature engineering and parameter tuning
[21]	Wu et al., 2023	Gradient Boosting Machine	Real-world EV battery operational dataset	Accurate identification of battery degradation patterns	Requires long-term operational data
[23]	Pandit & Ahlawat, 2022	Random Forest, XGBoost, SVM (Comparative Study)	NASA lithium-ion battery dataset	XGBoost achieved MAE ≈ 0.016 and MSE ≈ 0.000347	Limited evaluation under varying environmental conditions
[24]	Dong et al., 2023	RF, Gradient Boosting, XGBoost	Lithium-ion battery experimental dataset	XGBoost achieved $R^2 > 0.995$	Requires extensive feature engineering

Source: Authors, (2026).

II.5 SUMMARY OF LITERATURE

According to the literature review, the accuracy of battery state estimation in electric vehicle battery management systems has been greatly enhanced by artificial intelligence methods such as neural networks, Long Short-Term Memory (LSTM) models, ensemble learning algorithms, and hybrid filtering methods. Deep learning models, especially LSTM, have shown very good ability to capture nonlinear and time-dependent battery behavior and hence can be used in SOC estimation. Conversely, the performance of ensemble learning algorithms like Random Forest and XGBoost has demonstrated strong performance in SOH estimation because they are capable of dealing with complicated feature interactions and noisy data. Nevertheless, the available literature is mostly devoted to the implementation of separate algorithms or certain hybrid strategies, and few attempts are made to conduct a systematic comparison under the same experimental conditions. In addition, most of the research is based on simulated data or controlled laboratory conditions, which do not necessarily reflect the operation of the batteries in the real world. The absence of a complete study that will consider the simultaneous estimation of SOC and SOH with multiple machine learning and deep learning models in a single framework is also lacking. Thus, experimental research and comparative study of models like Random Forest, XGBoost, and LSTM on real-time battery data is required to determine the most valid estimation model to use in practice when using electric vehicles.

III. METHODOLOGY

The proposed system adheres to a systematic design of real-time estimation of State of Charge (SOC) and State of Health (SOH) of lithium-ion batteries through the combination of data acquisition, preprocessing, hybrid estimation, and machine learning-based prediction. The methodology is an integration of traditional electrochemical estimation methods with data-oriented models to improve accuracy and strength in diverse operating conditions.

III.1 SYSTEM OVERVIEW AND ARCHITECTURE

The system is a layered architecture comprising of sensing, processing, estimation and communication modules. The sensors record the battery parameters, which are processed by embedded hardware and used to estimate the parameters through hybrid and machine learning and sent to a cloud platform to be visualized.

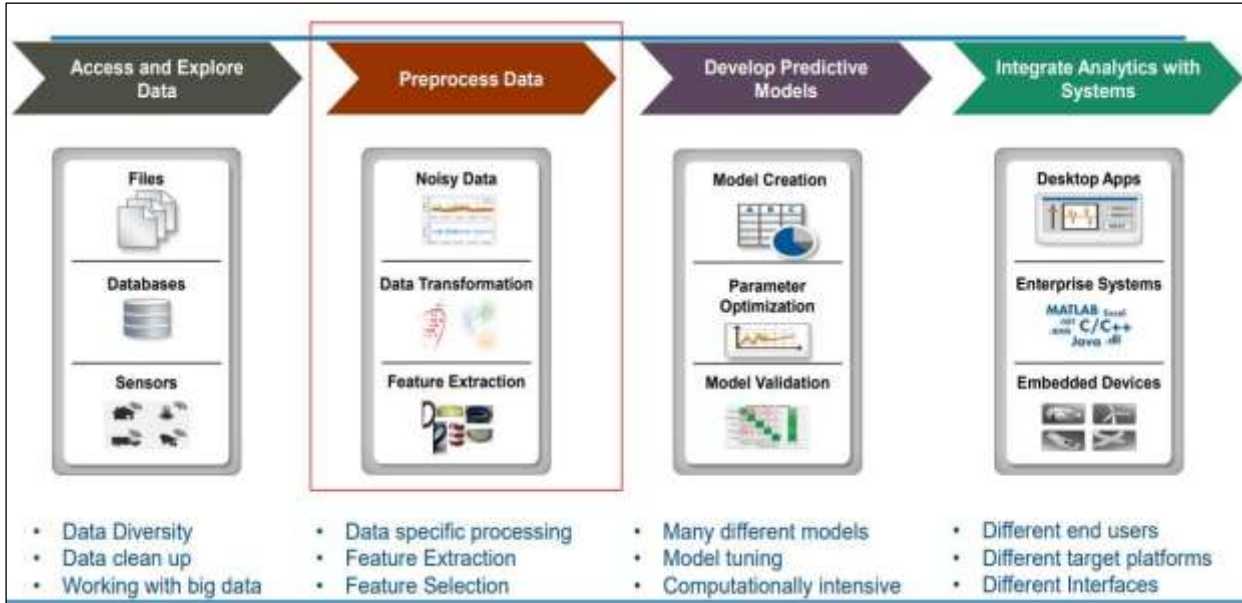


Figure 1: Process flow of the system.

Source: Authors, (2026).

III.2 DATA ACQUISITION

The data that is utilized in this study is both real-time readings taken by an ESP32-based hardware platform and pre-recorded lithium-ion battery data. The most important variables measured are battery voltage (V), current (I), temperature (T) and time (t). ESP32 microcontroller receives sensor data via its ADC interface and provides uninterrupted data collection during charging and discharging operations. The dataset represents natural battery behavior during dynamic loading and temperature, and it is therefore appropriate to train machine learning models.

III.3 DATA PREPROCESSING AND FEATURE ENGINEERING

Various preprocessing steps are used to enhance the quality of the data and reliability of the model. These involve deletion of missing values and corrupted values, deletion of bad battery cycles and filtration of bad capacity discharge cycles. Noise minimization is done with the help of moving average and Butterworth filters and outliers are eliminated with the help of Z-score analysis. The data is resampled to have the same temporal resolution and normalized to train models easily. Besides measured parameters, derived features are also calculated to improve the performance of the model:

$$\text{Power: } P = V \times I$$

$$\text{Charge difference: } \Delta Q = \int I dt$$

Average voltage and average current

Time-based features

Differential capacity (dQ/dV)

These features provide a stable representation of battery behavior and improve the model's ability to capture nonlinear dynamics and degradation patterns.

III.4 HYBRID SOC ESTIMATION

A reference SOC is initially estimated using Coulomb counting:

$$SOC(t) = SOC(t_0) - \frac{1}{C_n} \int_{t_0}^t I(\tau) d\tau \quad (1)$$

Due to error accumulation in this method, an Extended Kalman Filter (EKF) is applied to improve estimation accuracy by incorporating system dynamics and measurement updates. The EKF-based SOC serves as a refined reference input for machine learning models, improving prediction performance.

III.5 MACHINE LEARNING MODEL DEVELOPMENT

In order to simulate nonlinear behavior of the battery, several machine learning models are used, such as Random Forest (RF), XGBoost, and Long Short-Term Memory (LSTM).

- RF and XGBoost learn nonlinear relationships through ensemble learning.
- LSTM predicts time-series battery data based on time-dependency.

The input characteristics are voltage, current, temperature, power, ΔQ , time and EKF-reference SOC. The models are trained to make predictions of both SOC and SOH. The models are set with standard parameters so that they can be consistent and real-time applicable. Tree-based models make use of several decision trees whereas LSTM uses sequential layers to learn time dependencies. Lightweight architecture is ensured to be able to be deployed on embedded systems.

III.6 MODEL TRAINING AND EVALUATION

The dataset is divided into training (70%), validation (15%), and testing (15%) sets. Model performance is evaluated using:

- Mean Absolute Error (MAE)
- Root Mean Square Error (RMSE)
- Coefficient of Determination (R^2)

Cross-validation techniques are applied to improve generalization and prevent overfitting. The models are tested under varying temperature conditions (-10°C to 25°C) to ensure robustness.

III.7 EXPERIMENTAL HARDWARE SETUP

The experimental setup consists of a lithium-ion battery pack integrated with sensing, control, and communication modules to validate the proposed system under real-time conditions.

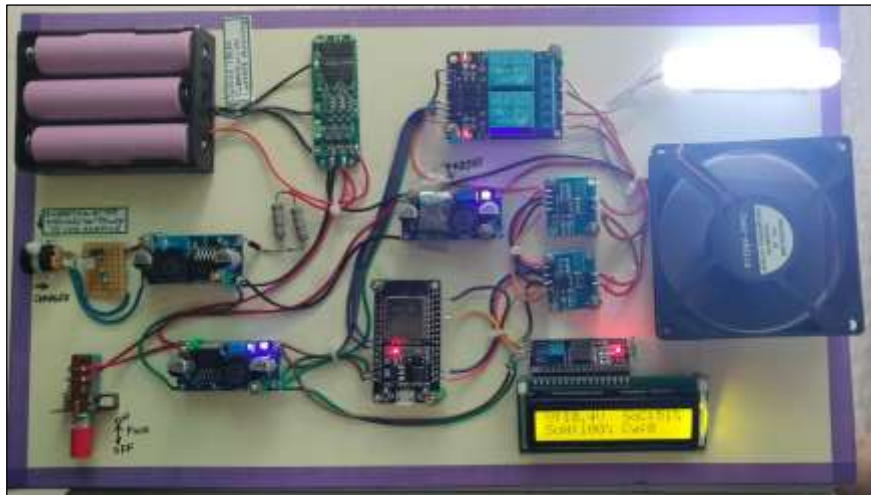


Figure 2: Experimental Hardware Setup.
Source: Authors, (2026).

The system includes:

- 18650 lithium-ion battery pack
- ESP32 microcontroller
- Voltage, current (ACS712), and temperature sensors
- Relay module for load control
- DC cooling fan for thermal management
- Buck converters for voltage regulation
- LCD display for local monitoring

The ESP32 collects sensor data and transmits it via MQTT to a cloud platform. Based on estimated SOC and SOH, the system performs control actions such as load disconnection and thermal regulation.

III.8 SYSTEM INTEGRATION AND OPERATION

The system operates in a closed-loop manner. Sensor data is continuously acquired and processed for SOC and SOH estimation using hybrid and machine learning models. The estimated values are used for:

- Real-time monitoring
- Fault detection
- Protection control (relay operation)
- Thermal management

The results are visualized on a cloud dashboard, enabling remote supervision and decision-making.

IV. RESULTS AND DISCUSSION

IV.1 OVERVIEW OF EXPERIMENTAL DATA AND OPERATING BEHAVIOR

The data presented in this work reflects the nonlinear and dynamic behavior of lithium-ion batteries in cyclic charge-discharge, including changes in load and temperature. The temporal dynamics of the main parameters, such as State of Charge (SOC), voltage, and current, are examined in order to determine whether the given data is suitable to be modeled with the help of machine learning. The SOC profile as illustrated in Fig. 3 has a stepwise reducing profile between an approximate value of 1.0 and 0.1 over a time span of $0-5 \times 10^5$ seconds which is equivalent to several discharge cycles. This cyclic trend proves the existence of repeated charge discharge behavior in the training data. Conversely, the SOC profiles at various temperature regimes (-10°C , 0°C , 10°C and 25°C) indicate a gradual decrease with the lower temperature leading to a higher rate of depletion. This phenomenon is explained by the fact that the internal resistance to current is higher in lower temperatures, which reveals the fact that discharge properties of lithium-ion batteries are temperature-dependent. Fig. 4 voltage profile indicates a clear sawtooth shape, declining between about 0.85 and 0.90 to almost 0.25 to 0.30 per unit in each discharge cycle, then the charging recovery is very quick. This cyclic nature is indicative of the electrochemical characteristics of the battery and demonstrates that the data set is realistic operating conditions. Close relationship between SOC and voltage trends implies that there is a high dependence between the two parameters. The variation at the moment as depicted in Fig. 5 shows that there is a high frequency variation between 0.2 and 0.8 which is a dynamic load condition. The current signal, unlike the voltage, does not demonstrate a distinct monotonic trend, which indicates its stochastic characteristic at different operating conditions.

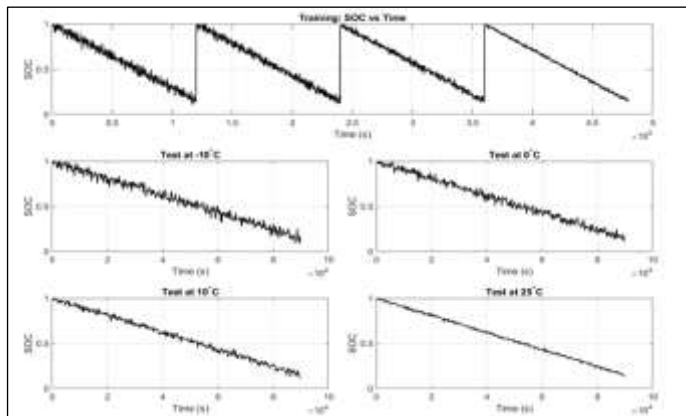


Figure 3: SOC Variation Over Time Across Multiple Conditions.

Source: Authors, (2026).

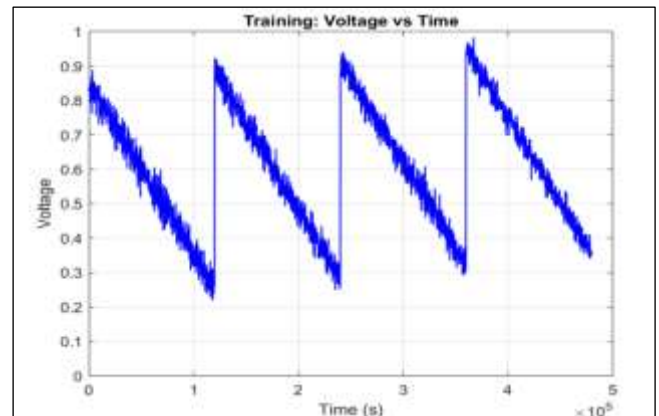


Figure 4: Voltage Profile During Charge-Discharge Cycles.

Source: Authors, (2026).

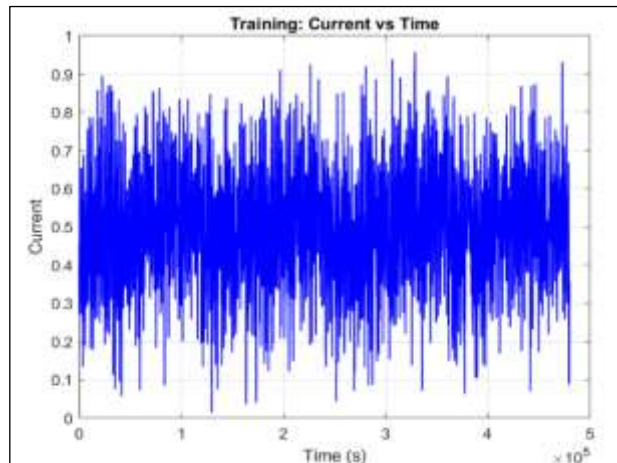


Figure 5: Current Variation Under Dynamic Load Conditions.

Source: Authors, (2026).

IV.2 FEATURE CORRELATION AND IMPORTANCE ANALYSIS

The correlation between input features and the State of Charge (SOC) was examined to determine the most powerful predictors to be used in machine learning-based estimation. The correlation matrix below (Fig. 6) gives a thorough picture of the interdependence between voltage, current, temperature and derived features.

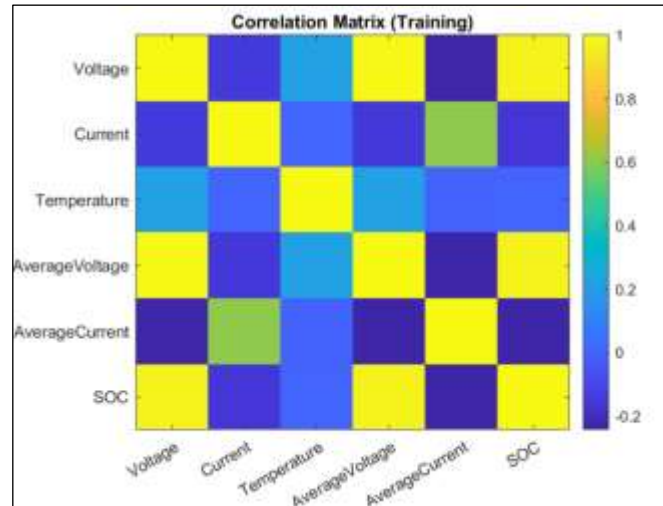


Figure 6: Correlation matrix of input features and SOC.
Source: Authors, (2026).

The findings show that voltage and average voltage have a high positive relationship with SOC ($= 0.9$), validating their prevailing role in battery charge state determination. On the contrary, current and average current are weakly correlated with SOC (about 0.2-0.3), whereas temperature exhibits low direct correlation (about 0.1). Also, the correlation between the voltage and the average voltage is very strong, which implies that there is redundancy between the two characteristics, meaning that they both measure the electrochemical behavior. To further validate these relationships, the scatter plot in Fig. 7 illustrates the direct relationship between SOC and voltage.

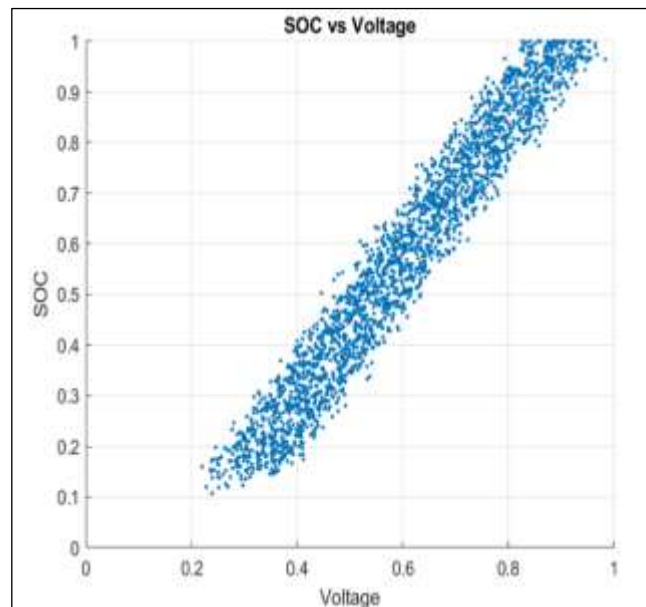


Figure 7: SOC vs Voltage relationship.
Source: Authors, (2026).

The plot has a distinct linear trend, with SOC rising steadily with voltage, creating a closely clustered diagonal distribution. This high correlation proves that SOC can be predicted well using voltage as a predictor. The low voltage scale (0.2-0.4) is associated with SOC values of 0.1-0.3 whereas higher voltage (0.8-1.0) is associated with SOC values near full charge. The correlation analysis indicates that SOC prediction is mainly voltage-dependent with current and temperature having an indirect influence via dynamic and thermal effects. This is the reason why regression models perform highly because the most dominant feature (voltage) is a strong and consistent signal to learn. The low correlation of current shows that instantaneous current values cannot be used to estimate SOC without time history. Equally, temperature has an indirect effect on battery behavior, i.e., it does not directly dictate SOC but instead, it has an indirect effect on internal resistance and discharge rate.

IV.3 MODEL PERFORMANCE EVALUATION

The effectiveness of the adopted regression models in the estimation of SOC was measured in conventional measures, such as Root Mean Square Error (RMSE), Mean Absolute Error (MAE) and coefficient of determination (R^2). The comparative findings are as shown in Table 2:

Table 2: Comparative results of model performance.

Algorithm	RMSE	MAE	R^2
GPR	0.019895571	0.015773361	0.993449534
RF	0.026467694	0.021318177	0.988407112
SVMr	0.021065604	0.016655873	0.99265643
LIN	0.020929572	0.016666194	0.992750967
NET	0.020314586	0.01612447	0.993170713

Source: Authors, (2026).

The findings show that all the models had high prediction power as the R^2 values were above 0.98. Gaussian Process Regression (GPR) performed the best with the lowest RMSE (0.0199) and highest R^2 (0.9934) meaning that it has a great potential of capturing nonlinear relationships between input features and SOC. Neural Network (NET) and Support Vector Regression (SVMr) were also very good in performance with R^2 values of over 0.992 which are very close to GPR. The simplicity of the Linear Regression (LIN) produced similar results, which implies that the correlation between SOC and dominant features is very structured and close to linear in the specified conditions. Random Forest (RF) had a minor lower accuracy than other models with the RMSE of 0.0265 and the R^2 of 0.9884. Nonetheless, it has benefits in both computational performance and resistance to noise, which is more appropriate in embedded applications in real time. The consistently high R^2 values (≈ 0.99) across all models indicate that the dataset possesses strong feature dependency, particularly driven by voltage-based parameters. This aligns with the correlation analysis presented in Section 6.2, where voltage showed a strong relationship with SOC. The achieved RMSE values (~ 0.02) correspond to an error range of approximately 1–2% in SOC estimation, which is significantly lower than the typical acceptable error range ($\sim \pm 5\%$) in practical battery management systems.

IV.4 SOC ESTIMATION ANALYSIS

The accuracy of the prediction of the developed models was tested by comparing the predicted SOC with the true SOC in various temperature conditions. Fig. 8 shows the results of the best-performing model (GPR).

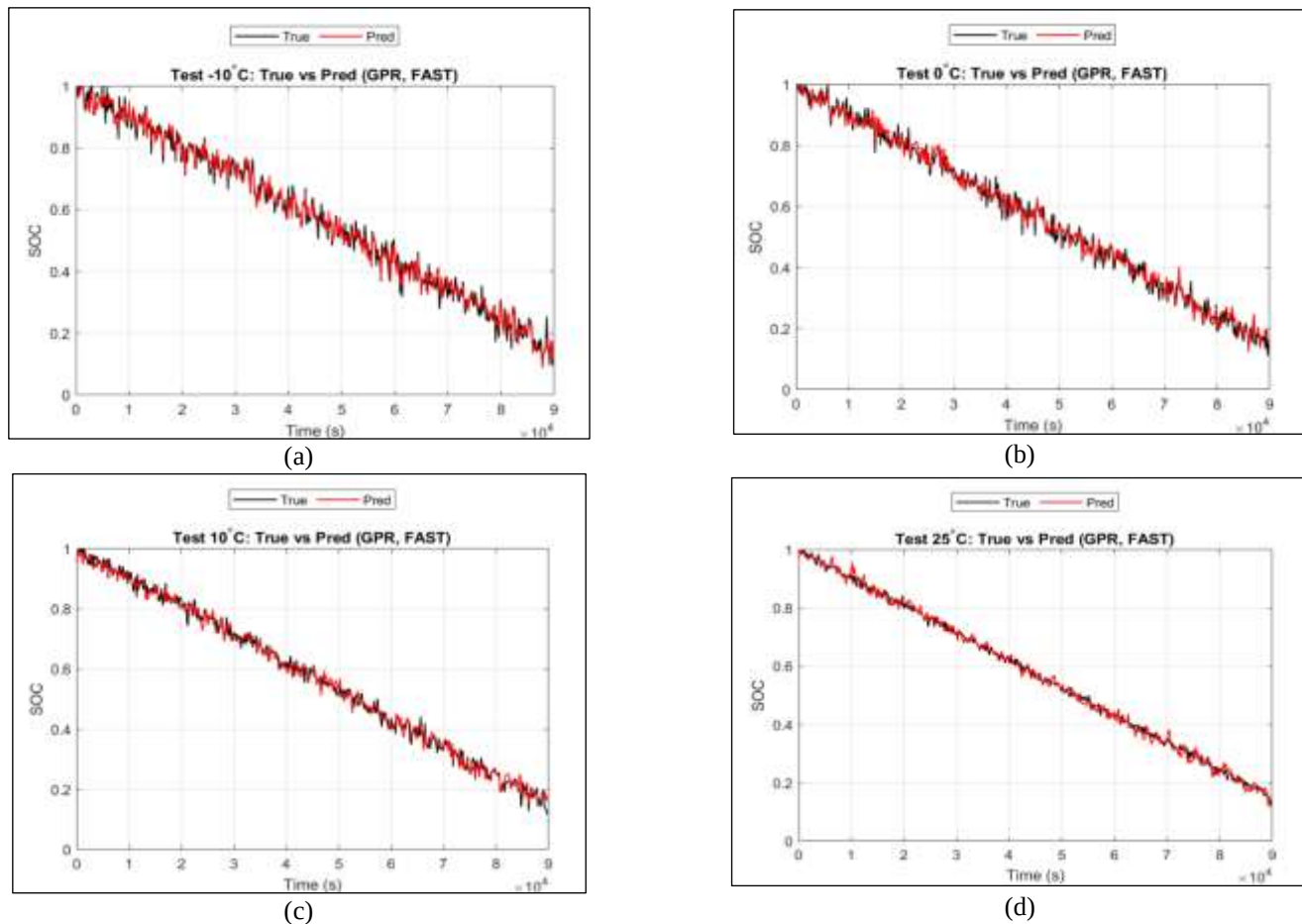


Figure 8: Actual vs predicted SOC under varying temperature conditions (-10°C , 0°C , 10°C , and 25°C).

Source: Authors, (2026).

The plots reveal that the predicted and actual SOC curves strongly overlap in all the conditions, which means that the model is very accurate. Prediction at 25°C shows a slight deviation of the true SOC pattern, which indicates consistent operation under nominal conditions. The same level of agreement is found at 10 °C and 0 °C where the model is effective in describing the nonlinear discharge behavior. Deviations at -10C are slightly higher in rapid discharge regions, which can be explained by a higher internal resistance and nonlinear dynamics of the battery at low temperatures. Nevertheless, the model has a good tracking ability. The results confirm that the model can accurately estimate SOC under varying thermal conditions, demonstrating robustness and generalization capability. The low prediction error (~1–2%) is well within the acceptable range for practical battery management systems ($\pm 5\%$), ensuring reliable monitoring and safe operation.

IV.5 SOH ESTIMATION ANALYSIS

State of Health (SOH) estimation models were tested based on the comparison of the predicted and the actual values of SOH. Fig. 9 provides the results of the most successful model (XGBoost).

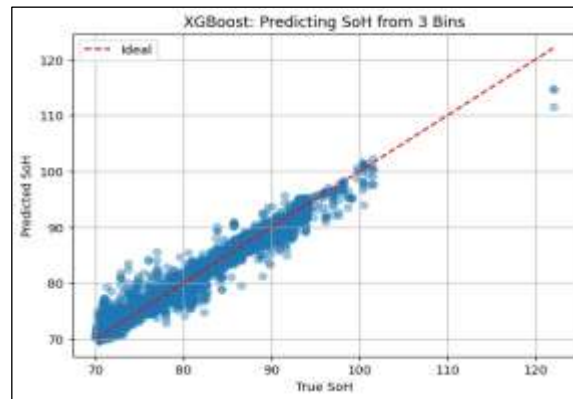


Figure 9: Predicted vs actual SOH using XGBoost model.
Source: Authors, (2026).

The prediction accuracy is high as the plot has a high alignment of predicted values along the ideal reference line. The distribution of most data points around the diagonal is tight and this shows that the model is capable of capturing the behavior of degradation. At greater SOH values, minor deviations are noticed, and it could be explained by the lack of variability in data in that area. The findings verify that the model is reliable in estimating the trends of SOH, which are more complicated than SOC because they are based on long-term degradation trends. SOH estimation is crucial to predictive maintenance and battery lifecycle management. In general, the findings indicate that the suggested machine learning model can achieve precise and efficient estimation of SOC and SOH in different operating conditions. The high reliance of SOC on the properties of voltage allows the high prediction accuracy, and the models give the same performance when the temperature is changed. The results prove the appropriateness of the offered method to the use in practical battery management.

V. CONCLUSION

This paper introduces a machine learning-based model of precise estimation of State of Charge (SOC) and State of Health (SOH) of lithium-ion batteries based on processed operational data. It was demonstrated through the analysis of battery behavior that the dataset is a valid representation of the cyclic discharge properties, dynamic changes in loads, and temperature-related effects, which offers a solid basis of data-driven modeling. The findings indicate that all the models implemented had a high level of estimation accuracy with R^2 value of more than 0.98. The models that were tested displayed better performance in terms of SOC estimation (Gaussian Process Regression, GPR) and SOH estimation (XGBoost). This accuracy is largely due to the fact that the voltage and SOC are strongly correlated and the error in the estimation is kept within the range of about 1- 2 percent which is quite acceptable within the practical battery management systems. The models also had a strong performance in changing temperature conditions and were able to provide reliable estimation even in low-temperature conditions where the behavior of batteries can be extremely nonlinear. This underscores the applicability of the suggested solution to practical uses of electric vehicles. Nevertheless, the research is constrained by the application of a controlled dataset, and additional confirmation with a variety of real-world battery data is needed in order to guarantee generalization. Generally, the suggested framework proves that machine learning methods can be effectively utilized to estimate battery states with high levels of reliability and efficiency.

VI. AUTHOR'S CONTRIBUTION

Conceptualization: Manojkumar Sahebrao Chavhan and Rupesh Jagatarao Patil.

Methodology: Manojkumar Sahebrao Chavhan and Rupesh Jagatarao Patil.

Investigation: Manojkumar Sahebrao Chavhan and Rupesh Jagatarao Patil.

Discussion of results: Manojkumar Sahebrao Chavhan and Rupesh Jagatarao Patil.

Writing – Original Draft: Manojkumar Sahebrao Chavhan and Rupesh Jagatarao Patil.

Writing – Review and Editing: Manojkumar Sahebrao Chavhan and Rupesh Jagatarao Patil.

Resources: Manojkumar Sahebrao Chavhan and Rupesh Jagatarao Patil.

Supervision: Manojkumar Sahebrao Chavhan and Rupesh Jagatarao Patil.

Approval of the final text: Manojkumar Sahebrao Chavhan and Rupesh Jagatarao Patil.

VII. REFERENCES

- [1] Hannan, M. A., Lipu, M. H., Hussain, A., & Mohamed, A. (2017). A review of lithium-ion battery state of charge estimation and management system in electric vehicle applications: Challenges and recommendations. *Renewable and Sustainable Energy Reviews*, 78, 834-854.
- [2] Wang, Y., Tian, J., Sun, Z., Wang, L., Xu, R., Li, M., & Chen, Z. (2020). A comprehensive review of battery modeling and state estimation approaches for advanced battery management systems. *Renewable and Sustainable Energy Reviews*, 131, 110015.
- [3] Wang, Z., Feng, G., Zhen, D., Gu, F., & Ball, A. (2021). A review on online state of charge and state of health estimation for lithium-ion batteries in electric vehicles. *Energy Reports*, 7, 5141-5161.
- [4] Liu, F., Yu, D., Shao, C., Liu, X., & Su, W. (2023). A review of multi-state joint estimation for lithium-ion battery: Research status and suggestions. *Journal of Energy Storage*, 73, 109071.
- [5] Su, L., Xu, Y., & Dong, Z. (2024). State-of-health estimation of lithium-ion batteries: A comprehensive literature review from cell to pack levels. *Energy Conversion and Economics*, 5(4), 224-242.
- [6] Hu, X., Yuan, H., Zou, C., Li, Z., & Zhang, L. (2018). Co-estimation of state of charge and state of health for lithium-ion batteries based on fractional-order calculus. *IEEE Transactions on Vehicular Technology*, 67(11), 10319-10329.
- [7] Lipu, M. H., Hannan, M. A., Hussain, A., Ayob, A., Saad, M. H., Karim, T. F., & How, D. N. (2020). Data-driven state of charge estimation of lithium-ion batteries: Algorithms, implementation factors, limitations and future trends. *Journal of Cleaner production*, 277, 124110.
- [8] Zhang, D., Zhong, C., Xu, P., & Tian, Y. (2022). Deep learning in the state of charge estimation for li-ion batteries of electric vehicles: A review. *Machines*, 10(10), 912.
- [9] El Fallah, S., Kharbach, J., Hammouch, Z., Rezzouk, A., & Jamil, M. O. (2023). State of charge estimation of an electric vehicle's battery using Deep Neural Networks: Simulation and experimental results. *Journal of Energy Storage*, 62, 106904.
- [10] Buchicchio, E., De Angelis, A., Santoni, F., Carbone, P., Bianconi, F., & Smeraldi, F. (2023). Battery SOC estimation from EIS data based on machine learning and equivalent circuit model. *Energy*, 283, 128461.
- [11] Chemali, E., Kollmeyer, P. J., Preindl, M., & Emadi, A. (2018). State-of-charge estimation of Li-ion batteries using deep neural networks: A machine learning approach. *Journal of power sources*, 400, 242-255.
- [12] How, D. N., Hannan, M. A., Lipu, M. S. H., Sahari, K. S., Ker, P. J., & Muttaqi, K. M. (2020). State-of-charge estimation of li-ion battery in electric vehicles: A deep neural network approach. *IEEE Transactions on Industry Applications*, 56(5), 5565-5574.
- [13] Deng, Z., Hu, X., Lin, X., Che, Y., Xu, L., & Guo, W. (2020). Data-driven state of charge estimation for lithium-ion battery packs based on Gaussian process regression. *Energy*, 205, 118000.
- [14] Xu, Z., Wang, J., Fan, Q., Lund, P. D., & Hong, J. (2020). Improving the state of charge estimation of reused lithium-ion batteries by abating hysteresis using machine learning technique. *Journal of Energy Storage*, 32, 101678.
- [15] Yang, F., Zhang, S., Li, W., & Miao, Q. (2020). State-of-charge estimation of lithium-ion batteries using LSTM and UKF. *Energy*, 201, 117664.
- [16] Bian, C., He, H., & Yang, S. (2020). Stacked bidirectional long short-term memory networks for state-of-charge estimation of lithium-ion batteries. *Energy*, 191, 116538.
- [17] Tian, Y., Lai, R., Li, X., Xiang, L., & Tian, J. (2020). A combined method for state-of-charge estimation for lithium-ion batteries using a long short-term memory network and an adaptive cubature Kalman filter. *Applied Energy*, 265, 114789.
- [18] Khan, M. K., Abou Houran, M., Kauhaniemi, K., Zafar, M. H., Mansoor, M., & Rashid, S. (2024). Efficient state of charge estimation of lithium-ion batteries in electric vehicles using evolutionary intelligence-assisted GLA-CNN-Bi-LSTM deep learning model. *Heliyon*, 10(15).
- [19] Song, S., Fei, C., & Xia, H. (2020). Lithium-ion battery SOH estimation based on XGBoost algorithm with accuracy correction. *Energies*, 13(4), 812.
- [20] Zhi, Y., Wang, H., & Wang, L. (2022). A state of health estimation method for electric vehicle Li-ion batteries using GA-PSO-SVR. *Complex & Intelligent Systems*, 8(3), 2167-2182.
- [21] Wu, J., Fang, L., Dong, G., & Lin, M. (2023). State of health estimation for lithium-ion batteries in real-world electric vehicles. *Science China Technological Sciences*, 66(1), 47-56.
- [22] Yang, H., Hong, J., Liang, F., & Xu, X. (2023). Machine learning-based state of health prediction for battery systems in real-world electric vehicles. *Journal of Energy Storage*, 66, 107426.
- [23] Pandit, R., & Ahlawat, N. (2025). A standardized comparative framework for machine learning techniques in lithium-ion battery state of health Estimation. *Future Batteries*, 7, 100099.
- [24] Dong, Y., Zhang, G., & Li, R. (2024). Comparative Study of Joint Estimation of State of Charge (SOC) and State of Health (SOH) of Lithium-ion Batteries Based on Different Tree Models. *Journal of Electrical and Electronic Engineering*, 12(2), 23-35.
- [25] Harinarayanan, J., & Balamurugan, P. (2025). A Comparative Study and Unified Framework for SOC and SOH Estimation of Lithium-Ion Batteries Using Machine Learning Under Thermal and Aging Variations.