EES Batteries

Accepted Manuscript

This article can be cited before page numbers have been issued, to do this please use: M. Gong, Y. Choi, H. M. Yang, S. B. Ma and D. Seo, *EES Batteries*, 2026, DOI: 10.1039/D6EB00066E.

View Article Online
DOI: 10.1039/D6EB00066E

Broader context for

Simultaneous Prediction of Three Key Li-Ion Battery Indicators through Multi-Task Learning

Accurate battery diagnosis is important for the global transition to renewable energy and the rapid growth of a sustainable second-life market. As the world shifts toward electric mobility and energy storage systems, ensuring the reliability and safety of battery infrastructure has become a critical socio-economic priority. However, assessing vital indicators such as the battery's remaining lifespan, internal resistance, and thermal safety behavior remains a major practical bottleneck. Traditional diagnostic methods require hours of testing. In addition, battery aging follows complex and non-linear degradation pathways, making accurate prediction more challenging. We overcome these limitations by introducing a multi-task learning framework that enables the rapid, simultaneous prediction of all three indicators through a brief diagnostic test. By capturing the correlations between different degradation indicators, our model significantly outperforms conventional single-task approaches, reducing prediction errors by up to 29.4%. This method leverages positive transfer and inductive bias, enabling high accuracy even in data-scarce environments where traditional models typically underperform. Ultimately, our research provides a scalable tool for the industry, accelerating safety assessments and enhancing recycling efficiency. By understanding these coupled aging patterns, we offer a pathway to optimize battery lifecycles, reduce industrial costs, and support more sustainable battery use and second-life strategies.

View Article Online
DOI: 10.1039/D6EB00066E

Simultaneous Prediction of Three Key Li-Ion Battery Indicators through Multi-Task Learning

Minjeong Gong[a], Yoonjung Choi[a], Han Mo Yang[b], Sang Bok Ma[b], Dong-Hwa Seo[a]*

Affiliation

[a] *Department of Materials Science and Engineering, Korea Advanced Institute of Science and Technology (KAIST), 291 Daehak-ro, Yuseong-gu, Daejeon 34141, Republic of Korea*

[b] *Battery Tech Department, SK ecoplant, 19 Yulgok-ro 2-gil, Jongno-gu, Seoul 03143, Republic of Korea*

* Corresponding author: dseo@kaist.ac.kr (Prof. D.-H. Seo)

Abstract

Accurate diagnosis of battery states is essential for ensuring operational safety, evaluating current performance, and assessing suitability for second-life applications. However, characterizing key indicators such as State of Health (SoH), Direct Current Internal Resistance (DCIR), and maximum cell temperature during operation (Max. temp.) remains highly time-intensive because it requires low-current measurements and separate testing procedures. The challenge is further compounded by the weak linearity and large variance in the relationships among these indicators. To address these challenges, we developed a Multi-Task Learning (MTL) framework that combined a rapid diagnostic protocol with simultaneous prediction of the three key battery indicators. Using features extracted from the rapid diagnostic protocol, a neural network model was trained to predict multiple targets concurrently. The proposed MTL framework outperformed single-task baselines, reducing the root mean squared error by 25.9% for SoH, 25.0% for the DCIR increase rate, and 29.4% for Max. temp. These performance gains arise from the joint learning of highly correlated targets, which enables the model to capture coupled degradation behavior and promotes positive transfer and inductive bias. By integrating a rapid diagnostic protocol with an MTL, this study provides an accurate and data-efficient framework for simultaneous prediction of multiple degradation-relevant battery indicators.

View Article Online
DOI: 10.1039/D6EB00066E

Introduction

Owing to their high energy density and long cyclability, Lithium-Ion Batteries (LIBs) are widely utilized in applications ranging from portable electronic devices to Electric Vehicles (EVs) and Energy Storage Systems (ESSs). However, battery degradation and lifespan are critically influenced by usage patterns, such as charging time or operating State of Charge (SoC) ranges, which induce distinct degradation mechanisms. Therefore, accurate estimation of the cell's current state is essential for ensuring operational safety and assessing the viability of spent batteries for second-life applications.

To quantify the degree of battery degradation, key indicators such as State of Health (SoH) and Direct Current Internal Resistance (DCIR) are widely used. SoH is defined as the ratio of the current maximum capacity to the capacity at the Beginning of Life (BoL), representing the remaining usable energy. DCIR reflects the internal resistance of the cell, which typically increases with cell degradation. However, direct measurement of these indicators is a time-consuming process, requiring a minimum total time of 7 hours per cell: approximately 4 hours for SoH measurement via a 0.5 C-rate (I_1) charging and discharging, and an additional 3 hours for DCIR measurement involving specific pulse tests such as $5\times I_1$ for 10 seconds at a Depth of Discharge (DoD) 50%.

In addition, the operating cell temperature is a critical factor determining battery safety.[1] As degradation progresses, the thickening of the Solid-Electrolyte Interphase (SEI) increases internal resistance, which in turn intensifies heat generation.[2] Crucially, if the cell temperature exceeds 70 ℃, the SEI layer begins to decompose, potentially triggering thermal runaway. Therefore, monitoring of the operating cell temperature is essential to prevent both accelerated degradation and catastrophic safety failures.

Since these indicators typically cannot be measured simultaneously, characterizing a large number of LIBs is a time-intensive process. To mitigate this challenge, researchers have proposed battery indicator prediction methods through Machine Learning (ML) models, with most studies focusing on single-target prediction. For instance, SoH prediction has been extensively studied using voltage profiles from galvanostatic measurements, such as partial Incremental Capacity (IC) curves or current pulse data at a specific State of Charge (SoC).[3-9] In contrast, DCIR prediction remains underexplored due to a lack of available data and the need for additional specific measurement protocols. While resistances derived from Electrochemical Impedance Spectroscopy (EIS) have been predicted using variables such as

View Article Online
DOI: 10.1039/D6EB00066E

voltage and temperature, EIS is fundamentally based on small-signal perturbations.[10] Therefore, it does not accurately reflect the internal resistance under high-current conditions typical of practical operation.[11]

While single-target models remain the prevalent approach across various fields, including battery research, domains analyzing complex and interrelated data increasingly leverage Multi-Task Learning (MTL) to uncover hidden correlations and optimize interdependent tasks simultaneously.[12-15] In an MTL framework, when multiple targets stem from common underlying causes—such as shared signatures—the hidden layers learn to extract these fundamental factors more robustly. By forcing the network parameters to optimize across all tasks simultaneously rather than overfitting to a single objective, this architecture inherently provides a strong regularization effect. Furthermore, when the gradients from different targets align during training, the optimization of one task intrinsically facilitates the others. This phenomenon, known as positive transfer, effectively amplifies the learning signals supplied to the shared representations from a given dataset. Consequently, especially in data-scarce environments, MTL consistently achieves higher predictive accuracy than Single-Task Learning (STL) by taking full advantage of the synergistic relationships among the targets.

Within the battery research field, although research has largely focused on single-target prediction, some studies have begun exploring the potential of MTL, particularly for predicting metrics such as Remaining Useful Life (RUL) and power. For instance, a notable study demonstrated that jointly forecasting capacity fade and power degradation can significantly enhance prediction accuracy by leveraging the physical correlation between these indicators.[16] Similarly, recent research has focused on the joint assessment of SoH and RUL by utilizing optimized health indicators extracted from multi-source data to capture complex degradation patterns.[17]

However, these approaches are inherently constrained by their data requirements. While some models depend on extensive historical cycling data to track long-term trends, others rely on features extracted from full charge-discharge cycles. In both cases, the necessity for prolonged data collection or exhaustive cycling data renders such methods impractical for real-time, on-board applications where rapid diagnosis is essential. Furthermore, the prevailing focus has remained on empirical performance gains rather than a detailed analysis of the underlying drivers or shared features that facilitate co-learning. Despite achieving higher accuracy, these models remain somewhat opaque, as the specific input attributes prioritized

View Article Online
DOI: 10.1039/D6EB00066E

across different tasks have yet to be fully elucidated. Consequently, the internal synergy between targets remains a 'black box', leaving it unclear whether the model is capturing genuine physical correlations.

In battery systems, SoH degradation, increased DCIR, and elevated operating temperatures are not independent phenomena. Rather, they are driven concurrently by common electrochemical mechanisms, such as impedance growth, increased polarization, and accelerated heat generation. By employing an MTL framework, the shared hidden layers extract a unified latent representation that intrinsically captures these coupled degradation mechanisms. This joint optimization offers a significant advantage over STL in achieving superior generalization. Consequently, MTL can deliver highly accurate and robust predictions, making it effective for practical applications where training data is scarce.

Herein, we propose a simultaneous prediction model for three key battery indicators utilizing an MTL framework, eliminating the need for pre-settings and additional measurements. The dataset was constructed using 18650-sized cylindrical LIBs ($LiNi_xCo_yAl_zO_2$/graphite) degraded under various C-rates. The battery indicator prediction process consists of two sequential parts: diagnostic protocol classification and battery indicator regression. Building upon our previous work, which validated the classification performance using a 10-second discharge voltage at a 1.0 C-rate, we utilized the voltage data corresponding to the identified diagnostic protocols as input.[18]

The model simultaneously predicts SoH, DCIR, and the maximum temperature during 0.5 C-rate charging and discharging (Max. temp.). By employing an MTL based on Neural Network (NN), the model effectively captures the correlations among these indicators, demonstrating superior performance compared to STL baselines. Furthermore, analyses using gradient cosine similarity and Integrated Gradients (IG) confirmed synergistic interactions between targets during training. Finally, a t-distributed Stochastic Neighbor Embedding (t-SNE) visualization of high-impact features revealed that a concise subset of only seven features is sufficient to represent the current degradation state.

Results

Dataset construction and interrelation of battery indicators

The overall framework for battery indicator predictions is illustrated in **Fig. 1**. This process builds upon our previous work, which demonstrated superior performance in SoH

View Article Online
DOI: 10.1039/D6EB00066E

prediction.[18] It consists of two sequential steps: diagnostic protocol classification and battery indicator regression. In this study, we employ an MTL model to simultaneously predict three key targets: SoH, DCIR, and Max. temp. The process begins by acquiring the Open Circuit Voltage (OCV) and a 1.0 C-rate discharge voltage for 10 seconds from a battery in an unknown state. These voltages are used to identify the appropriate diagnostic protocol. Specifically, the protocols are categorized into three types based on SoC and are designed to prevent over-discharge while ensuring that the battery does not reach its upper or lower voltage limits. Subsequently, the specific voltage data corresponding to the classified diagnostic protocol serve as inputs for the MTL model. In our previous work, we confirmed a diagnostic protocol classification accuracy of 0.9985 by utilizing key features such as OCV, skewness, and resistance extracted from 10-second 1.0 C-rate discharge profiles.[18] Based on these verified results, the primary focus of this work is shifted toward the simultaneous prediction of these battery indicators.

Data were collected using a total of 44 18650-sized cylindrical LIBs (Samsung SDI, $LiNi_xCo_yAl_zO_2$/Graphite, 2.85 Ah), as illustrated in **Fig. 2a**. The cells were degraded under various C-rates (charge: 0.5, 1.0, and 2.0 C-rate/discharge: 0.5, 1.0, 2.0, and 4.0 C-rate). Following the degradation cycle, a reset cycle with a 0.5 C-rate charge and discharge was performed to mitigate any polarization effects accumulated during the degradation cycle. Subsequently, three key battery indicators were measured. First, the discharge capacity at a 0.5 C-rate (I_1) was obtained for the SoH calculation. Second, the DCIR was measured via a 10-second discharge pulse at a 2.5 C-rate ($I_2 = 5 \times I_1$) at a 50% DoD. Notably, the relative DCIR increase rate—normalized by the BoL value—was utilized as the target variable to account for form-factor-dependent initial variations. Finally, the maximum temperature (Max. temp.) was obtained during the SoH measurement (0.5 C-rate charge and discharge) as shown in **Fig. 2b**. Given that the Max. temp. was measured under low-current conditions, the potential for rapid temperature increase is substantial, strongly indicating that thermal issues will be markedly exacerbated under high C-rate operations (**Fig. S1**). The statistical distributions of these three collected indicators are presented in **Fig. 2c-e**. Subsequently, the diagnostic protocol was conducted at 5% SoC intervals from 100% to 0%. This resulted in a total dataset of 16,391 samples, encompassing diverse degradation and SoC levels.

These indicators—SoH, DCIR increase rate, and Max. temp.—are critical metrics indicative of cell degradation.[19, 20] The interrelationships among these indicators were

View Article Online
DOI: 10.1039/D6EB00066E

confirmed (**Fig. 2f-h**). Specifically, SoH demonstrated a strong correlation with the DCIR increase rate, accompanied by a widening variance as degradation progressed. Conversely, while the Max. temp. was correlated with both SoH and DCIR increase rate, it exhibited weaker linear relationships and significantly larger variances. Consequently, to achieve a holistic diagnosis of the battery state, it is essential to evaluate all three indicators rather than relying on a single indicator.

Performance comparison of MTL and STL models

The MTL model was architected with five shared hidden layers, each containing 256 nodes, followed by three task-specific layers for each indicator. During the architecture optimization process, 15% of the training data was strictly allocated as a validation set to monitor performance, thereby preventing overfitting. To enable learning of diverse degradation mechanisms, data points from the same cell at various stages were treated as independent samples and randomly distributed across the training and test sets. Crucially, the test set remained entirely isolated from every phase of the optimization process, including the structural tuning performed via the validation set. Consequently, the final results reported herein reflect the model's rigorous generalization performance on novel, unseen samples, ensuring the reliability of the predictive metrics.

By leveraging features extracted from the diagnostic protocol, the MTL model simultaneously predicted the three battery indicators (**Fig. 3a-c**). The model demonstrated superior performance on the training and test dataset, achieving R^2 scores of 0.991 and 0.988 for SoH, 0.997 and 0.997 for DCIR increase rate, 0.921 and 0.883 for Max. temp., respectively. Furthermore, by training the model on a comprehensive dataset covering the full SoC range, we ensured that the framework can accurately predict the three indicators using diagnostic protocol data regardless of the cell's current SoC.

For comparative analysis, three STL models were trained for each individual target battery indicator, as shown in **Fig. 3d-f**. To establish a rigorous baseline, each STL model was individually optimized using an AutoML framework that evaluated approximately 20 different architectures. Within this framework, a k-fold cross-validation strategy ($k = 7$) was implemented strictly on the training subset to robustly evaluate and select the optimal algorithm for each task. This resulted in the selection of RandomForest regressor for SoH and DCIR increase rate, and ExtraTrees regressor for Max. temp. Furthermore, feature selection was applied to these STL baselines by removing features with low feature

View Article Online
DOI: 10.1039/D6EB00066E

importance to ensure the highest possible predictive performance for each specific task. Consistent with our MTL approach, these optimized models were evaluated solely on the independent test set. These baseline models yielded training and test R^2 scores of 0.998 and 0.984 for SoH, 0.997 and 0.994 for DCIR increase rate, and 0.811 and 0.770 for Max. temp., respectively. Consequently, the proposed MTL approach proved superior to the STL baselines. The improvement was particularly significant in terms of RMSE reduction, recording reductions of 25.9% for SoH, 25.0% for the DCIR increase rate, and 29.4% for Max. temp.

Mechanistic understanding of MTL performance improvement

The factors contributing to the improved performance of the MTL framework were further investigated. As shown in **Fig. 4a**, the losses, defined as the mean squared error, exhibited a continuous decline during training. Notably, the loss for Max. temp. converged at a slower rate compared to the other indicators. This indicates that Max. temp. prediction was inherently more difficult than that of the other battery indicators.

To elucidate the primary factors contributing to the performance enhancement, a gradient cosine similarity analysis was conducted (**Fig. 4b**). Cosine similarity quantifies the directional alignment between two vectors, serving as a metric widely applied in diverse fields, including biomedical sciences.[21, 22] In the context of this study, we computed the cosine similarities between the gradient vectors derived from the loss functions of the different targets. This approach enables analysis of how the shared layer parameters are updated during training, specifically by revealing whether the gradients from different tasks are aligned or conflicting.

The analysis revealed that the mean gradient cosine similarities of Max. temp. with respect to SoH and DCIR increase rate were approximately 0.175 and 0.272, respectively. These positive values indicate that the weight updates occurred in an aligned direction, thereby facilitating positive transfer across tasks, where the learning process of one task benefits from that of the others. In contrast, the cosine similarity between SoH and DCIR increase rate remained negative, with an average value of -0.239, and gradually decreased during training, indicating partial gradient misalignment.

Despite this explicit gradient conflict, the loss functions for both SoH and DCIR increase rate converged simultaneously, and the MTL model consistently outperformed STL NN models with identical backbone architectures (**Fig. S2**). This observation suggests

View Article Online
DOI: 10.1039/D6EB00066E

that shared representations in the MTL framework may help capture common degradation-related features while mitigating task-specific overfitting. Rather than indicating complete gradient agreement, the coexistence of partial alignment and misalignment appears to contribute to improved generalization performance, potentially acting as an implicit regularization effect within the shared layers.[23]

Building on this observation, the slightly lower training error of STL likely reflects task-specific fitting rather than better generalization. In contrast, MTL constrains the model through joint optimization of correlated targets, promoting shared and more generalizable representations. The coexistence of gradient alignment and conflict, therefore, appears to act as an implicit regularizer, leading to improved test performance despite slightly higher training error.

Finally, IG analysis was employed to quantify the contribution of individual features to the target indicators. Mathematically, the IG value is derived by accumulating the gradients along a linear interpolation path from a baseline to the input and multiplying this integral by the difference between the input and the baseline.[24-26] Analogous to SHapley Additive exPlanations (SHAP), a positive IG value indicates that a feature contributes to an increase in the target prediction, whereas a negative value suggests a suppressive effect. However, IG was selected over SHAP for this study due to its superior computational efficiency and its direct utilization of model gradients, which aligns well with the differentiable nature of neural networks.

The top four features with the highest mean absolute IG values for each battery indicator are shown in **Fig. 4c-e**. As highlighted in the colored boxes, three of the top four features consistently appeared across all target indicators. This significant overlap suggests that the MTL model leveraged the shared features to facilitate cross-task information sharing, thereby confirming the occurrence of positive transfer.

Among the various features, 2.0C_dch_P_skew emerged as a critical feature for all indicators. Physically, the voltage profile tends to exhibit a steeper slope as the cell degrades, leading to a corresponding increase in the value of 2.0C_dch_P_skew (**Fig. S3a**).[27] This degradation characteristic was accurately reflected in the IG values for both DCIR increase rate and Max. temp. predictions. In contrast, the contribution of 2.0C_dch_P_skew to SoH prediction exhibited a positive correlation with the feature value itself. Notably, 2.0C_dch_P_skew was identified as the most impactful feature for Max. temp. prediction.

View Article Online
DOI: 10.1039/D6EB00066E

Since elevated cell temperatures enhance electrochemical kinetics, they can induce a temporary increase in discharge capacity (**Fig. S4**). Consequently, 2.0C_dch_P_skew, serving as a proxy for this thermal effect, accounts for the phenomenon where a temperature rise contributes to an apparent SoH increase.[28, 29] This intricate relationship demonstrates the MTL model's capability to capture complex and simultaneous interactions within the battery system.

For SoH prediction, 2.0C_P_hyst emerged as the most influential feature. As the cell degrades, the increase in polarization effects leads to a pronounced difference between the end voltages of the charge and discharge pulses.[30] Consequently, the IG value exhibited an inverse relationship with 2.0C_P_hyst, increasing as the feature value decreased. Similarly, the voltage relaxation during the rest time following a current pulse became more pronounced compared to the BoL, a phenomenon also derived from increased polarization. This trend resulted in an increase in 1.65C_dch_R_kurt, which is the kurtosis of the voltage distribution during the rest time, as degradation progressed (**Fig. S3b**).

Regarding the DCIR increase rate prediction, the resistance-related features—specifically 2.0C_ch_P_$R_{0.1\text{-}5s}$, 1.65C_dch_R_$R_{0.1\text{-}5s}$, and 2.0C_ch_R_$R_{0.1\text{-}5s}$—demonstrated the highest impact. For these features, a direct correlation was observed, where an increase in their values corresponded to higher IG values, reflecting the physical rise in internal resistance.

Additionally, SHAP analysis was performed using 1,000 randomly selected samples as the background distribution to identify general importance trends while minimizing computational overhead (**Fig. S5**). The top-ranking features showed slight discrepancies between SHAP and IG, which is attributed to their different calculation mechanisms: SHAP approximates feature contributions based on game-theoretic expectations against a baseline, whereas IG integrates the gradients along a path from a reference input. Despite these methodological differences, both SHAP and IG results consistently revealed a high degree of overlapping features across all targets, demonstrating that the MTL model effectively leverages shared physical insights to enhance the predictive performance of each task.

The seven high-impact features derived from the IG analysis were visualized using t-SNE to examine whether the selected feature subset retains degradation-consistent local structure in the reduced feature space. Because t-SNE primarily preserves local neighborhood relationships, this analysis was used to assess local continuity rather than to

View Article Online
DOI: 10.1039/D6EB00066E

infer a rigorously ordered global degradation trajectory. As shown in **Fig. 5a-c**, neighboring regions in the embedding were associated with similar values of SoH, DCIR increase rate, and Max. temp., indicating that the selected features preserve meaningful variation related to battery degradation. These results suggest that a concise subset of seven high-impact features captures essential degradation-related information and can provide a useful qualitative representation of battery state. For complementary comparison, Principal Component Analysis (PCA) on the same feature set showed broadly consistent target gradients, although the local transitions were less clearly resolved than in t-SNE (**Fig. S6**).

Data efficiency of MTL under limited training data

Beyond the mechanistic insights discussed above, the MTL framework exhibited a clear advantage in data efficiency, maintaining superior predictive performance on the held-out test set even when the amount of training data was substantially reduced. As shown in **Fig. 5d**, the predictive performance on the held-out test set (15% of the total dataset) was compared while systematically varying the fraction of the remaining 85% training data used for model training. A distinct performance gap was observed when the availability of training data was limited. Notably, utilizing only 10% of the training data resulted in substantial performance gains with the MTL approach. Specifically, the RMSE was reduced by 23.0% for SoH, 53.2% for the DCIR increase rate, and 5.6% for Max. temp. relative to the STL baselines. This trend is particularly notable because NN models are generally vulnerable to performance degradation when training data are limited. However, this study confirms that the proposed MTL approach significantly enhances model robustness and prediction accuracy, even under data-scarce conditions. Such performance gains are driven by positive transfer and inductive bias acquired through co-training on highly correlated battery indicators, ensuring stable predictions despite limited training data.

Discussion

In this study, we proposed a robust MTL framework integrated with a rapid diagnostic protocol to simultaneously predict three critical battery indicators using short current pulse data applicable across the full SoC range. By leveraging their inherent correlations among the targets and utilizing a diverse array of directly collected degradation-related factors, the proposed MTL model significantly outperformed the STL baselines for SoH, DCIR increase rate, and Max. temp. Notably, the STL approach exhibited suboptimal performance for Max.

View Article Online
DOI: 10.1039/D6EB00066E

temp. prediction, particularly in the high-temperature region, due to the weak correlation between the temperature and the input data. However, the application of the MTL framework successfully overcame this limitation. The model demonstrated a substantial improvement in predictive accuracy for Max. temp., achieving reductions of 30.5% in MAE, 29.4% in RMSE, and 30.2% in MAPE. Beyond these performance metrics, our methodology demonstrates that multi-task knowledge sharing can effectively overcome the limitations of data-constrained battery research, providing a scalable template for broader electrochemical and materials systems.

It is important to note that the correlations between the diagnostic features and the target indicators arise from underlying electrochemical relationships rather than from implicit information leakage. The features are derived from voltage responses to controlled current pulses and rest periods, which reflect the intrinsic state of the battery but do not contain the target labels themselves. Therefore, the model learns physically meaningful mappings between measurable signals and degradation indicators, consistent with the fundamental principles of battery diagnostics.

To ensure the reliability of the dataset, the target indicators and diagnostic features were obtained through separate measurement procedures. The target indicators, including SoH, DCIR increase rate, and Max. temp., were independently measured after degradation, whereas the input features were extracted from voltage responses during the diagnostic protocol. In addition, a reset cycle was introduced before the diagnostic measurements to remove residual kinetic polarization accumulated during prior degradation cycles.[31] This procedure allowed the diagnostic features to reflect the intrinsic electrochemical state of the battery under standardized conditions. To further evaluate the possible influence of cell-specific correlations, we performed an additional cell-wise split analysis, in which all data from one physical cell were excluded from the training set and used as the test set. As shown in **Fig. S7**, the resulting errors were comparable to those from the original evaluation, indicating that the model was not primarily driven by cell-specific memorization but instead captured degradation-related features shared across different cells. While these procedures support the reliability of the dataset, they do not eliminate the need for comprehensive data collection across diverse degradation environments, which remains essential for accurately predicting various cell states.

As shown in **Fig. S8**, the correlation between SoH and the DCIR increase rate is highly dependent on the charging C-rate during the degradation cycle. This is because varying cycling

View Article Online
DOI: 10.1039/D6EB00066E

conditions trigger distinct degradation mechanisms, such as lithium plating, SEI layer growth, the loss of active material, and micro-cracking.[32-36] Consequently, if the training dataset lacks data from specific degradation pathways, the model's performance degrades significantly when evaluating such unseen conditions (**Fig. S9**). To ensure robust generalization, a well-balanced dataset encompassing diverse aging conditions is essential. Furthermore, generative models can be utilized to mitigate data scarcity issues, while Physics-Informed Neural Networks (PINNs), which integrate governing physical equations into the learning process, are being actively explored to enhance model reliability.

The significance of this study lies in the synergistic integration of a rapid diagnostic protocol with an MTL architecture, a combination that effectively addresses the challenge of limited data availability in battery research. Our framework demonstrates how a pulse-based protocol, designed to be applicable across the full range of SoC, can be coupled with shared feature representation to achieve high-fidelity diagnostics for multiple indicators simultaneously. This approach is particularly valuable for research and industrial contexts where characterizing diverse degradation pathways is experimentally expensive and time-intensive. Furthermore, by consolidating multiple estimation tasks into a single unified model, this architecture provides practical advantages in terms of computational efficiency and streamlined model management compared to independent task-specific models. Ultimately, this methodology serves as a versatile and scalable template for broader materials science research, providing a robust pathway for predicting interdependent properties in complex systems where experimental data is difficult or costly to obtain.

Conclusion

We integrated a rapid diagnostic protocol with an MTL architecture to achieve SoC-independent battery diagnostics, overcoming the limitations of conventional state-dependent models. By learning multiple correlated indicators simultaneously from a diverse dataset, our framework maintains superior performance across the full SoC range. The proposed model demonstrated high accuracy on the test set, achieving R2 scores of 0.988 for SoH, 0.997 for the DCIR increase rate, and 0.883 for Max. temp. Notably, the MTL framework significantly outperformed the STL baselines, reducing the RMSE by 25.9% for SoH, 25.0% for DCIR increase rate, and 29.4% for Max. temp. These performance gains are attributed to the joint learning of correlated targets, which promotes positive transfer and inductive bias, and also

View Article Online
DOI: 10.1039/D6EB00066E

improves robustness under limited-data conditions. More importantly, this study demonstrates that combining a rapid diagnostic protocol with an MTL framework is an effective strategy for simultaneously improving predictive accuracy and data efficiency in battery diagnostics. These findings provide a practical foundation for scalable diagnosis of multiple correlated battery indicators and suggest that this approach can be extended to other battery metrics and materials systems in data-scarce environments.

Experimental Section

Data generation

Commercial 18650-sized cylindrical Lithidum-Ion Battery (LIB) cells, INR18650-29E (NCA/Graphite, Samsung SDI), were used for data generation. We used a 5 V and 12 A maximum cycler (BTS-4008-5V12A-S1, Neware) for the cell degradation cycle test, discharge capacity check for State of Health (SoH) calculation, Direct Current Internal Resistance (DCIR) check, and three types of diagnostic protocol tests. All the LIB cells were tested in a chamber maintained in Room Temperature (RT, 25 ℃). Each LIB cell was cycled in the voltage range of 2.5 - 4.2 V with different charge and discharge C-rate conditions (charge: 0.5, 1.0, 2.0 C-rate, discharge: 0.5, 1.0, 2.0, 4.0 C-rate): a Constant Current (CC) / Constant Voltage (CV) mode for charging step (end condition: 0.02 C-rate cut-off) and a CC mode for discharge step at RT. The sample size of each cycle test was 3 EA except for the test condition with 4.0 C-rate discharge, 6 EA.

After the degradation cycle, to remove historical information during degradation, a reset cycle was conducted with 0.5 C-rate charging and discharging. Then, SoH, DCIR, and diagnostic protocol data were collected. The discharge capacity was measured at 0.5 C-rate (I_1) to check SoH, and then, DCIR was measured at 2.5 C-rate ($I_2 = 5*I_1$) discharge for 10 seconds at Depth of Discharge (DoD) 50%. The cell temperature was measured during all tests by attaching sensors to the external case of the cell near the Ni tap from the CT measurement. After all measurements, the diagnostic protocol tests were conducted, using the same method as in the previous study.[18, 37] The diagnostic protocol tests were categorized into three distinct SoC ranges: 0, 5% SoC (Diagnostic protocol 1, Low SoC), 10 to 75% SoC (Diagnostic protocol 2, medium SoC), and 80 to 100% SoC (Diagnostic protocol 3, High SoC). To set the target SoC levels, cells were discharged from 100% to 0% at 1.0 C-rate in 5% SoC intervals. At each SoC setpoint, diagnostic tests were conducted in

View Article Online
DOI: 10.1039/D6EB00066E

descending order from diagnostic protocol 3 to 1. During these tests, current pulses of 1.65 C-rate and 2.0 C-rate were applied for 5 seconds, followed by a 20-second rest time to collect voltage response data. Notably, for diagnostic protocol 1 and 3, a reduced current of 0.1 C-rate was applied during the discharge and charge steps, respectively, to ensure compliance with voltage cut-off conditions. We repeated the measurement every 30 cycles when SoH was 80% or above, and every 15 cycles when it fell below 80%.

Data preprocessing

From the collected data, the SoC and SoH, and DCIR increase rate were calculated using the following equations.

$$SoC\ [\%] = \frac{Remaining\ capacity}{Fully\ discharge\ capacity} \times 100$$

$$SoH\ [\%] = \frac{Discharge\ capacity}{Discharge\ capacity\ at\ BoL} \times 100$$

$$DCIR\ [\Omega] = \frac{V_1 - V_2}{I_2 - I_1}\ (V_1, V_2 = onset, offset\ voltage\ upon\ applying\ I_2)$$

$$DCIR\ increase\ rate\ [\%] = \frac{DCIR}{DCIR\ at\ BoL} \times 100$$

And maximum temperature (Max. temp.) was collected from the maximum temperature value during 0.5 C-rate charging and discharging for SoH calculation.

Input features were extracted from voltage responses obtained during diagnostic protocol measurements. The detailed methodology for feature generation is provided in **Supplementary Note S1**.

Machine Learning

To optimize the Multi-Task Learning (MTL) model for predicting three battery indicators (SoH, DCIR increase rate, Max. temp), comprehensive hyperparameter tuning was conducted. The multi-task neural network was constructed with 58 generated features in the input layer and the three battery indicators as the output layer. The activation functions for the hidden and output layers were ReLU and linear functions, respectively. The hidden layers were optimized by using Bayesian optimization via the KerasTuner library. The search space for the Bayesian optimization included the number of hidden layers (n = 2, 3, 4, 5, 6) and the number of nodes per hidden layer (m = 16, 32, 64, 128, 256). The

View Article Online
DOI: 10.1039/D6EB00066E

optimization was executed with 5 repetitions per trial for 200 iterations, using Mean Squared Error (MSE) as loss function. The maximum number of epochs and the batch size were set to 1,500 and 512, respectively.

To improve overall performance, the loss weights for SoH, the DCIR increase rate, and Max. temp. were set to 10, 1, and 100. This weighting strategy was specifically designed to balance the contribution of each task by compensating for the significant differences in their target ranges and statistical variances. Such an approach prevents the optimization process from being disproportionately biased toward indicators with larger numerical magnitudes, ensuring stable and concurrent learning across all tasks. Furthermore, higher weights were allocated to more challenging tasks, such as Max. temp., to compensate for its relatively lower correlation with the input features compared to the physically direct relationship observed for the DCIR increase rate. Early stopping was applied with a patience of 200 epochs to prevent overfitting. Lastly, to enhance learning stability, the learning rate of the optimizer was set to decay exponentially, starting from an initial value of 0.001 and multiplying 0.96 every 3,000 iterations.

For Single-Task Learning (STL), feature selection and model optimization, such as model selection and hyperparameter tuning, were conducted for each target (SoH, DCIR increase rate, and Max. temp.). Because most features were related to resistance, highly correlated features were removed to reduce model complexity. To remove these highly correlated features, the Pearson correlation coefficients among all generated features were compared, and if the absolute coefficient was higher than 0.75, all features except the one having the highest correlation with the target value were removed. After removing redundant features and applying min-max scaling, the optimal number of features and the best-performing ML model were selected by performing 30 iterations of 7-fold cross-validation using the PyCaret library in Python[38]. Based on the selected features and the ML models, hyperparameter tuning was conducted for each model using Bayesian optimization. (**Table. S1-4**)

To evaluate the performance of the three STL models and the MTL model, 15% of the full dataset at each SoC level was first held out as an independent test set, ensuring an even distribution across full SoC ranges. This test set was not used during any stage of model selection, hyperparameter tuning, architecture optimization, or early stopping. The remaining 85% of the data was used as the training pool. For the MTL model, 15% of the

View Article Online
DOI: 10.1039/D6EB00066E

training pool was further separated as a validation set for architecture optimization and early stopping. For the STL models, model selection and hyperparameter optimization were performed using 7-fold cross-validation within the training pool. After all optimization procedures were completed, final performance was evaluated only on the held-out test set using R^2 score, Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and Mean Absolute Percentage Error (MAPE).

$$R^2\ score = 1 - \frac{\sum_{i=1}^{n}(y^i - \hat{y}^i)^2}{\sum_{i=1}^{n}(y^i - \bar{y})^2}$$

$$MAE = \frac{\sum_{i=1}^{n}|y^i - \hat{y}^i|}{n}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(y^i - \hat{y}^i)^2}{n}}$$

$$MAPE = \frac{100}{n}\sum_{i=1}^{n}\left|\frac{y^i - \hat{y}^i}{y^i}\right|$$

where y^i was true value, $\hat{y}^i$ was predicted value of i[th] data, $\bar{y}$ was mean of true values, and n was the amount of data.

Analysis of the trained MTL model and t-SNE

To analyze the MTL model training process, a cosine similarity analysis between two targets at the shared layers was conducted. The similarity metric $Sim_i(A, B)$ at epoch i is defined as:

$$g_i(A) = \nabla_{\theta_{shared}} L_A$$

$$Sim_i(A, B) = \frac{g_i(A) \cdot g_i(B)}{|g_i(A)|\,|g_i(B)|}$$

where $g_i(A)$ and $g_i(B)$ denote the flattened gradient vectors of the target A and B losses with respect to the shared parameter, respectively. The cosine similarities of all combinations of the target were calculated every epoch.

IG analysis was performed to quantify the impacts of each feature, and $IG_i(x)$ was calculated as followes:

$$IG_i(x) = (x_i - x'_i) \times \int_{\alpha=0}^{1} \frac{\partial F(x' + \alpha \times (x - x'))}{\partial x_i} d\alpha$$

where x, x_i, x'_i, F, and α are input values, the mean value is the baseline, input feature

View Article Online
DOI: 10.1039/D6EB00066E

value for IG calculation, prediction model, and interpolation coefficient.[24] For simple calculation, the Riemann approximation was conducted as follows:

$$IG_i(x) \approx (x_i - x'_i) \times \frac{1}{m} \sum_{k=1}^{m} \frac{\partial F(x' + \frac{k}{m} \times (x - x'))}{\partial x_i}$$

where m is the number of steps. In this study, m was 50.

For visualization of high-dimensional data, t-distributed Stochastic Neighbor Embedding (t-SNE), which approximates Gaussian distributions that model point neighborhoods in the original space with heavy-tailed t-distributions in the embedding space, was applied.[39] By adopting these two different distributions, t-SNE maintains local neighborhoods and creates stronger repulsion between dissimilar points to avoid crowding. The dimension of data was reduced from 7 to 2.

Author contributions

M. G., Y. C. and D.-H. S conceived the concept and designed the entire process, with D.-H. Seo supervising the project. With advice from H. M. Y. and S. B. M., M. G. and Y. C. developed the entire process from data collection to machine learning model training. Y. C. designed diagnostic protocols and collected all of data. M. G. conducted data preprocessing, machine learning model optimization. Thereafter, M. G., Y. C., H. M. Y., S. B. M., and D.-H. S. wrote and revised the manuscript. All the authors reviewed and provided feedback on the manuscript.

Conflicts of interest

The authors declare no competing interests.

Data availability

The code and processed data used in this work have been deposited at Zenodo at https://doi.org/10.5281/zenodo.19204562.

Acknowledgments

This research was supported by SK ecoplant and the National Research Foundation of Korea (NRF) funded by Ministry of Science and ICT (RS-2024-00435493 and RS-2024-

View Article Online
DOI: 10.1039/D6EB00066E

00451160).

Declaration of generative AI and AI-assisted technologies in the writing process

During the preparation of this work, the authors used Gemini (by Google) to improve readability and language. After using this tool, the authors reviewed and edited the content as needed and took full responsibility for the content of the publication.

References

1. X. Zhu, Z. Sun, Z. Wang, H. Wang, N. Lin and C. Shan, Journal of Energy Storage, 2023, 68, 107768.
2. I. Andriunas, Z. Milojevic, N. Wade and P. K. Das, Journal of Power Sources, 2022, 525, 231126.
3. J. Zhu, Y. Wang, Y. Huang, R. Bhushan Gopaluni, Y. Cao, M. Heere, M. J. Mühlbauer, L. Mereacre, H. Dai, X. Liu, A. Senyshyn, X. Wei, M. Knapp and H. Ehrenberg, Nature Communications, 2022, 13, 2261.
4. S. Park, H. Lee, Z. K. Scott-Nevros, D. Lim, D.-H. Seo, Y. Choi, H. Lim and D. Kim, Materials Horizons, 2023, 10, 1274-1281.
5. S. Tao, R. Ma, Z. Zhao, G. Ma, L. Su, H. Chang, Y. Chen, H. Liu, Z. Liang, T. Cao, H. Ji, Z. Han, M. Lu, H. Yang, Z. Wen, J. Yao, R. Yu, G. Wei, Y. Li, X. Zhang, T. Xu and G. Zhou, Nature Communications, 2024, 15, 10154.
6. D. I. Stroe and E. Schaltz, IEEE Transactions on Industry Applications, 2020, 56, 678-685.
7. S. Zhang, B. Zhai, X. Guo, K. Wang, N. Peng and X. Zhang, Journal of Energy Storage, 2019, 26, 100951.
8. Q. Li, R. Song and Y. Wei, Journal of Energy Storage, 2025, 118, 116078.
9. D. Kim, K. Koo, J. J. Jeong, T. Goh and S. W. Kim, Energies, 2013, 6, 5538-5551.
10. C. Nguyen Van and D. T. Quang, International Journal of Electrochemical Science, 2023, 18, 100166.
11. A. Barai, K. Uddin, W. D. Widanage, A. McGordon and P. Jennings, Scientific Reports, 2018, 8, 21.
12. X. Liu, P. He, W. Chen and J. Gao, ArXiv, 2019, abs/1901.11504.
13. Z. Zhang, P. Luo, C. C. Loy and X. Tang, Year.
14. J. Jumper, R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K.

View Article Online
DOI: 10.1039/D6EB00066E

Tunyasuvunakool, R. Bates, A. Žídek, A. Potapenko, A. Bridgland, C. Meyer, S. A. A. Kohl, A. J. Ballard, A. Cowie, B. Romera-Paredes, S. Nikolov, R. Jain, J. Adler, T. Back, S. Petersen, D. Reiman, E. Clancy, M. Zielinski, M. Steinegger, M. Pacholska, T. Berghammer, S. Bodenstein, D. Silver, O. Vinyals, A. W. Senior, K. Kavukcuoglu, P. Kohli and D. Hassabis, Nature, 2021, 596, 583-589.

15. H. Tang, J. Liu, M. Zhao and X. Gong, presented in part at the Proceedings of the 14th ACM Conference on Recommender Systems, Virtual Event, Brazil, 2020.
16. W. Li, H. Zhang, B. van Vlijmen, P. Dechent and D. U. Sauer, Energy Storage Materials, 2022, 53, 453-466.
17. D. Li, Y. Xing and X. Li, Journal of Energy Storage, 2026, 161, 121823.
18. M. Gong, Y. Choi, H. M. Yang, S. B. Ma and D.-H. Seo, Energy Storage Materials, 2026, 88, 105060.
19. M. Fischer, M. J. Brand, A. Karger, M. R. Gomez, M. Rehm, J. Natterer and A. Jossen, Journal of Power Sources, 2025, 656, 237921.
20. W. Shen, N. Wang, J. Zhang, F. Wang and G. Zhang, ACS Omega, 2022, 7, 44733-44742.
21. V. Tshitoyan, J. Dagdelen, L. Weston, A. Dunn, Z. Rong, O. Kononova, K. A. Persson, G. Ceder and A. Jain, Nature, 2019, 571, 95-98.
22. A. Rives, J. Meier, T. Sercu, S. Goyal, Z. Lin, J. Liu, D. Guo, M. Ott, C. L. Zitnick, J. Ma and R. Fergus, Proceedings of the National Academy of Sciences, 2021, 118, e2016239118.
23. R. Caruana, Machine Learning, 1997, 28, 41-75.
24. M. Sundararajan, A. Taly and Q. Yan, Year.
25. R. Sayres, A. Taly, E. Rahimy, K. Blumer, D. Coz, N. Hammel, J. Krause, A. Narayanaswamy, Z. Rastegar, D. Wu, S. Xu, S. Barb, A. Joseph, M. Shumski, J. Smith, A. B. Sood, G. S. Corrado, L. Peng and D. R. Webster, Ophthalmology, 2019, 126, 552-564.
26. K. Preuer, G. Klambauer, F. Rippmann, S. Hochreiter and T. Unterthiner, in *Explainable AI: interpreting, explaining and visualizing deep learning*, Springer, 2019, pp. 331-345.
27. B. Stiaszny, J. C. Ziegler, E. E. Krauß, J. P. Schmidt and E. Ivers-Tiffée, Journal of Power Sources, 2014, 251, 439-450.
28. F. Leng, C. M. Tan and M. Pecht, Scientific Reports, 2015, 5, 12967.
29. S. Ma, M. Jiang, P. Tao, C. Song, J. Wu, J. Wang, T. Deng and W. Shang, Progress in

View Article Online
DOI: 10.1039/D6EB00066E

Natural Science: Materials International, 2018, 28, 653-666.

30. C. Schmitt, D. Kopljar and K. A. Friedrich, Journal of Energy Storage, 2025, 120, 116486.
31. A. Geslin, B. van Vlijmen, X. Cui, A. Bhargava, P. A. Asinger, R. D. Braatz and W. C. Chueh, Joule, 2023, 7, 1956-1965.
32. H. Wang, S. Frisco, E. Gottlieb, R. Yuan and J. F. Whitacre, Journal of Power Sources, 2019, 426, 67-73.
33. T. Waldmann, M. Wilka, M. Kasper, M. Fleischhammer and M. Wohlfahrt-Mehrens, Journal of Power Sources, 2014, 262, 129-135.
34. J. Vetter, P. Novák, M. R. Wagner, C. Veit, K. C. Möller, J. O. Besenhard, M. Winter, M. Wohlfahrt-Mehrens, C. Vogler and A. Hammouche, Journal of Power Sources, 2005, 147, 269-281.
35. D. N. Wong, D. A. Wetz, A. M. Mansour and J. M. Heinzel, Year.
36. M. Dubarry, C. Truchot, B. Y. Liaw, K. Gering, S. Sazhin, D. Jamison and C. Michelbacher, Journal of The Electrochemical Society, 2013, 160, A191-A199.
37. M. Gong, Y. Choi, H. M. Yang, S. B. Ma and D.-H. Seo, 2025.
38. PyCaret: An open source, low-code machine learning library in Python, https://pycaret.org/, (accessed.)
39. L. Van Der Maaten, The journal of machine learning research, 2014, 15, 3221-3245.

View Article Online
DOI: 10.1039/D6EB00066E

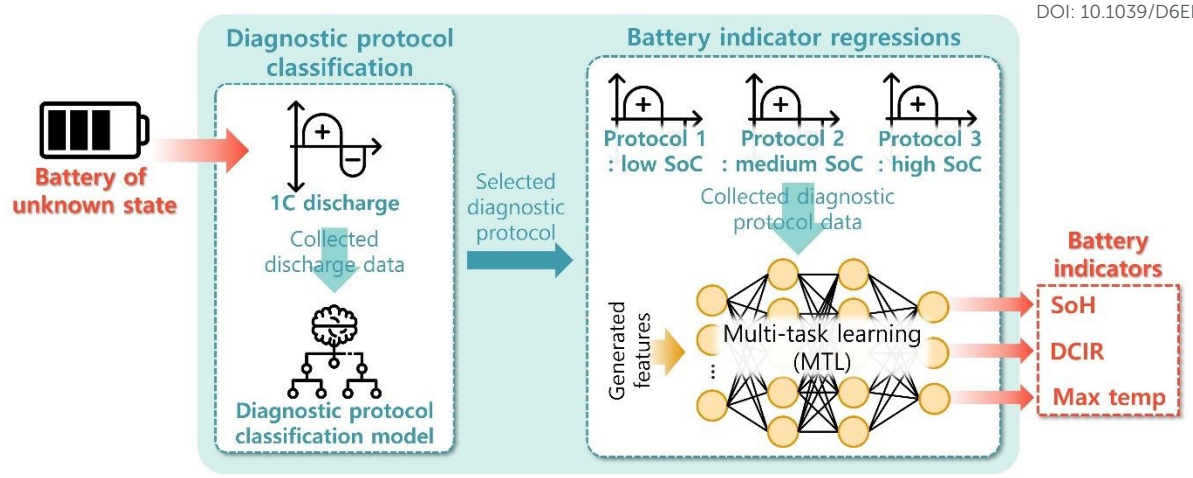

Fig. 1 Overall battery indicators prediction process for the batteries through multi-task learning (MTL). MTL simultaneously predicts 3 battery indicators, state of health (SoH), direct current internal resistance (DCIR), and maximum temperature during 0.5 C-rate charge and discharge (Max. temp.).

View Article Online
DOI: 10.1039/D6EB00066E

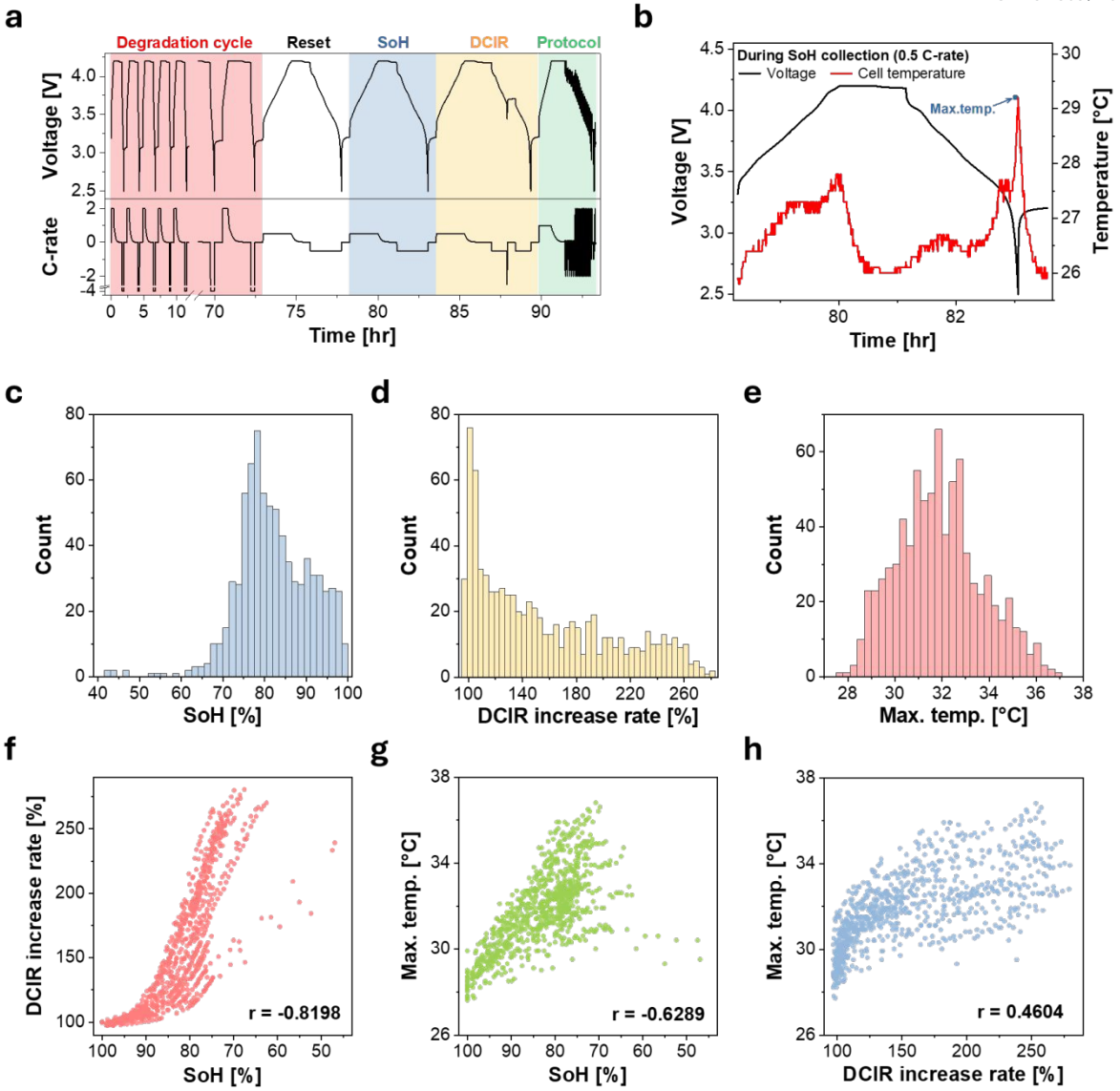

Fig. 2 Data collection and relations between battery indicators **(a)** Data collection method for battery indicator prediction. 3 battery indicators (SoH, DCIR, and Max. temp.) were measured after degradation and reset cycle. Then, diagnostic protocol data were recorded at 5% intervals of the SoCs. Reset cycle was conducted with 0.5 C-rate charge and discharge after degradation cycle. SoH were calculated from capacity with 0.5 C-rate charge and discharge. DCIR were collected from resistance after 5 × I_1 (I_1 = 0.5 C-rate) current pulse for 10 seconds at depth of discharge (DoD) 50%. DCIR increase rate were calculated by the ratio of DCIR to DICR of beginning of life (BoL). **(b)** Max. temp. were collected maximum temperature during SoH measurement. The distribution of the collected battery indicator values **(c)** SoH, **(d)** DCIR increase rate, and **(e)** Max. temp. The relations between battery indicators **(f)** SoH and DCIR increase rate, **(g)** SoH and Max. temp., and **(h)** DCIR increase rate and Max. temp. (r = Pearson

View Article Online
DOI: 10.1039/D6EB00066E

correlation coefficient)

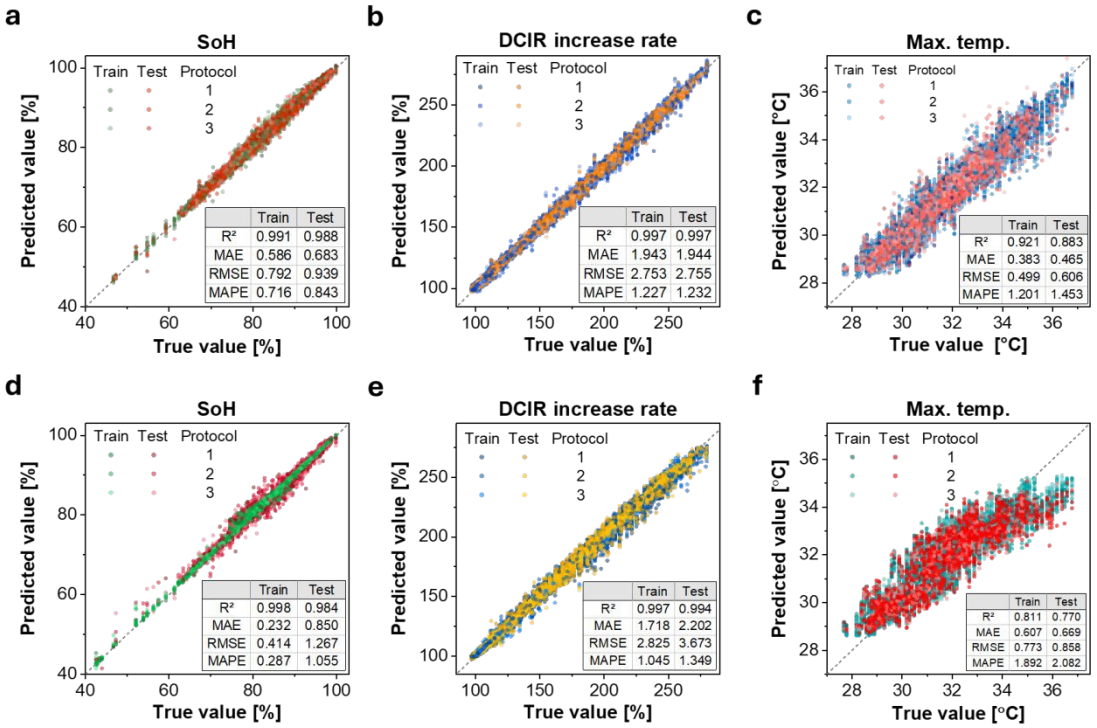

Fig. 3 Battery indicator prediction results of MTL **(a)** SoH, **(b)** DCIR increase rate, and **(c)** Max. temp. Battery indicator prediction results of single-task learning (STL) **(d)** SoH, **(e)** DCIR increase rate, and **(f)** Max. temp. The test data ratio was 15%.

View Article Online
DOI: 10.1039/D6EB00066E

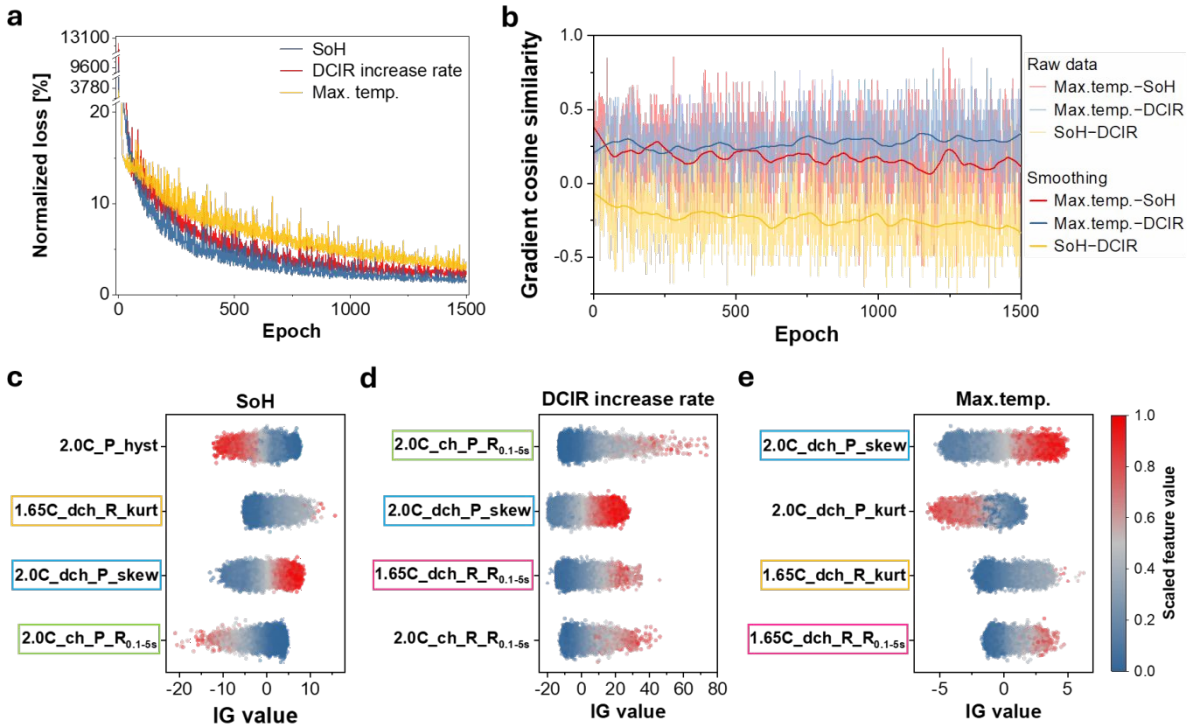

Fig. 4 Analysis of prediction performance improvement factors for MTL **(a)** Normalized loss of three battery indicators for each epoch. The losses were normalized based on the difference between the maximum and minimum value of target. **(b)** Gradient cosine similarity analysis during MTL model training. Cosine similarity for each epoch was calculated based on the cosine of the losses between the two targets. Integrated gradient (IG) analysis results of top 4 features for each battery indicator predictions **(c)** SoH, **(d)** DCIR increase rate, and **(e)** Max. temp. The colored (yellow, sky blue, light green and pink) boxes represented overlapped features.

View Article Online
DOI: 10.1039/D6EB00066E

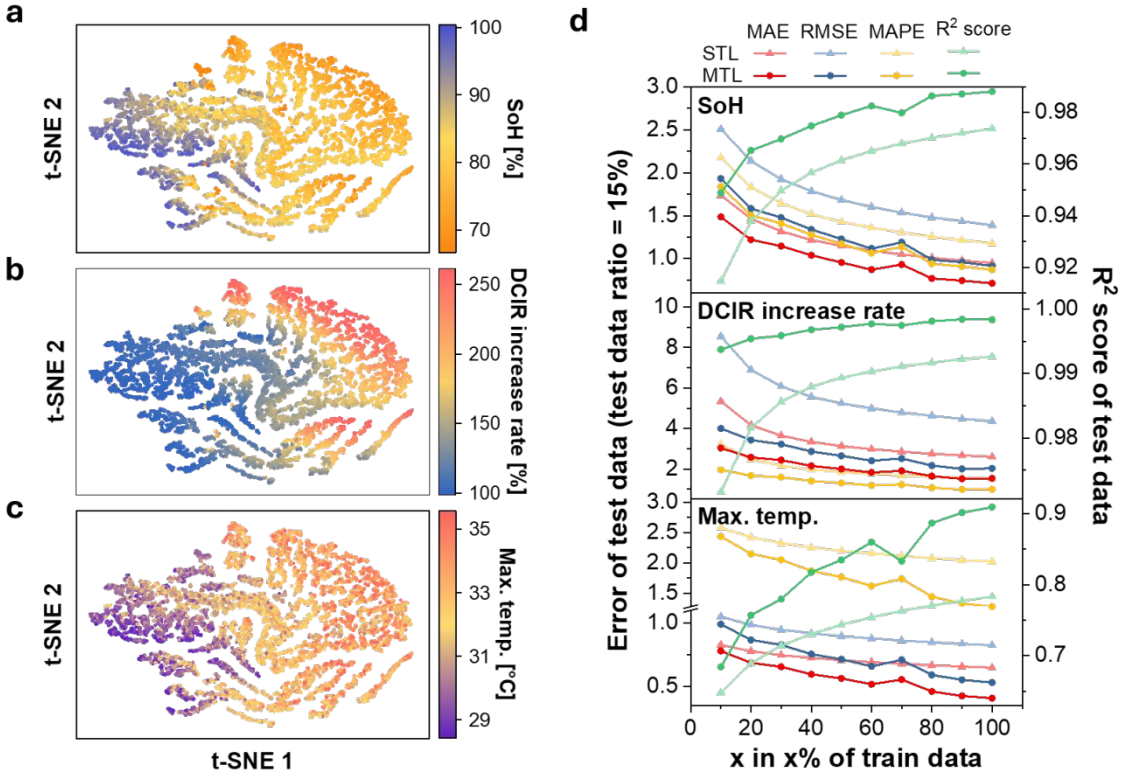

Fig. 5 t-SNE (t-distributed stochastic neighbor embedding) results obtained from 7 selected features derived from the top 4 IG values per battery indicator prediction. Color mappings were represented **(a)** SoH, **(b)** DCIR increase rate, and **(c)** Max. temp. **(d)** Prediction performance of STL and MTL according to the training data ratio. The value was averaged over 100 repetitions. The error was calculated using the test data (test data ratio = 15%), which was evaluated on the trained model that utilized x% of the total training data.

View Article Online
DOI: 10.1039/D6EB00066E

Data availability statement for

Simultaneous Prediction of Three Key Li-Ion Battery Indicators through Multi-Task Learning

Data for this article, including code and raw data files are available at Zenodo at https://doi.org/10.5281/zenodo.19204562.

The data supporting this article have been included as part of the Supplementary Information: Supplementary Note. S1, Fig S1-S9, and Table S1-S7.