*sustainability*

MDPI

*Article*

# Optimizing Lithium-Ion Battery Performance: Integrating Machine Learning and Explainable AI for Enhanced Energy Management

**Saadin Oyucu [1,*], Betül Ersöz [2], Şeref Sağıroğlu [2], Ahmet Aksöz [3] and Emre Biçer [4]**

[1] Faculty of Engineering, Department of Computer Engineering, Adıyaman University, 02040 Adıyaman, Türkiye
[2] Artificial Intelligence and Big Data Analytics Security R&D Center, Gazi University, 06570 Ankara, Türkiye; betul.ersoz@gazi.edu.tr (B.E.); ss@gazi.edu.tr (Ş.S.)
[3] Mobilers Team, Sivas Cumhuriyet University, 58050 Sivas, Türkiye; aaksoz@cumhuriyet.edu.tr
[4] Battery Research Laboratory, Faculty of Engineering and Natural Sciences, Sivas University of Science and Technology, 58010 Sivas, Türkiye; emre.bicer@sivas.edu.tr
* Correspondence: saadinoyucu@adiyaman.edu.tr

**Abstract:** Managing the capacity of lithium-ion batteries (LiBs) accurately, particularly in large-scale applications, enhances the cost-effectiveness of energy storage systems. Less frequent replacement or maintenance of LiBs results in cost savings in the long term. Therefore, in this study, AdaBoost, gradient boosting, XGBoost, LightGBM, CatBoost, and ensemble learning models were employed to predict the discharge capacity of LiBs. The prediction performances of each model were compared based on mean absolute error (MAE), mean squared error (MSE), and R-squared values. The research findings reveal that the LightGBM model exhibited the lowest MAE (0.103) and MSE (0.019) values and the highest R-squared (0.887) value, thus demonstrating the strongest correlation in predictions. Gradient boosting and XGBoost models showed similar performance levels but ranked just below LightGBM. The competitive performance of the ensemble model indicates that combining multiple models could lead to an overall performance improvement. Furthermore, the study incorporates an analysis of key features affecting model predictions using SHAP (Shapley additive explanations) values within the framework of explainable artificial intelligence (XAI). This analysis evaluates the impact of features such as temperature, cycle index, voltage, and current on predictions, revealing a significant effect of temperature on discharge capacity. The results of this study emphasize the potential of machine learning models in LiB management within the XAI framework and demonstrate how these technologies could play a strategic role in optimizing energy storage systems.

**Keywords:** Li-ion; BMS; SoH estimation; ensemble learning; explainable AI; SHAP

check for updates

**Citation:** Oyucu, S.; Ersöz, B.; Sağıroğlu, Ş.; Aksöz, A.; Biçer, E. Optimizing Lithium-Ion Battery Performance: Integrating Machine Learning and Explainable AI for Enhanced Energy Management. *Sustainability* **2024**, *16*, 4755. https://doi.org/10.3390/su16114755

Academic Editors: Mobyen Uddin Ahmed and Rosina O. Weber

Received: 13 May 2024
Revised: 31 May 2024
Accepted: 31 May 2024
Published: 3 June 2024

**Copyright:** © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

## 1. Introduction

Li-ion batteries (LiBs) are crucial energy sources for electric vehicles (EVs), offering advantages such as high energy density, lightweight, low self-discharge rates, fast charging capabilities, and minimal maintenance requirements. These qualities have established LiBs as the preferred power option for EVs across various applications. The shift toward electrification has further underscored LiBs' appeal as environmentally friendly alternatives, emitting fewer greenhouse gases than traditional fossil fuel vehicles and reducing overall carbon footprints [1].

Despite their widespread use and benefits, LiBs face challenges related to gradual performance decline due to chemical and physical factors. This necessitates the development of efficient battery management systems to address high costs and limited repair opportunities. Accurately predicting battery health indicators is crucial for effectively managing performance and ensuring long-term reliability [2].

The state of health (SoH) metric is pivotal in evaluating LiB performance, serving as a crucial parameter for various applications, including portable electronics, EVs, and grid-scale energy storage. SoH reflects a battery's current condition and capacity relative to its original state, guiding predictions of remaining useful life and overall performance [3–6]. It encompasses critical aspects such as capacity degradation, internal resistance changes, cycle life, and safety considerations, playing a pivotal role in enhancing safety, sustainability, and cost-effectiveness across industries [7–9].

SoH evaluation is essential for maintaining range predictability, charging efficiency, and overall vehicle performance. For grid-scale applications like renewable energy integration and peak shaving, understanding battery SoH maximizes energy utilization, optimizes investments, and mitigates operational risks [10].

Advancements in SoH assessment methodologies, including machine learning algorithms and diagnostic technologies, are driving the development of more intelligent battery management systems capable of real-time monitoring, predictive maintenance, and adaptive control strategies. This proactive approach not only prolongs battery lifespan but also promotes sustainable practices by minimizing premature replacements and optimizing resource utilization and recycling efforts [11,12]. Namely, various advanced techniques are available for predicting the performance of lithium-ion batteries, including molecular dynamics simulations and density functional theory (DFT). Molecular dynamics simulations provide insights into the atomic and molecular interactions within the battery, helping us to understand the behavior of materials under different conditions. DFT, on the other hand, allows for the calculation of electronic properties and can predict how changes at the atomic level impact overall battery performance. These techniques, along with others such as machine learning models and electrochemical modeling, offer comprehensive tools for researchers to explore and enhance battery materials and design. By leveraging these methods, significant advancements can be made in optimizing battery performance, lifespan, and safety, contributing to the development of more efficient and reliable energy storage systems [13,14]. The significance of SoH in Li-ion batteries extends beyond performance metrics, representing a strategic imperative for innovation, sustainability, and reliability in energy storage technologies.

The implementation of explainable artificial intelligence (XAI) techniques in lithium-ion batteries is crucial as it enhances the transparency and interpretability of predictive models, allowing for better understanding and management of battery performance and health. The main novelty of this work is the integration of XAI techniques as an alternative to traditional methods used in SoH assessment of LiBs. Current battery management systems are still inadequate for detecting degradation and aging processes affecting battery performance promptly. The proposed XAI-based approach describes machine learning and artificial intelligence models that work like a closed box to predict battery health indicators more accurately. Thus, the features affecting the models have the potential to maximize lifetime and minimize maintenance costs. In addition, this method aims to reduce environmental impact and make resource use more sustainable by optimizing battery replacement times. Therefore, within the scope of the study, a strategic innovation is presented in terms of sustainability and reliability in the field of battery technologies, especially with the use of XAI.

Remaining useful life (RUL) and SoH metrics act as key determinants of battery performance, durability, and reliability across a broad spectrum of applications. Understanding the interplay between LiBs, RUL, and SoH is essential for optimizing energy storage systems, enhancing operational efficiency, and ensuring sustainable practices in the evolving battery technology landscape. To make it happen, explainable artificial intelligence is utilized to estimate the SoH of a Li-ion battery; as far as we know, this paper is the first time to tackle such an estimation method. The integration of explainable artificial intelligence (XAI) techniques has emerged as a critical aspect in assessing RUL and SoH metrics for Li-ion in various applications. As industries increasingly rely on complex systems and

data-driven approaches, the need for transparent and interpretable AI models to predict RUL and monitor SoH has become paramount [15].

In our study, we applied various machine learning techniques, including principal component analysis (PCA), linear regression, ridge regression, k-nearest neighbors (k-NN), random forest, polynomial regression, and gradient boosting, to predict the SoH of Li-ion batteries. We compared the performance of these models in terms of accuracy, computational efficiency, and interpretability. Specifically, we focused on the explainability of each model by using techniques such as SHAP (Shapley additive explanations) values, which provide insights into the contribution of each feature to the model's predictions. Our results demonstrate that while complex models like random forest and gradient boosting offer higher accuracy, simpler models such as linear regression and k-NN, combined with explainable AI techniques, provide a more transparent understanding of the factors influencing battery degradation. Also, we believe that these models are not only applicable to Li-ion batteries but also Na-ion batteries [13,14,16,17]. This study focuses on predicting the discharge capacity of lithium-ion batteries, integrating machine learning (ML) and XAI techniques in this domain. In the second section, the materials and methodology used are explained in detail, while in the third section, experimental results and findings obtained from SHAP (Shapley additive explanations) analyses are presented. The fourth section discusses the obtained results, and in the fifth and final section, the study's conclusions are summarized, along with future trends and recommendations.

## 2. Materials and Methods

The structure of the study, the dataset used, and the machine learning and artificial intelligence methods employed are presented in this section. To ensure the study's reproducibility and enhance the methodology's transparency, each method and the parameters used are detailed extensively. Scientific justifications for the selection of each material and method were thoroughly discussed. A flowchart illustrating, step-by-step, how a model was developed using machine learning algorithms such as adaptive boosting (AdaBoost), gradient boosting, extreme gradient boosting (XGBoost), light gradient boosting machine (LightGBM), and category boosting (CatBoost) employed in the study is provided in Figure 1.

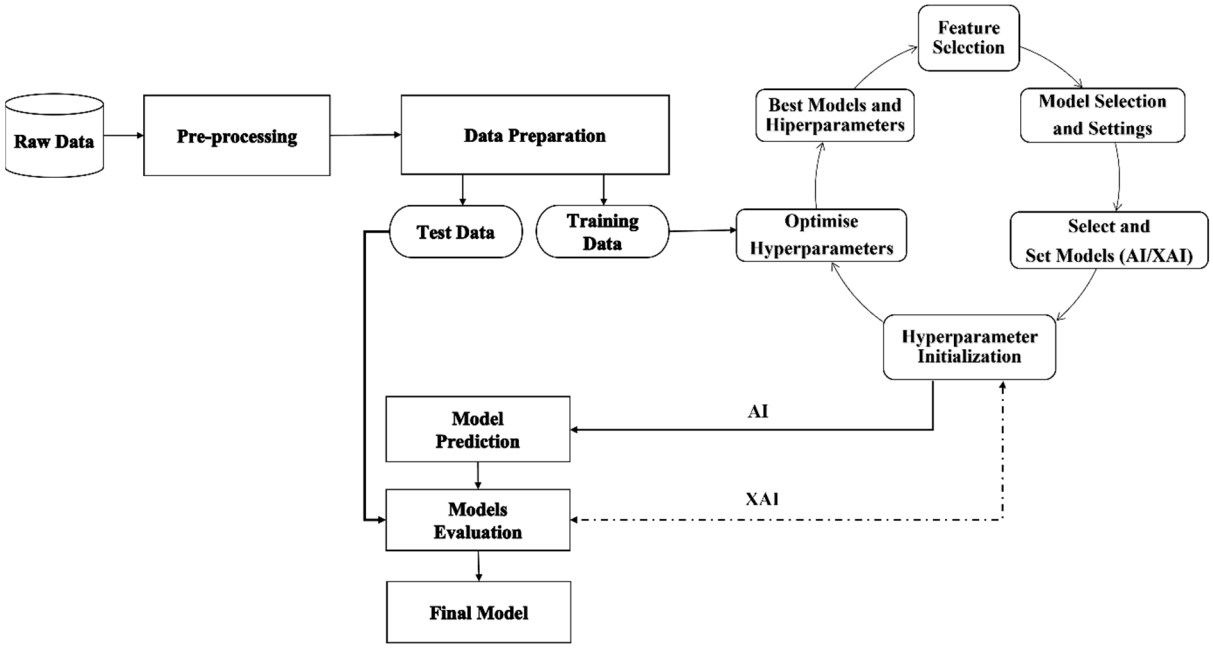



**Figure 1.** Flowchart of models.

The flowchart depicted in Figure 1 gives more details about the training and test procedures of the presented work. This flowchart clearly illustrates the operations performed at each step of the model development process, ensuring transparency and repeatability. The training and test datasets were applied to the models after the necessary preprocessing steps were applied to the raw data. In the final step, the model was evaluated.

### *2.1. Dataset Definition*

The dataset used in this study consists of 45,699 records containing the characteristics of battery discharge cycles. The dataset was divided into 80% for training and 20% for testing. A total of 38,845 data points were used for training, while 6854 data points were used for testing. Features such as "Temperature", "Current (A)", "Voltage (V)", "Cycle Index", and "Discharge Capacity (Ah)" were necessary to describe the electrical performance of batteries included in the dataset. The dataset used in the study was obtained from a comprehensive database containing the characteristics of battery discharge cycles. Various parameters measured during the battery discharge process were recorded in each cycle [2].

This study includes LIB samples tested for 750 cycles. All the charge/discharge tests were conducted within the voltage range of 3.0 to 4.4 V, using various C-rates and temperatures. Each test condition involved 8 cells, totaling 192 cells (Table 1). For each condition, the 8 cells were subjected to four different temperatures (10, 25, 45, and 60), three different discharge C-rates (0.7 C, 1 C, and 2 C), and two different charge cut-off C-rates (C/5 and C/40) using the constant current constant voltage (CCCV) protocol. All the tests were sourced from open-access data provided by Michael Pecht at the Center for Advanced Life Cycle Engineering (CALCE), University of Maryland. The experiments were conducted using LiCoO2 (cathode)–graphite (anode) cells [18]. The charge and discharge datasets obtained from these tests were subsequently used to train machine learning models. The first 300 cycles in the dataset were used for training in the preprocessing step. In this regard, 450 cycles between 301 and 750 were removed from the training and test datasets. The dataset underwent various preprocessing steps to make it more suitable for direct use. Among these processes, handling missing data was performed, and it was observed that there were no missing values in the dataset. Detection and correction of outliers were ensured. This process was carried out to provide a balanced approach during the training and evaluation stages of the model. The first 300 cycles of the dataset were used for training and test purposes.

**Table 1.** Test procedures applied to cells [18].

| Discharge C-Rate | Ambient Temperature (°C) | | | | Discharge Cut-Off C-Rate |
|---|---|---|---|---|---|
| | 10 | 25 | 45 | 60 | |
| 0.7 C | Test No.1 | Test No.7 | Test No.13 | Test No.19 | C/5 |
| | Test No.2 | Test No.8 | Test No.14 | Test No.20 | C/40 |
| 1 C | Test No.3 | Test No.9 | Test No.15 | Test No.21 | C/5 |
| | Test No.4 | Test No.10 | Test No.16 | Test No.22 | C/40 |
| 2 C | Test No.5 | Test No.11 | Test No.17 | Test No.23 | C/5 |
| | Test No.6 | Test No.12 | Test No.18 | Test No.24 | C/40 |

### *2.2. Model Selection*

This study evaluated and compared machine learning algorithms such as AdaBoost, gradient boosting, XGBoost, LightGBM, and CatBoost for model training. The selection of these algorithms was based on their unique characteristics and advantages. Each algorithm has its strengths and weaknesses depending on the dataset's features and desired outcomes. Comparing and evaluating these algorithms is essential for determining the most suitable one for a specific dataset. Ensemble methods like AdaBoost, gradient boosting, XGBoost, LightGBM, and CatBoost were among these algorithms. Each algorithm underwent training

on the dataset's training subset and was then evaluated for performance on the test subset. Additionally, a general overview of SHAP, one of the XAI methods, was introduced to provide and interpret the data of the results and the LightGBM model.

AdaBoost, gradient boosting, XGBoost, LightGBM, and CatBoost are boosted tree-based algorithms widely used in machine learning [19]. AdaBoost improves performance by creating a series of weak classifiers that correct classification errors. AdaBoost is mainly known for its sensitivity to noisy data and outliers. However, while it performs well on small and relatively simple datasets, the risk of overfitting increases on complex data structures [20].

Gradient boosting builds a model based on error correction by optimizing consecutive weak learners. Each new tree is constructed to reduce the errors of the previous one. Gradient boosting provides an advantage with its high degree of customizability. However, optimizing model-tuning parameters can be a time-consuming process [19]. XGBoost is designed to enhance the gradient boosting method in terms of scalability and speed. It stands out for its fast operation on large datasets and regularization features that enhance the model's performance. XGBoost can handle missing data and provides built-in support for cross-validation [21].

LightGBM is an algorithm capable of processing large datasets and consuming less memory. As the dataset's size and number of features increase, the advantages of LightGBM become more apparent. It utilizes a leaf-wise algorithm that grows datasets in a leaf-wise manner. This leaf-wise structure enables the model to learn faster [19]. CatBoost has explicitly been optimized to process categorical variables. This algorithm may automatically convert categorical features, thus reducing the need for feature engineering. CatBoost generally possesses a robust structure designed to reduce the risk of overfitting. However, iterative parameter tuning may still be required for optimal performance [22].

Our dataset's large scale and diverse features significantly influenced the selection of these algorithms. Each algorithm was evaluated based on its capacity to adapt to the complexity and size of the dataset. The ability of XGBoost and LightGBM algorithms to effectively handle large datasets and CatBoost's capability to process categorical data are reasons they were specifically preferred for the dataset. The SHAP method was employed to enhance the interpretability and transparency of the model [23]. SHAP allows us to understand which features influence model predictions and to what extent. Especially in critical applications like predicting the lifecycle of LiBs, it strengthens decision-making processes and guides how the model may be improved. Therefore, the integration of boosted tree-based algorithms and the SHAP method used in this study enhances the model's performance and adds significant scientific value by elucidating the dynamics underlying model decisions.

## 3. Experimental Results

Performance evaluations of the examined algorithms, comparative analyses, and the prediction success of each model on SoH of LiBs are developed and discussed in this section under different subsections. Additionally, the effects of the SHAP method on understanding model predictions and determining the decisions are then evaluated.

### *3.1. Model Evaluation Metrics*

The performance of the models was evaluated using standard regression metrics such as MSE (mean squared error), MAE (mean absolute error), and $R^2$ (R-squared). These metrics measure how well the models' predictions agree with actual values and how well the models generalized the task with high accuracy [24].

MSE values measure how far each model's predictions are from the actual values. A lower MSE indicates better model performance.

$$\mathrm{MSE} = \frac{1}{N}\sum_{i}^{N}\left(Y_i - Y_{predict(i)}\right)^2 \quad (1)$$

$R^2$ (R-squared): $R^2$ values measure how much the model explains the variance of the dependent variable and of the independent variables. An $R^2$ value approaching 1 indicates that the model explains most of the dataset and provides a good fit.

$$R^2 = 1 - \frac{\sum_i^N \left(Y_{predict(i)} - Y_i\right)^2}{\sum_i^N (Y_i - Y_{mean})^2} \quad (2)$$

MAE values measure how far each model's predictions are from actual values. A lower MAE indicates better model performance:

$$MAE = \frac{1}{N}\sum_{i=1}^{N}\left|Y_i - Y_{predict}\right| \quad (3)$$

### *3.2. Model Training*

Five learning algorithms were used for model training: AdaBoost, gradient boosting, XGBoost, LightGBM, and CatBoost. These algorithms aim to create a robust model by combining multiple weak learners using ensemble learning methods [25]. These methods involve combining predictions from multiple individual models created using different algorithms or configurations of the same algorithm to create a more robust model [26]. Ensemble learning improves prediction performance by balancing the weaknesses of a single model. This approach significantly enhances the accuracy of model predictions by leveraging the strengths of multiple models and compensating for their weaknesses. However, the effectiveness of this method may vary, and evaluating their performance on specific datasets is crucial to determine their suitability for a particular task. The dataset used in this study and the importance of ensemble learning modeling enable us to achieve prediction accuracy and reliability, especially in predicting SoH of LiBs, which individual models may not achieve alone. This contributes to the model providing more effective and reliable results in real-world applications. The hyperparameters of the models used in the study are detailed in Table 2.

**Table 2.** Hyperparameters of models.

| Models | Hyperparameters |
|---|---|
| AdaBoost | n_estimators = 100<br>learning_rate = 0.2<br>loss = "linear" |
| Gradient boosting | n_estimators = 100<br>learning_rate = 0.1<br>max_depth = 3<br>min_samples_split = 2<br>min_samples_leaf = 1 |
| XGBoost | n_estimators = 100<br>learning_rate = 0.1<br>max_depth = 3<br>subsample = 0.8<br>colsample_bytree = 0.8<br>gamma = 0<br>min_child_weight = 1 |
| LightGBM | n_estimators = 100<br>random_state = 42 |
| CatBoost | n_estimators = 100<br>random_state = 42<br>verbose = 0 |

The model hyperparameters in Table 2 were determined for AdaBoost, gradient boosting, XGBoost, LightGBM, and CatBoost models. For each model, the number of iterations ("n_estimators") was set to 100, while the learning rates and other specific parameters were chosen to optimize the model's fit to the data and the speed of the learning process. In particular, the "loss" function was set to "linear" in AdaBoost, branching criteria in gradient boosting, sampling and splitting criteria in XGBoost, and "random_state" settings in LightGBM and CatBoost to ensure repeatability. These parameters play a decisive role in model performance by determining how much deep learning the models will perform on the data during training and their adaptation speed. Lastly, the CatBoost model was configured with a "random_state" to 42, and a "verbose" level to 0. These hyperparameters directly affect how deeply each model will learn during the training process, how quickly it will adapt, and how well it will fit the data, thus playing a decisive role in the overall performance of the model.

### *3.3. Model Comparisons Results*

Performance metrics obtained when various boosting ensemble learning methods (AdaBoost, gradient boosting, XGBoost, LightGBM, CatBoost) applied to our dataset are shown in Table 3. Each metric is used to measure the accuracy of the model's predictions.

**Table 3.** Model comparisons results.

| Models | MAE | MSE | R-Squared |
|---|---|---|---|
| AdaBoost | 0.134 | 0.041 | 0.763 |
| Gradient boosting | 0.108 | 0.023 | 0.864 |
| XGBoost | 0.110 | 0.023 | 0.864 |
| LightGBM | 0.103 | 0.019 | 0.887 |
| CatBoost | 0.104 | 0.020 | 0.881 |

Table 3 displays the performance of six different models in predicting SoH of LiBs. Among these models, the LightGBM model demonstrated the highest performance with 0.103 MAE, 0.019 MSE, and 0.887 R-squared. Figure 2 provides a comparison of the predictions of different machine learning models with the actual results.

In Figure 2, the accuracy of predictions made by different machine learning models on discharge capacities of LiB is visually presented and compared. The graphical representation displays the actual discharge capacity values (Ah) on the horizontal axis and the predicted discharge capacity values (Ah) on the vertical axis. In an ideal scenario, all points would lie on the dashed black line (y = x) drawn on the graph. As shown from the graphs in Figure 2, the results of the AdaBoost model exhibit a wide distribution, indicating low accuracy in its predictions. On the other hand, the LightGBM model demonstrated a performance closer to the actual data, with most of its predictions being very close to the actual values. Similarly, CatBoost has also produced forecasts with high accuracy.

LightGBM is a boosting algorithm optimized for performance and efficiency. Innovative approaches such as histogram-based learning and leaf-based tree growth enable it to work quickly and effectively on large, high-dimensional datasets. These features make LightGBM an algorithm of choice, especially in large-scale machine learning projects. While tree growth occurs level-wise in other boosting algorithms (for example, XGBoost), leaf-wise growth is used in LightGBM. Leaf-wise growth deepens trees, starting with the leaf with the highest error rate. This way, more complex decision boundaries with better generalization capacity can be created compared to trees of the same depth.

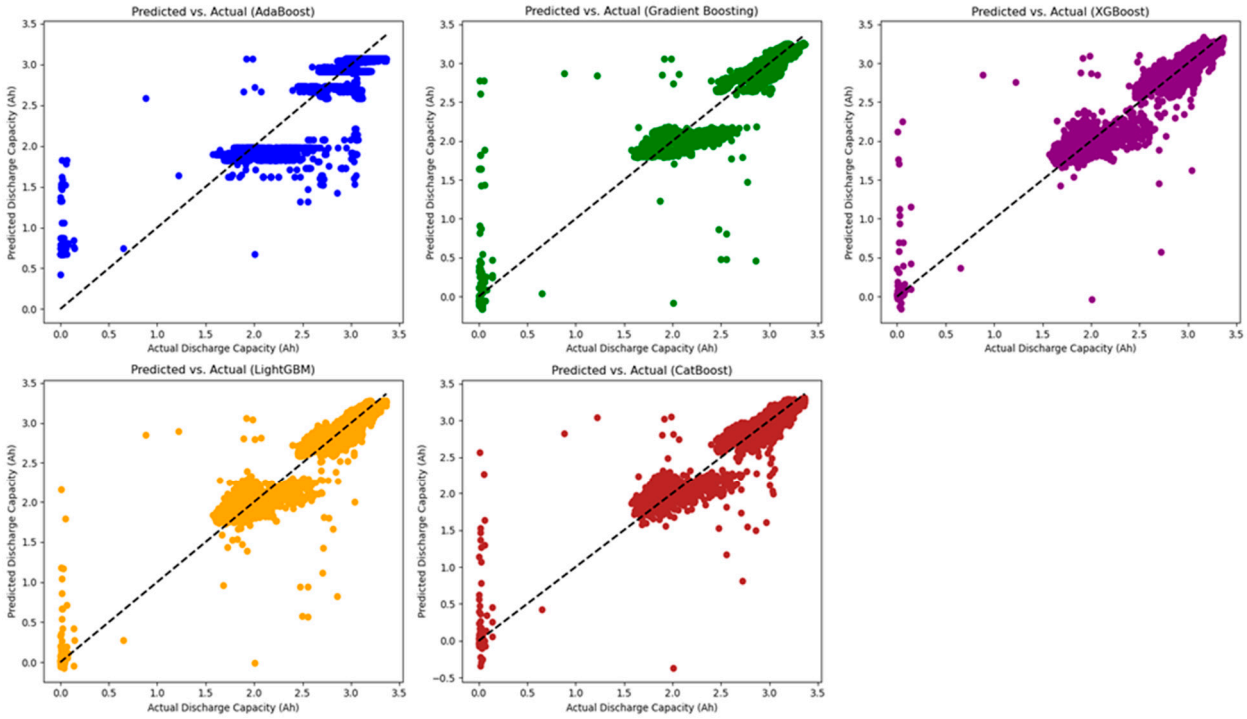



**Figure 2.** Predicted vs. actual values of models.

### *3.4. Ensemble Learning Model Results*

As part of the study, a voting-based regression model called VotingRegressor was developed. VotingRegressor combines predictions from different regression models to create an ensemble model. In this setup, the model comprises predictors from five regression algorithms: gradient boosting, XGBoost, AdaBoost, LightGBM, and CatBoost.

In the ensemble model, each submodel must be trained beforehand and be capable of making predictions. VotingRegressor aggregates the outputs from each of these predictors and generates the final prediction by averaging them [27]. This method may enhance overall prediction performance by leveraging the strengths of different models and balancing out the weaknesses of individual models. The ensemble model developed within the scope of the study demonstrated consistent performance among models, with 0.105 MAE, 0.020 MSE, and 0.884 R-squared. Figure 3 provides a comparison of the prediction results of the ensemble learning model with the actual results. The graph in Figure 3 illustrates that many predictions cluster around the line, indicating that the model might make predictions close to the actual values in most cases. However, there are noticeable deviations at low- and high-capacity values. It can be said that the model's accuracy may be lower in these specific ranges. Additionally, the clustering of predictions around 1.5 Ah and 2.5 Ah indicates that the model performs better in these value ranges.

Despite the ensemble model yielding good results, the LightGBM model demonstrated superior performance compared to the ensemble model. Among the reasons for this are LightGBM's ability to adapt particularly well to the characteristics of the dataset and its effectiveness in handling complexities in the dataset more efficiently with its leaf-wise growth method. On the other hand, in the ensemble model, combining predictions from different models with a simple average may not always adequately reflect the specificities and strengths of individual models. Considering these factors, the superior performance of LightGBM compared to other models highlights the importance of proper data preprocessing and model parameterization.

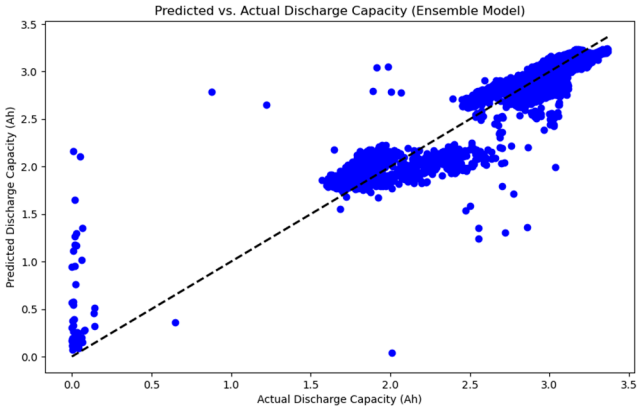



**Figure 3.** The actual and predicted ensemble model.

### *3.5. Explainability Analysis of AI Models*

In this study, the LightGBM model, which exhibited the best performance among the models developed for predicting LiBs' discharge capacities, was subjected to explainability analysis using SHAP [28]. SHAP allows for the quantitative evaluation of the effects of features contributing to a model's predictions through Shapley values derived from game theory. This analysis was conducted to unveil the strongest and weakest impacts on predictions, thus providing a deeper understanding of how the model operated and making the dynamics underlying predictions comprehensible. This analysis visually and understandably presents which features the model assigns more weight to and the effects of these features on prediction outcomes [29]. In Figure 4, the impact of four main features used in LiB discharge capacity predictions by the LightGBM model is explained using the SHAP model. After all the dataset features and predictions are made, the decisions made by the LightGBM algorithm, which gives the most successful results and performance, are interpreted with the SHAP technique. When interpreting with the SHAP technique, the impact ratios of the input feature "Temperature", "Current (A)", "Voltage (V)", and "Cycle Index", which are feature names, in determining the output feature "Discharge Capacity" are evaluated.

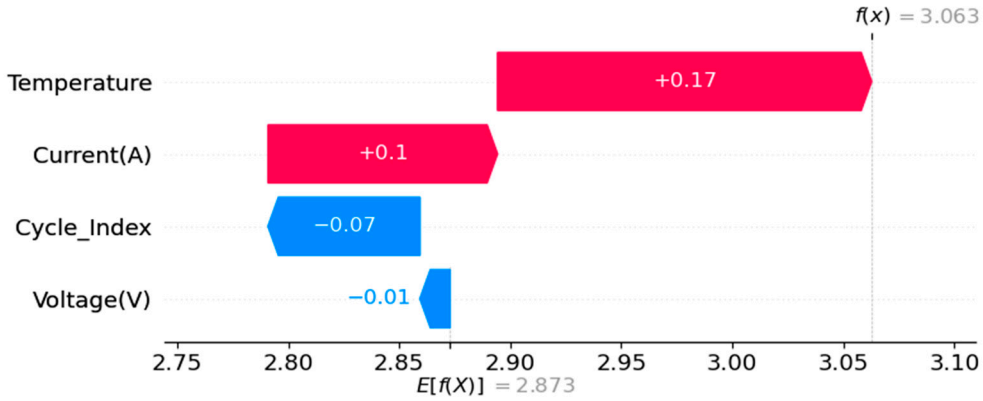



**Figure 4.** SHAP value waterfall for LGBM model.

In Figure 4, the "Temperature" and "Current (A)" features are represented by positive SHAP values, indicating that these features increase the model predictions. Particularly, "Temperature" has the highest positive impact (+0.17), suggesting that an increase in temperature significantly enhances the prediction of discharge capacity. "Current (A)" has a lower positive effect (+0.1) on the predictions. On the other hand, the "Cycle_Index" and "Voltage(V)" features are represented by negative SHAP values, indicating that these two features decrease the model predictions. "Cycle_Index" (−0.07) is the most significant factor reducing model predictions, while the effect of "Voltage (V)" is relatively smaller

(−0.01). With E[f(X)] value determined as 2.873, adding SHAP values results in the final model prediction f(X) reaching 3.063. This clearly demonstrates how the total effect of features contributes to changes in model predictions. Figure 5 provides the graph of average absolute SHAP values.

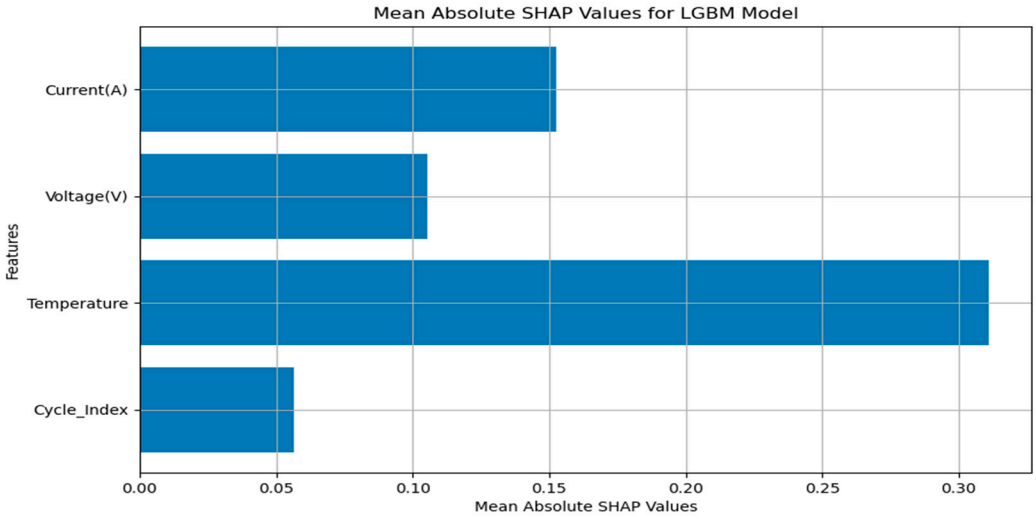



**Figure 5.** LightGBM model parameters explainability with SHAP model values.

According to the data presented in Figure 5, the "Temperature" feature has the highest impact on the predictions of "Discharge Capacity (Ah)", and SHAP values generally range between −1.5 and 0.5. The SHAP values for the "Current" feature range between −0.5 and 0.3, with most values lying in the positive region. While the SHAP values for the "Voltage" feature vary between −0.5 and 0.1, the SHAP values for the "Cycle_Index" feature mostly range between −0.3 and −0.1, predominantly exhibiting negative effects.

Figure 6 depicts the contribution of features step by step for a given forecast sample. Numerical values are as follows: The feature increases the model output moving to the right on the axis and decreases it moving to the left. The absolute magnitude of the values indicates the strength of the feature's effect. Colors represent high or low values of the features. Red indicates high values while blue indicates low values.

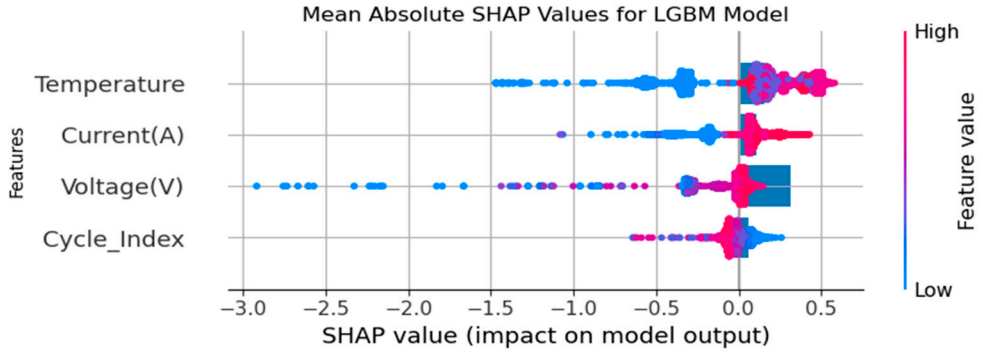



**Figure 6.** Mean absolute SHAP values for the LightGBM model.

As shown in the LightGBM model described in Figures 4 and 6, temperature has a significant impact on discharge capacity. Low-temperature values reduce discharge capacity, and higher values positively affect discharge capacity. The LightGBM model emerges as the best-performing model among the five models, with the lowest MSE and MAE and the highest R-squared value, indicating the strongest correlation between predictions and actual values. Gradient boosting and XGBoost are at similar performance levels, ranking just behind LightGBM [30]. The performance of the ensemble model is

competitive, indicating that combining these models may lead to an overall well-performing model. SHAP measures how much influence a model has on each instance and the change this interaction causes. These values indicate the importance levels of the features that contribute to the model's predictions. Each value expresses how significant the change in estimates are due to a particular feature.

**Cycle_Index** shows the importance of the effect of the cycle index on the predictions. The cycle index acts as a temporary timestamp throughout a cycle. Generally, it represents the number of times a cycle is repeated or the amount of time elapsed throughout a cycle. A high SHAP value indicates that the cycle index plays an important role in the predictions.

**Temperature** might often be an important factor in a machine learning model in several applications. For instance, temperature plays a vital role in energy efficiency, chemical reaction rates, and many other physical processes. A high SHAP value indicates that temperature significantly impacts the forecast.

**Voltage (V)** represents the energy level in an electrical system. Voltage is an essential factor that affects a system's energy efficiency and performance. A high SHAP value indicates that voltage is essential to the predictions.

**Current (A)** presents the energy flow in an electrical unit. The height of the current indicates how much energy a circuit consumes or produces. A high SHAP value suggests that the current significantly impacts the predictions. Interpreting these values is essential to understand which features of the model are most important in the projections. For instance, features with high SHAP values may be assigned more weight in the model's predictions, while features with low values may be less influential. This information might be used to determine which features need to be worked on to improve the model's performance.

## 4. Discussion

Recent research has made significant progress in predicting LiB performance management in EVs using ML techniques.

- Chandran [31] and Dineva [32] both demonstrated the effectiveness of ML algorithms, such as artificial neural networks and ensemble-boosted tree models, in estimating the state of charge and terminal voltage of LiB, respectively.
- Poh [33] provided a comprehensive review of ML methods for state estimation in EV applications, highlighting the potential of these techniques in addressing the challenges of time-varying and nonlinear battery traits.
- Astaneh [34] further advanced the field by presenting a methodology for simulating the performance of LiB packs in EVs, achieving high accuracy in voltage prediction and battery pack temperature estimation. These studies collectively underscore the promising role of ML in enhancing the performance management of LiBs in EVs. In the survey, boosting worked iteratively, adding a new model at each iteration. The latest model tries to correct the mistakes of the previous model. This process is accomplished by stepping along the gradient of the loss function. This is also true for a specific application such as boost gradients, but, in general, the boosting technique does not include regularization terms that control the complexity of the model and is generally aimed at reducing the model's errors.

In addition, various studies aiming to predict the state of LiBs using different methodologies were summarized. In a study in [31], the methodology involved using six machine learning algorithms (ANN, SVM, LR, GPR, EBa, EBo) to estimate LiBs' state of charge in EVs, with a comparison of their performance indices. Participants received interventions related to studying battery charging and discharging characteristics in ML models, such as analyzing the signal's energy, temperature variations, voltage curves, and specific characteristics of battery behavior. The performance of six ML algorithms and MSE, RMSE, and MAE metrics were used. ANN and GPR were the best methods for state-of-charge estimation, outperforming other machine learning algorithms, with an 85% MAE. These models accurately predicted the battery state of charge, assisting in battery selection for specific applications.

In the studies:

- The methodology involved conducting WLTP discharge tests on an NMC cell and using historical data to build boosted tree models for voltage prediction. Measurements of WLTP discharge tests were taken at different temperatures on an NMC cell, along with terminal voltage, discharge rate, and temperatures at four points in [32], and prediction of cell voltage as outcome sequences into the future (battery SoC prediction). The study conducted WLTP discharge tests on an NMC cell and found that a direct multistep-ahead forecasting strategy with standard machine learning techniques is efficient and applicable in battery SoC prediction.
- The methodology involved proposing a multioutput convolved Gaussian-process (MCGP) model for capacity estimation of LiB cells in EVs and validating it with experimental data from specific battery cells. Capacity estimation of LiB cells, SoC estimation accuracy, and a battery cell's remaining useful life (RUL) was conducted in [35]. The study introduced a multioutput convolved Gaussian-process (MCGP) model for accurately estimating the capacity of LiB cells in EVs. The proposed model can improve the accuracy of SoC estimation and serve as a precise tool for predicting the RUL of a battery cell.
- A battery aging and temperature-aware predictive energy management strategy for parallel hybrid electric vehicles was used to model predictive control (MPC) and optimize with Pontryagin's minimization principle (PMP) [36]. These are battery aging and temperature-aware predictive energy management strategies for parallel hybrid electric vehicles. The study developed a battery aging and temperature-aware predictive energy management strategy, with the PMP method superior to dynamic programming, and showed that a battery temperature-aware strategy can reduce total energy consumption.
- Artificial intelligence techniques for battery SoC estimation, utilizing the Panasonic 18650PF dataset, conducting preprocessing and feature extraction, and evaluating model performance with various metrics were given in [31]. The participants did not receive any interventions as this study was focused on the application of artificial intelligence techniques for estimating battery state of charge (SoC) in LiBs.
- Estimation of battery state of charge (SoC) using artificial intelligence techniques is a focus for the ANN model's performance ($R^2$ value of 0.9925). The ANN model outperformed other AI techniques in estimating battery state of charge (SoC) with a high $R^2$ value of 0.9925, showcasing its ability to capture underlying patterns for precise estimates.

Many studies have focused on improving prediction accuracy on LiB capacities; however, examples of the use of XAI methodologies in this field are rarely encountered. Therefore, the current study conducted using SHAP contributes significantly to the literature. SHAP values are utilized to understand which features, and to what extent, influence a model's predictions. Particularly, variables affecting the performance of lithium-ion batteries and their contributions to predictions have been analyzed in detail. The present study enabled a better understanding of the models used to predict the state and capacity of LiBs, thereby increasing the reliability of these predictions.

The findings discussed in our experimental results chapter are compared with existing machine learning methodologies used in the prediction of the state and capacity of LiBs. This study makes a significant contribution in providing strategic insights for the optimization and development of battery management systems, particularly through the use of SHAP-based XAI methodologies. Our findings show that SHAP values play a critical role in understanding which features influence model predictions and to what extent. This approach extends its applications by deepening the body of knowledge in battery technologies. In particular, the ability of ANN and GPR models to predict SoC with higher accuracy than other machine learning algorithms highlights the capacity of these techniques to capture battery behavior. As a result, this study improves the relia-

bility of models used to predict the performance of LiBs and enables the development of management systems that can make battery use more efficient.

Additionally, through this approach, strategic information might be provided for the optimization and development of battery management systems, allowing for more efficient battery usage. Therefore, the SHAP-based XAI study deepens the knowledge of battery technologies and expands their applications.

## 5. Conclusions and Future Trends

This study specifically evaluated the performance of various machine learning models in predicting the discharge capacity of LiBs by focusing on AdaBoost, gradient boosting, XGBoost, LightGBM, CatBoost, and an ensemble learning model. The findings demonstrated that LightGBM outperformed other models with the lowest MAE and MSE values and the highest R-squared value, indicating a strong correlation between predicted and actual values. Both gradient boosting and XGBoost exhibited similar performance levels but slightly lagged behind LightGBM.

The ensemble learning model showcased competitive performance by emphasizing the effectiveness of combining multiple models to obtain a robust prediction framework. The utilization of SHAP values for explaining the most successful model, LightGBM, highlighted the significance of various features such as temperature, cycle index, voltage, and current in influencing predictions. Particularly, temperature emerged as a critical determinant. XAI analyses conducted revealed that high temperatures negatively affect discharge capacity, aligning with physical expectations regarding battery performance.

The results of this study demonstrate that machine learning models are theoretically effective for predicting the condition and capacity of LiBs. However, the applicability of these models in real-world scenarios involves practical challenges beyond the adequacy of the algorithm, such as model complexity, computational requirements, and real-time data processing capacity, which may limit the integration of the model into industrial applications. In battery management systems for electric vehicles, machine learning models can optimize battery life and maximize range by feeding in real-time data such as traffic conditions, driving style, and weather. In smart grids that incorporate renewable energy, these models predict energy demand, improving energy efficiency and reducing operational costs. In factories, machine learning models can predict failures by determining maintenance schedules for battery-powered equipment. Such real-time applications demonstrate the practical value of the model and its effective integration into industrial applications. Implementing more sophisticated BMS algorithms that incorporate machine learning and AI will better monitor and manage battery health, ensuring optimal charging and discharging cycles to prolong battery life.

Future studies may explore more complex ensemble methods integrating newer or more diverse machine learning models to further enhance prediction accuracy. Integrating these models into real-time monitoring systems for electric vehicles could provide dynamic predictions adaptable to changing conditions, thus optimizing battery usage and extending battery life. A deeper analysis regarding additional features that may affect battery performance, such as environmental factors or material degradation, will provide more comprehensive insights. Extending prediction-based models to other battery and storage technologies could broaden their applicability and impact across different energy systems.

Continued focus on XAI models, particularly in sensitive applications like electric vehicles where understanding model predictions may significantly influence design and operational decisions, will be crucial. As machine learning models become more prevalent in critical applications, developing standards and regulatory frameworks to ensure their reliability and security will become imperative. The inclusion of XAI not only ensures transparency but also provides actionable insights that can shape the future of battery technology development. As demand for sustainable energy solutions continues to rise, developing more robust prediction models and explaining them with XAI methods

will be crucial in elucidating the physical properties of LiBs and the relationship among prediction algorithms.

Once Li-ion batteries reach the end of their life due to performance decline, it is essential to address their disposal and recycling responsibly. Additionally, refurbishing spent batteries for secondary use in less demanding applications can extend their lifecycle and contribute to a circular economy. Developing efficient and sustainable processes for handling end-of-life lithium-ion batteries is crucial for minimizing environmental impact and supporting the growing demand for battery materials in an eco-friendly manner.

**Author Contributions:** Methodology, S.O. and A.A.; Software, S.O. and B.E.; Validation, B.E.; Investigation, E.B.; Resources, A.A.; Writing—original draft, B.E.; Writing—review & editing, Ş.S. and E.B.; Supervision, S.O., Ş.S. and E.B.; Project administration, S.O.; Funding acquisition, Ş.S. and A.A. All authors have read and agreed to the published version of the manuscript.

**Funding:** This paper was supported by the European Union's Horizon Europe research and innovation programme under "Next Generation of Multifunctional, Modular and Scalable Solid State Batteries System" (EXTENDED), Grant Agreement No. 101102278.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The original contributions presented in the study are included in the article, further inquiries can be directed to the corresponding author.

**Acknowledgments:** This research was conducted collaboratively by the MOBILERS team at Sivas Cumhuriyet University and the Battery Research Group at Sivas University of Science and Technology (SBTU).

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Oyucu, S.; Doğan, F.; Aksöz, A.; Biçer, E. Comparative Analysis of Commonly Used Machine Learning Approaches for Li-Ion Battery Performance Prediction and Management in Electric Vehicles. *Appl. Sci.* **2024**, *14*, 2306. [CrossRef]
2. Oyucu, S.; Dümen, S.; Duru, İ.; Aksöz, A.; Biçer, E. Discharge Capacity Estimation for Li-Ion Batteries: A Comparative Study. *Symmetry* **2024**, *16*, 436. [CrossRef]
3. Klass, V.; Behm, M.; Lindbergh, G. Capturing lithium-ion battery dynamics with support vector machine-based battery model. *J. Power Sources* **2015**, *298*, 92–101. [CrossRef]
4. Laadjal, K.; Cardoso, A.J.M. Estimation of Lithium-Ion Batteries State-Condition in Electric Vehicle Applications: Issues and State of the Art. *Electronics* **2021**, *10*, 1588. [CrossRef]
5. Li, Y.; Luo, L.; Zhang, C.; Liu, H. State of Health Assessment for Lithium-Ion Batteries Using Incremental Energy Analysis and Bidirectional Long Short-Term Memory. *World Electr. Veh. J.* **2023**, *14*, 188. [CrossRef]
6. Sarmah, S.B.; Kalita, P.; Garg, A.; Niu, X.-D.; Zhang, X.-W.; Peng, X.; Bhattacharjee, D. A Review of State of Health Estimation of Energy Storage Systems: Challenges and Possible Solutions for Futuristic Applications of Li-Ion Battery Packs in Electric Vehicles. *J. Electrochem. Energy Convers. Storage* **2019**, *16*, 040801. [CrossRef]
7. Harper, G.D.J.; Kendrick, E.; Anderson, P.A.; Mrozik, W.; Christensen, P.; Lambert, S.; Greenwood, D.; Das, P.K.; Ahmeid, M.; Milojevic, Z.; et al. Roadmap for a sustainable circular economy in lithium-ion and future battery technologies. *J. Phys. Energy* **2023**, *5*, 021501. [CrossRef]
8. Lu, L.; Han, X.; Li, J.; Hua, J.; Ouyang, M. A review on the key issues for lithium-ion battery management in electric vehicles. *J. Power Sources* **2013**, *226*, 272–288. [CrossRef]
9. Madani, S.S.; Ziebert, C.; Marzband, M. Thermal Characteristics and Safety Aspects of Lithium-Ion Batteries: An In-Depth Review. *Symmetry* **2023**, *15*, 1925. [CrossRef]
10. Chen, T.; Jin, Y.; Lv, H.; Yang, A.; Liu, M.; Chen, B.; Xie, Y.; Chen, Q. Applications of Lithium-Ion Batteries in Grid-Scale Energy Storage Systems. *Trans. Tianjin Univ.* **2020**, *26*, 208–217. [CrossRef]
11. Chen, J.C.; Chen, T.-L.; Liu, W.-J.; Cheng, C.; Li, M.-G. Combining empirical mode decomposition and deep recurrent neural networks for predictive maintenance of lithium-ion battery. *Adv. Eng. Inform.* **2021**, *50*, 101405. [CrossRef]
12. Lipu, M.S.H.; Miah, S.; Jamal, T.; Rahman, T.; Ansari, S.; Rahman, S.; Ashique, R.H.; Shihavuddin, A.S.M.; Shakib, M.N. Artificial Intelligence Approaches for Advanced Battery Management System in Electric Vehicle Applications: A Statistical Analysis towards Future Research Opportunities. *Vehicles* **2024**, *6*, 22–70. [CrossRef]
13. He, Q.; Yu, B.; Li, Z.; Zhao, Y. Density Functional Theory for Battery Materials. *Energy Environ. Mater.* **2019**, *2*, 264–279. [CrossRef]

14. Zhang, C.; Cheng, J.; Chen, Y.; Chan, M.K.Y.; Cai, Q.; Carvalho, R.P.; Marchiori, C.F.N.; Brandell, D.; Araujo, C.M.; Chen, M.; et al. 2023 roadmap on molecular modelling of electrochemical energy materials. *J. Phys. Energy* **2023**, *5*, 041501. [CrossRef]
15. Machlev, R.; Heistrene, L.; Perl, M.; Levy, K.; Belikov, J.; Mannor, S.; Levron, Y. Explainable Artificial Intelligence (XAI) techniques for energy and power systems: Review, challenges and opportunities. *Energy AI* **2022**, *9*, 100169. [CrossRef]
16. Alam, M.W.; BaQais, A.; Rahman, M.M.; Aamir, M.; Abuzir, A.; Mushtaq, S.; Amin, M.N.; Khan, M.S. Investigation on In Situ Carbon-Coated $ZnFe_2O_4$ as Advanced Anode Material for Li-Ion Batteries. *Gels* **2022**, *8*, 305. [CrossRef] [PubMed]
17. Jayachitra, J.; Joshua, J.R.; Balamurugan, A.; Sivakumar, N.; Sharmila, V.; Shanavas, S.; Abu Haija, M.; Alam, M.W.; BaQais, A. High electrode performance of hydrothermally developed activated C coated O3–$NaFeO_2$ electrode for Na-ion batteries applications. *Ceram. Int.* **2023**, *49*, 48–56. [CrossRef]
18. Diao, W.; Saxena, S.; Pecht, M. Accelerated cycle life testing and capacity degradation modeling of $LiCoO_2$-graphite cells. *J. Power Sources* **2019**, *435*, 226830. [CrossRef]
19. Bentéjac, C.; Csörgő, A.; Martínez-Muñoz, G. A Comparative Analysis of Gradient Boosting Algorithms. *Arti. Intelli. Rev.* **2021**, *54*, 1937–1967. [CrossRef]
20. Cai, B.; Li, M.; Yang, H.; Wang, C.; Chen, Y. State of Charge Estimation of Lithium-Ion Battery Based on Back Propagation Neural Network and AdaBoost Algorithm. *Energies* **2023**, *16*, 7824. [CrossRef]
21. Chen, T.; He, T.; Benesty, M. XGBoost: eXtreme Gradient Boosting. R Package Version 0.4-2. 2015. Available online: https://rdocumentation.org/packages/xgboost/versions/0.4-2 (accessed on 12 May 2024).
22. Al Daoud, E. Comparison between XGBoost, LightGBM and CatBoost Using a Home Credit Dataset. *Int. J. Comput. Inf. Eng.* **2019**, *13*, 6–10.
23. Nordin, N.; Zainol, Z.; Noor, M.H.M.; Chan, L.F. An explainable predictive model for suicide attempt risk using an ensemble learning and Shapley Additive Explanations (SHAP) approach. *Asian J. Psychiatry* **2023**, *79*, 103316. [CrossRef] [PubMed]
24. De Myttenaere, A.; Golden, B.; Le Grand, B.; Rossi, F. Mean Absolute Percentage Error for regression models. *Neurocomputing* **2016**, *192*, 38–48. [CrossRef]
25. Yin, J.; Li, N. Ensemble learning models with a Bayesian optimization algorithm for mineral prospectivity mapping. *Ore Geol. Rev.* **2022**, *145*, 104916. [CrossRef]
26. Hasan, R. Revitalizing the Electric Grid: A Machine Learning Paradigm for Ensuring Stability in the U.S.A. *J. Comput. Sci. Technol. Stud.* **2024**, *6*, 141–154. [CrossRef]
27. Phyo, P.P.; Byun, Y.C.; Park, N. Short-Term Energy Forecasting Using Machine-Learning-Based Ensemble Voting Regression. *Symmetry* **2022**, *14*, 160. [CrossRef]
28. Aldrees, A.; Khan, M.; Taha, A.T.B.; Ali, M. Evaluation of water quality indexes with novel machine learning and SHapley Additive ExPlanation (SHAP) approaches. *J. Water Process. Eng.* **2024**, *58*, 104789. [CrossRef]
29. Ma, K. Integrated hybrid modeling and SHAP (SHapley Additive exPlanations) to predict and explain the adsorption properties of thermoplastic polyurethane (TPU) porous materials. *RSC Adv.* **2024**, *14*, 10348–10357. [CrossRef] [PubMed]
30. Mohammed, A.; Kora, R. A comprehensive review on ensemble deep learning: Opportunities and challenges. *J. King Saud Univ.-Comput. Inf. Sci.* **2023**, *35*, 757–774. [CrossRef]
31. Chandran, V.; Patil, C.K.; Karthick, A.; Ganeshaperumal, D.; Rahim, R.; Ghosh, A. State of Charge Estimation of Lithium-Ion Battery for Electric Vehicles Using Machine Learning Algorithms. *World Electr. Veh. J.* **2021**, *12*, 38. [CrossRef]
32. Dineva, A.; Kocsis, S.S.; Vajda, I. Data-driven terminal voltage prediction of li-ion batteries under dynamic loads. In Proceedings of the 2020 21st International Symposium on Electrical Apparatus & Technologies (SIELA), Bourgas, Bulgaria, 3–6 June 2020; pp. 1–5. [CrossRef]
33. Poh, W.Q.T.; Xu, Y.; Tan, R.T.P. A Review of Machine Learning Applications for Li-Ion Battery State Estimation in Electric Vehicles. In Proceedings of the 2022 IEEE PES Innovative Smart Grid Technologies-Asia (ISGT Asia), Singapore, 1–5 November 2022. [CrossRef]
34. Astaneh, M.; Andric, J.; Löfdahl, L.; Maggiolo, D.; Stopp, P.; Moghaddam, M.; Chapuis, M.; Ström, H. Calibration Optimization Methodology for Lithium-Ion Battery Pack Model for Electric Vehicles in Mining Applications. *Energies* **2020**, *13*, 3532. [CrossRef]
35. Chehade, A.A.; Hussein, A.A. A Multi-Output Convolved Gaussian Process Model for Capacity Estimation of Electric Vehicle Li-ion Battery Cells. In Proceedings of the 2019 IEEE Transportation Electrification Conference and Expo (ITEC), Detroit, MI, USA, 19–21 June 2019.
36. Du, R.; Hu, X.; Xie, S.; Hu, L.; Zhang, Z.; Lin, X. Battery aging-and temperature-aware predictive energy management for hybrid electric vehicles. *J. Power Sources* **2020**, *473*, 228568. [CrossRef]


**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.