Research Article

Metaheuristic-based machine learning models for accurate battery remaining useful life prediction in renewable energy storage system

Shravani GANDHI*📄, **Ashlesha CHOPANE**📄, **Anushka PARKHI**📄, **Sakshi DESHMUKH**📄, **Smita KULKARNI**📄

¹Department of Electronics and Telecommunication Engineering, MIT Academy of Engineering, Pune, 412105, India

ARTICLE INFO

ABSTRACT

This work focuses on properly estimating lithium-ion battery Remaining Useful Life in renewable energy storage and electric vehicle applications. This is important since batteries are essential to sustainable energy systems and electric car demand is expanding. Accurately predicting Remaining Useful Life (RUL) is essential for maximizing battery safety and efficiency. Furthermore, these predictions play a critical role in ensuring that energy storage systems remain both economically viable and environmentally sustainable. The study uses a dataset from the Hawaii Natural Energy Institute, containing data on battery lifecycle performance metrics. The machine learning models used include random forests, decision trees, linear regression, K-nearest neighbors (KNN), support vector machines (SVM), AdaBoost, and extreme gradient boosting. To maximize predictive precision, every model undergoes a process of hyperparameter tuning, data preprocessing, and the evaluation of feature importance. Comparative measures like RMSE, MAE, and R² assess model predictive capability. The research shows that Random Forest and XGBoost models excel, with Random Forest earning an RMSE of 23.26 and an R2 score of 99.68. XGBoost demonstrated strong performance, including a decreased RMSE of 20.64 and an R² score of 99.61. Research indicates that ensemble learning techniques, such as XGBoost and Random Forest, provide highly reliable forecasts for Remaining Useful Life, facilitating improved maintenance and oversight of battery systems. The novelty of this study lies in its comprehensive comparative analysis of multiple machine-learning models optimized with metaheuristic techniques for battery Remaining Useful Life predictions. In contrast to earlier research that typically utilized single-model methods, this study investigates an ensemble of machine learning models with a focus on optimization and feature importance to enhance predictive precision. The results show that feature selection, data processing, and hyperparameter optimization have improved the model's performance. This supports machine learning's role in preventive maintenance, battery longevity, and sustainable energy. This study introduces a robust, data-centric framework that combines various techniques to enhance remaining useful life predictions, serving as a beneficial resource for practical battery management systems.

Cite this article as: Gandhi S, Chopane A, Parkhi A, Deshmukh S, Kulkarni S. Metaheuristic-based machine learning models for accurate battery remaining useful life prediction in renewable energy storage system. Sigma J Eng Nat Sci 2026;44(1):528–544.

***Corresponding author.**
*E-mail address: shravani.gandhi@mitaoe.ac.in

INTRODUCTION

As the demand for sustainable energy solutions grows, accurately predicting the remaining useful life (RUL) of lithium-ion batteries has become crucial, especially for electric vehicles (EVs) and renewable energy storage. While existing literature presents various machine-learning approaches for RUL prediction, many studies rely on traditional algorithms that often lack the precision needed for real-world applications. For example, support vector machines (SVM) and basic regression techniques frequently overlook the complexities of battery degradation and diverse operating conditions. Advancing sustainable development requires reducing greenhouse gas emissions, thus promoting electric vehicles and innovative energy storage solutions. As consumer electronics increasingly depend on battery technology and renewable energy expands, the energy storage sector has become pivotal. Metaheuristic solutions enhance process efficiency, sustainability, and energy consumption.

Lithium-ion batteries (LIBs), known for their high energy density and durability, are the primary choice for powering electric vehicles. Notwithstanding recent progress, certain voids persist in the current research. Numerous predictive frameworks fail to extensively utilize contemporary machine learning methods, especially metaheuristic optimization approaches, which have the potential to boost accuracy. Additionally, most studies inadequately address the real-time implementation of predictive models, limiting their applicability in dynamic environments where timely decision-making is essential. Forecasting RUL is essential for evaluating battery safety, improving performance, and maximizing the efficiency and sustainability of energy storage systems. Efficient battery management supports sustainable development goals, such as affordable clean energy and climate action [1], and promotes renewable energy adoption and innovation. RUL estimations, represented through multiple indicators, offer early warnings of failure that are critical for system prognostics and health management. Methods for modelling the degradation of complex systems generally encompass both physics-based and data-driven strategies. A probabilistic adaptive estimator assesses the state of energy (SOE) and state of charge (SOC) using neural networks [2].

One-dimensional models analyze battery decline through SEI generation kinetics and solvent transport, while a refined, semi-empirical aging model incorporates factors like internal resistance and capacity degradation [3]. A comparison study using RMS and MSE as metrics [4] found that XGBoost and Gradient Boost were the best at predicting Remaining Useful Life. Another study using LSTM techniques and temperature, current, and voltage as metrics [5] also found high accuracy. A KNN regression approach, optimized with particle swarm optimization (PSO), offers a simple yet accurate model for capacity assessment [6]. Artificial intelligence is fundamental to enhancing battery behavior, as explored in a bibliometric review of machine learning techniques for battery forecasting [7]. Additionally, previous research has confirmed the effectiveness of models such as PNN, LSTM, and SVR using the NASA battery dataset to refine practical performance [8]. Moreover, [9] emphasizes the significance of SOH estimates in electric vehicles and highlights advancements in filtering techniques and co-estimation methods. Investigators examine the degradation variables of lead-acid batteries to refine battery control and reliability, while a unique ensemble learning technique exhibits superior RUL prediction precision relative to conventional methods [10,11].

Researchers have also proposed integrated modelling approaches to improve the prediction accuracy of lithium-ion battery lifespan [12]. Despite challenges related to computational costs and data deficits, researchers have noted the effectiveness of neural networks, particularly RNNs and LSTMs, in managing complex data [13]. Several studies have compared different machine learning algorithms, including XGBoost and Gradient Boost, using various performance metrics [14]. Furthermore, studies have underscored the importance of parameters like temperature and voltage in LSTM-based predictions [15]. The literature highlights the importance of accurate Remaining Useful Life predictions in enhancing battery management systems for electric vehicles [16-21]. To manage uncertainties and fluctuating operational states in batteries, investigators have utilized methodologies including unscented Kalman filters (UKF) and hidden Markov models (HMM) [22-24]. Furthermore, studies focusing on SOC estimation and battery parameter estimation have demonstrated improved accuracy through advanced techniques, supporting effective battery management [23,25].

Recent results indicate that combining model-based methodologies with machine learning can improve RUL estimations and refine battery consumption, furthering sustainability objectives [26]. The study on heart failure utilizes machine learning techniques to predict patient survival probabilities, underscoring the significance of predictive analytics in healthcare. Similarly, this research investigates metaheuristic-based models for accurate prediction of the RUL of batteries, aiming to improve operational efficiency in renewable energy systems [27]. Research regarding the estimation of beam structure natural frequencies highlights the successful implementation of machine learning techniques within engineering fields. This parallels the research on metaheuristic-based models for accurately predicting the Remaining Useful Life (RUL) of batteries in renewable energy systems [28]. The use of machine learning to generate precise forecasts of patient counts and mortality during the COVID-19 pandemic mirrors the methodologies applied in research on battery Remaining Useful Life estimation [29,30]. This study aims to bridge existing gaps by utilizing advanced machine learning models, specifically Random Forest and XGBoost, enhanced with metaheuristic optimization techniques. By isolating essential

performance markers such as discharge intervals, voltage dynamics, and cycle indices, this study demonstrates a rigorous approach to battery data extraction for predictive modeling. Emphasizing the importance of preprocessing steps like data standardization, feature selection, and handling missing values; these techniques ensure accurate RUL predictions by refining input data and enhancing model reliability. This investigation provides an in-depth analysis of error metrics and feature significance, demonstrating that these models achieve higher precision in Remaining Useful Life estimation compared to conventional techniques. The study presents a robust dataset from the Hawaii Natural Energy Institute and presents a systematic approach that integrates effective data preparation and hyperparameter tuning. By optimising RUL forecasts, this study contributes to proactive maintenance strategies, waste reduction, and safety assurance, thereby promoting sustainable energy practices. Furthermore, the findings extend beyond the battery sector, impacting industries such as manufacturing, healthcare, and infrastructure monitoring, where predictive maintenance can drive significant resource efficiencies. Additionally, this work promotes green energy initiatives by delivering consistent and precise estimations of battery durability, which improves the management of materials and lengthens the functional period of storage units.

The proposed models are key to enhancing renewable energy storage system efficiency and reliability, paving the way for broader applications in electric vehicles and renewable energy grids. The originality of this research stems from its comprehensive strategy of embedding sophisticated artificial intelligence within battery control frameworks, thereby bolstering the dependability and environmental viability of energy storage technologies. The study lays the groundwork for future research prioritizing real-time implementation capabilities and exploring diverse battery chemistries aligned with global climate goals and the urgent need for innovative energy solutions.

DATASET FOR REMAINING USEFUL LIFE PREDICTION

The Hawaii Natural Energy Institute conducted an experimental investigation to evaluate the performance of fourteen NMC-LCO 18650 lithium-ion batteries. This extensive dataset facilitated the extraction of features that characterized the voltage and current behavior of the batteries throughout their lifecycles. The development of predictive models to estimate the Remaining Useful Life of the batteries is dependent upon the presence of these features. The summarized dataset encapsulates the critical performance metrics for the 14 batteries, providing a comprehensive overview of their degradation patterns and operational efficiency.

Cycle Index: This column identifies each cycle in the dataset individually and acts as a chronological index. Every input represents a complete charge-discharge cycle,

enabling a sequential analysis of the battery's performance over time. To effectively track wear-and-tear patterns over time, the cycle index is indispensable. It allows researchers to establish a direct relationship between performance decay and the number of cycles completed. Depletion Duration (s): The 'Discharge Time' field tracks the interval in seconds for the battery to reach a fully drained state during every cycle, serving as a vital temporal indicator for evaluating the speed of energy exhaustion. Variations in discharge time across cycles can indicate changes in the battery's capacity and efficiency, thus serving as a key indicator of its health and performance.

Interval 3.6-3.4V (s): This field quantifies the period, in seconds, required for the battery's potential to decrease from 3.6 to 3.4 volts throughout the discharge phase. This detailed measurement provides insights into the voltage drop kinetics within a specific voltage window, revealing information about the internal resistance and overall discharge dynamics of the battery. Rapid voltage decrements may signal increased internal resistance or diminished capacity.

Max. Voltage Discharge. (V): This field captures the highest voltage level measured during every cycle's discharge stage, providing a baseline in volts. This maximum value represents the battery's electrical reaction while under load, which serves as a metric for its immediate power output capacity. Monitoring this parameter helps understand the voltage's behavior and identify any anomalies that could indicate potential issues.

Min. Voltage Charg. (V): This data point identifies the lowest voltage recorded throughout the recharging interval of each cycle. This value is critical for evaluating the minimum electrical threshold during the charging process, offering a perspective on the battery's recovery traits and its overall charging effectiveness. It serves as a critical indicator of the battery's ability to recuperate from deep discharge states.

Time at 4.15V (s): This column shows how long, in seconds, the battery sustains a voltage level of 4.15 volts during the charging process. The length of this constant voltage stage indicates the battery's saturation patterns and its proficiency in absorbing energy as it nears full capacity. This measurement is vital for assessing how stable and effective the established charging procedure remains.

Time constant current(s): The time constant current column likely refers to the time constant associated with the current in the battery system, measured in seconds. Often represented by τ (tau), this variable signifies the kinetic reaction of the battery's internal chemical mechanisms. Grasping the speed at which current adjusts following a shift in voltage is vital, as it dictates the behaviour of both charge and discharge curves.

Charging time (s): This column records the total duration, measured in seconds, required to complete the charging process for each cycle. It spans the complete interval from the start to the conclusion of the charging process,

thereby offering a comprehensive perspective on the battery's energy absorption rate and effectiveness. Shifts in this duration may indicate fluctuations in internal resistance and storage capacity, rendering it a crucial variable for evaluating the battery's overall lifespan.

RUL (Remaining Useful Life): The RUL column contains estimations of the remaining life of the battery, based on its current performance characteristics. This prognostic indicator estimates the remaining functional duration of the battery prior to hitting a set expiration threshold, such as a substantial decline in storage capacity or a rise in internal impedance. Remaining Useful Life is crucial for maintenance planning, performance forecasting, and overall management of the battery system.

This dataset, enriched with these detailed features, serves as a robust foundation for predictive analytics aimed at estimating battery lifespans and enhancing the understanding of battery degradation mechanisms.

Figure 1 displays several violin plots that illustrate the distribution of various performance metrics of 18650 lithium-ion batteries across different bound cycle indices. The initial graph, illustrating the duration of discharge in seconds, reveals a steady spread of time intervals across different cycle groupings, though the wide range of data points emphasizes the inherent inconsistency in these discharge

periods. The Decrement 3.6-3.4V (s) plot reveals that the time decrement between 3.6V and 3.4V remains fairly consistent across cycle bins but has some noticeable outliers, indicating fluctuations in performance. The Max. The voltage discharge (V) plot shows that the maximum discharge voltage slightly varies across cycle bins but generally remains within a specific range, suggesting stable maximum discharge voltages across cycles.

In contrast, the Min. The voltage charge (V) plot shows significant variability in the minimum charge voltage across bins, pointing to fluctuations in the minimum voltage required for charging. The graph for 'Time at 4.15V (s)' shows substantial fluctuation and numerous extreme values, suggesting that the periods spent at this specific voltage threshold are highly irregular. The time constant current (s) plot reveals that most data points are concentrated around lower values, with a few higher outliers, suggesting that the time spent at constant current is generally low but can vary significantly. In a similar fashion, the 'Charging Time (s)' graph demonstrates that most data points cluster within the lower range, though the presence of prominent anomalies underscores the fluctuations in recharge periods.

The dataset includes key features such as discharge time, voltage variations, and cycle index, which serve as input for machine learning models to predict Remaining Useful Life.
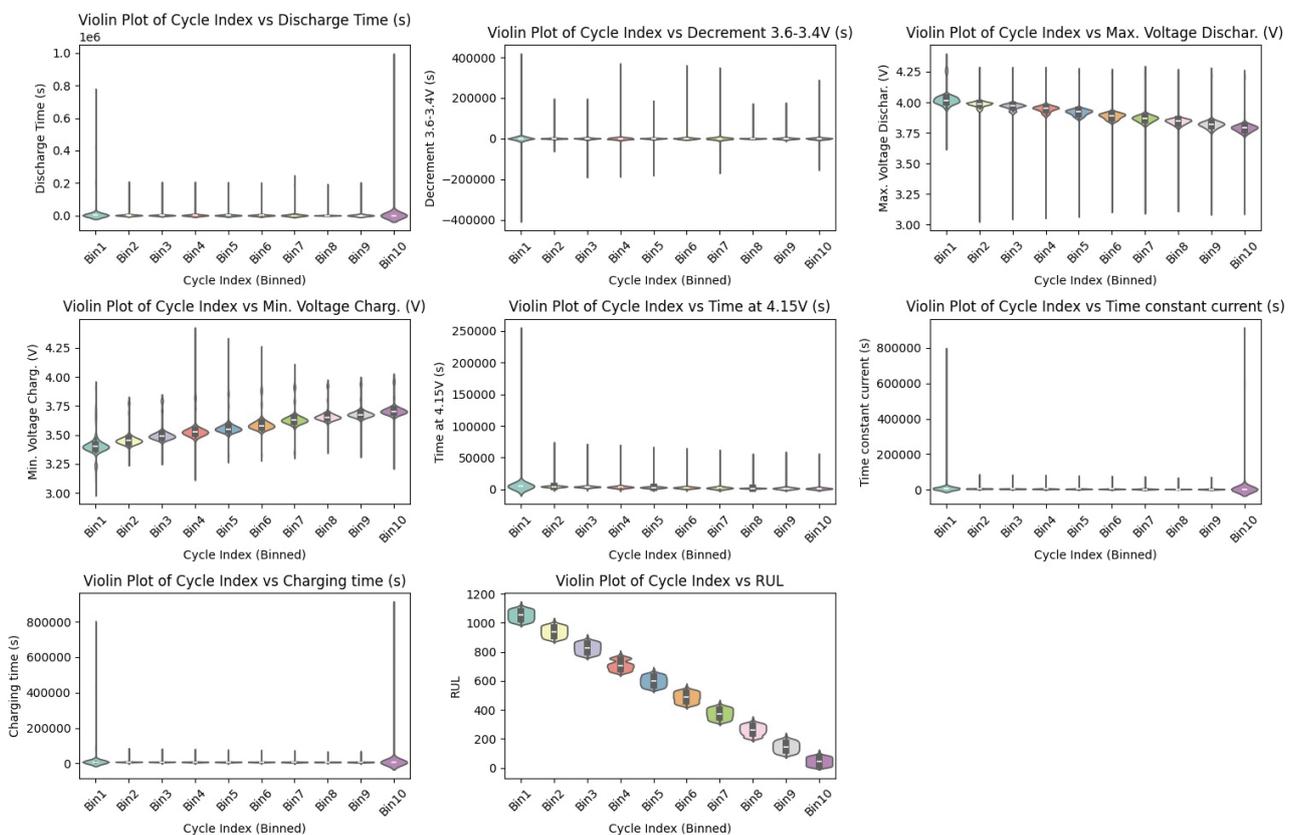

Figure 1. Violin plots of battery performance metrics across cycle index bins.

The resulting data undergoes normalization and organization to ensure seamless integration with algorithms such as Random Forest and XGBoost, enabling precise estimations of Remaining Useful Life across a wide range of battery states. The dataset undergoes a structured testing process to validate model accuracy and ensure practical applicability. The steps include:

1. Partitioning the data into training and evaluation subsets while ensuring that feature characteristics remain uniform across both groups.
2. Performing k-fold cross-validation on the training data to prevent overfitting.
3. Evaluating models on the testing set using metrics like root mean square error (RMSE) and R^2 scores.
4. Comparing model outputs against baseline results to validate improvement and statistical significance.

Heatmaps are important tools for data analysis because they make it easier to see trends, relationships, and outliers in large sets of data. They are particularly useful for quickly detecting relationships between variables, which is crucial for understanding the underlying structure of the data.

Thermal visualizations improve the clarity of numerical analysis by employing color scales to depict data magnitudes, ensuring the information remains understandable for a general audience. This visual approach supports data-driven decision-making by providing clear evidence of relationships and trends within the data. Heat maps are also valuable for identifying trends and variations over time or across different categories, thereby aiding in the construction of predictive models and improving the efficiency of data analysis.

Figure 2 provides a heat map that displays the correlation matrix of various battery performance metrics. The color-coded system indicates the strengths and directions of the relationships between these metrics. The coefficients of association range from -1 to 1, where crimson hues signify a direct relationship and azure tones indicate an inverse correlation. The heat map shows that there is a strong negative
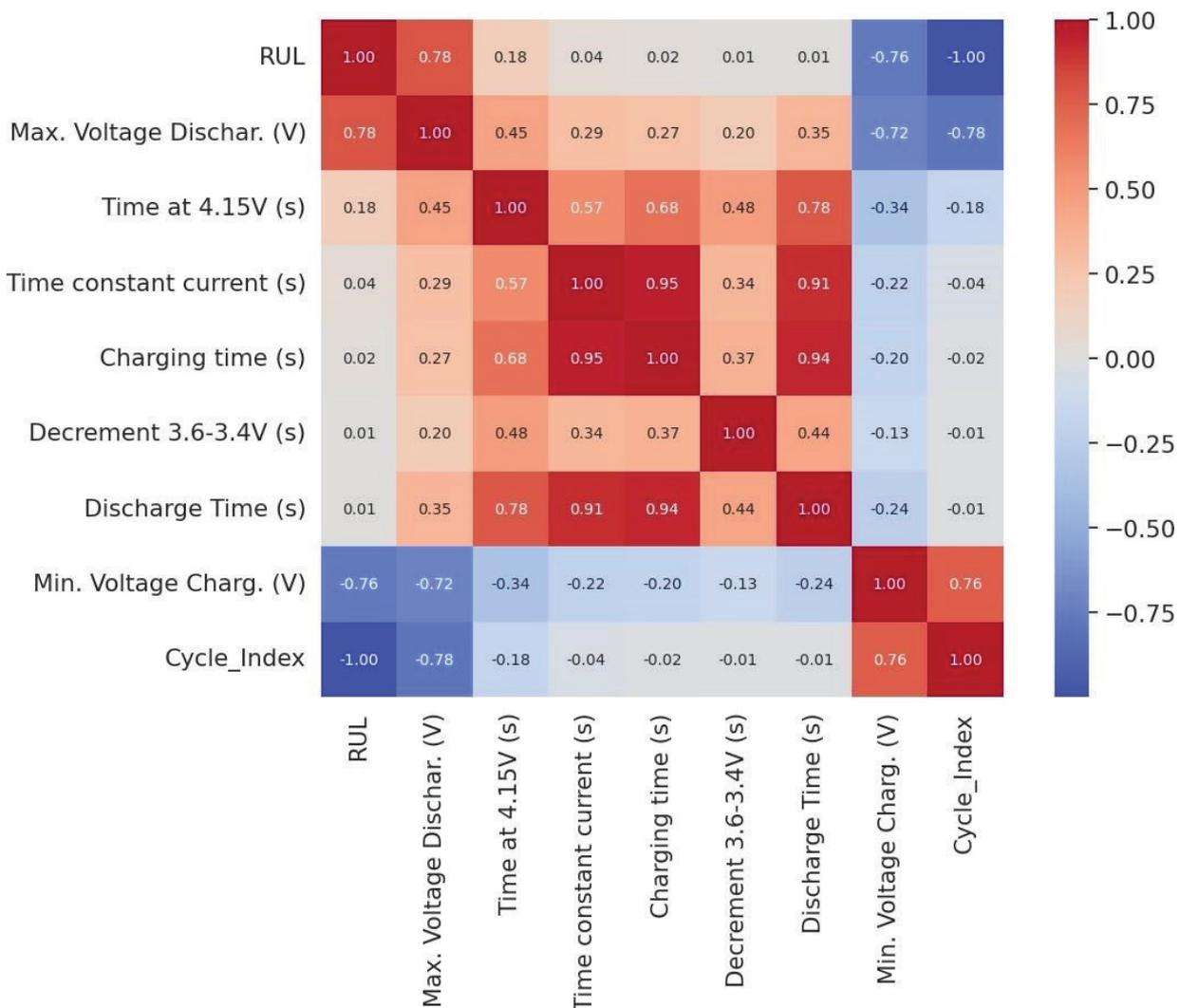

Figure 2. Correlation heatmap of battery performance metrics.
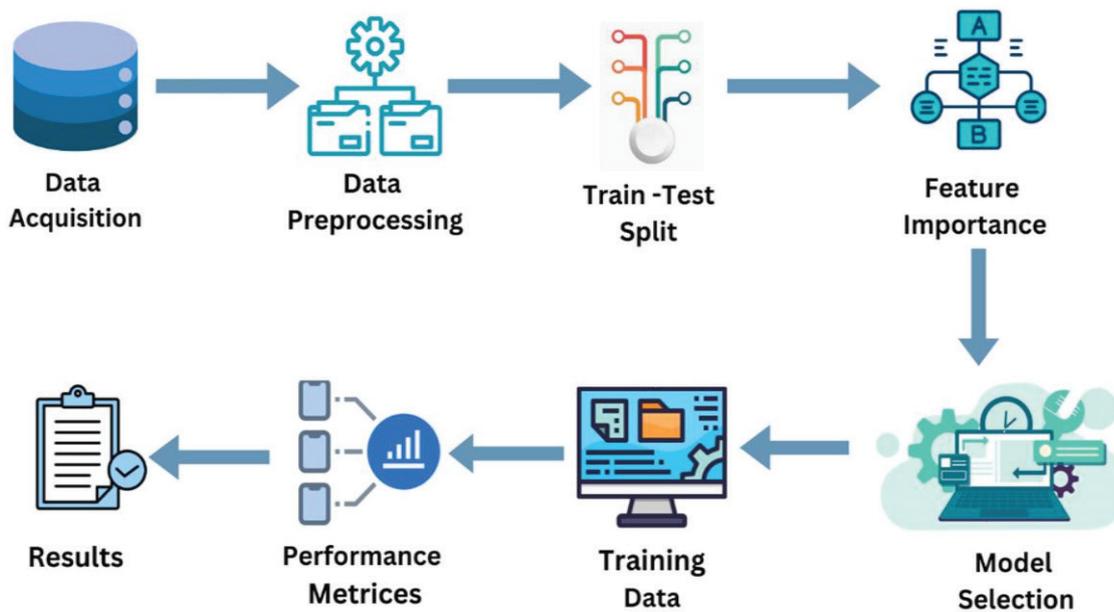
Figure 3. Machine learning project pipeline.

correlation (-1.00) between the RUL and the Cycle Index. This means that the Remaining Useful Life goes down a lot as the number of charge-discharge cycles goes up. Furthermore, a robust direct association (0.76) is observed between the RUL and the minimum voltage charge, suggesting that extended battery longevity is associated with elevated minimum thresholds during the charging process. The maximum voltage discharge also shows a moderately positive correlation (0.78) with RUL.

Time-based metrics such as constant current, charging time, and discharge time exhibit high correlations with each other, highlighting their interdependencies. The minimum charging potential displays a significant inverse relationship with both the Remaining Useful Life and the peak discharge voltage, yielding coefficients of -0.76 and -0.72, respectively. These insights demonstrate the critical factors influencing battery life and performance and underscore the importance of managing cycle counts and voltage levels to optimize battery longevity. Consequently, this thermal matrix offers an all-encompassing overview of the interdependencies among different battery health indicators, facilitating more effective system oversight and the strategic scheduling of upkeep.

FRAMEWORK FOR REMAINING USEFUL LIFE PREDICTION

Within the domain of prognostics and health management (PHM), the precise estimation of a system's RUL is a fundamental requirement. This paper introduces a novel framework integrating metaheuristic optimization techniques and machine learning models to enhance prediction.

Metaheuristic techniques fine-tune the configurations and architectures of machine learning frameworks, enhancing both their precision and stability. This strategy mitigates common drawbacks, including overfitting and inadequate adaptability, found in traditional approaches. Validated with a comprehensive dataset, the proposed framework demonstrates significant performance improvements, offering a more reliable tool for predictive maintenance. The diagram that illustrates the machine learning project pipeline and the proposed predictive model for RUL prediction is also given.

Figure 3 illustrates the typical workflow of a machine learning project. It begins with data acquisition, gathering raw data from various sources. Subsequently, the data preprocessing phase focuses on scrubbing and organizing the raw information to ensure it is ready for analysis. The train-test split stage then splits the data into training and test sets. Feature importance identifies the most influential variables. Model selection follows, choosing the best algorithm for the task. The chosen model is then trained on the training data. Performance metrics evaluate the model's effectiveness. Finally, the model generates results that showcase its predictive capabilities and performance.

BLOCK DIAGRAM OF A PROPOSED PREDICTIVE MODEL FOR RUL PREDICTION

The conceptual framework for estimating Remaining Useful Life (RUL) through machine learning techniques is illustrated in Figure 4. It starts with data collection and preprocessing, including data representation, standardization, and handling null values. The significance of specific

variables is evaluated through a correlation matrix, mutual information scores, and specialized selection algorithms; following this, the data is partitioned into training and validation subsets. Various machine learning algorithms, including linear regression, random forest, bagging regressor, k-nearest neighbour (KNN), decision tree, support vector machine (SVM), AdaBoost, and XGBoost, are employed for RUL prediction. The process concludes with selecting the best algorithm for accurate RUL prediction.

MATERIALS AND METHODS

The research compared metaheuristic-optimized frameworks against conventional benchmarks, such as linear regression and SVM, to measure enhancements in forecasting precision and the minimization of error metrics. Unlike traditional models, metaheuristic techniques apply global optimization, capturing complex, non-linear degradation patterns in battery data more effectively. The research utilized inferential statistics, specifically t-tests and Analysis of Variance (ANOVA), to verify that the recorded improvements—such as lowered RMSE and higher R-squared values—attained statistical significance and were not merely coincidental. These evaluations substantiate the superior precision and reliability that metaheuristic strategies contribute to RUL estimation.

Linear Regression

Linear regression offers a powerful tool for analyzing the battery dataset, revealing insights into factors like discharge time, voltages, and cycle durations that influence RUL. The coefficients are uncovered to minimize discrepancies between predicted and actual RUL values by fitting a linear model. Despite its simplicity, this approach provides valuable insights into the intricate dynamics of battery performance. It functions as a fundamental element of predictive modeling, providing transparent outcomes that clarify linear dependencies in the data and effectively informing protocols for battery upkeep and replacement."From Simple Linear Regression:
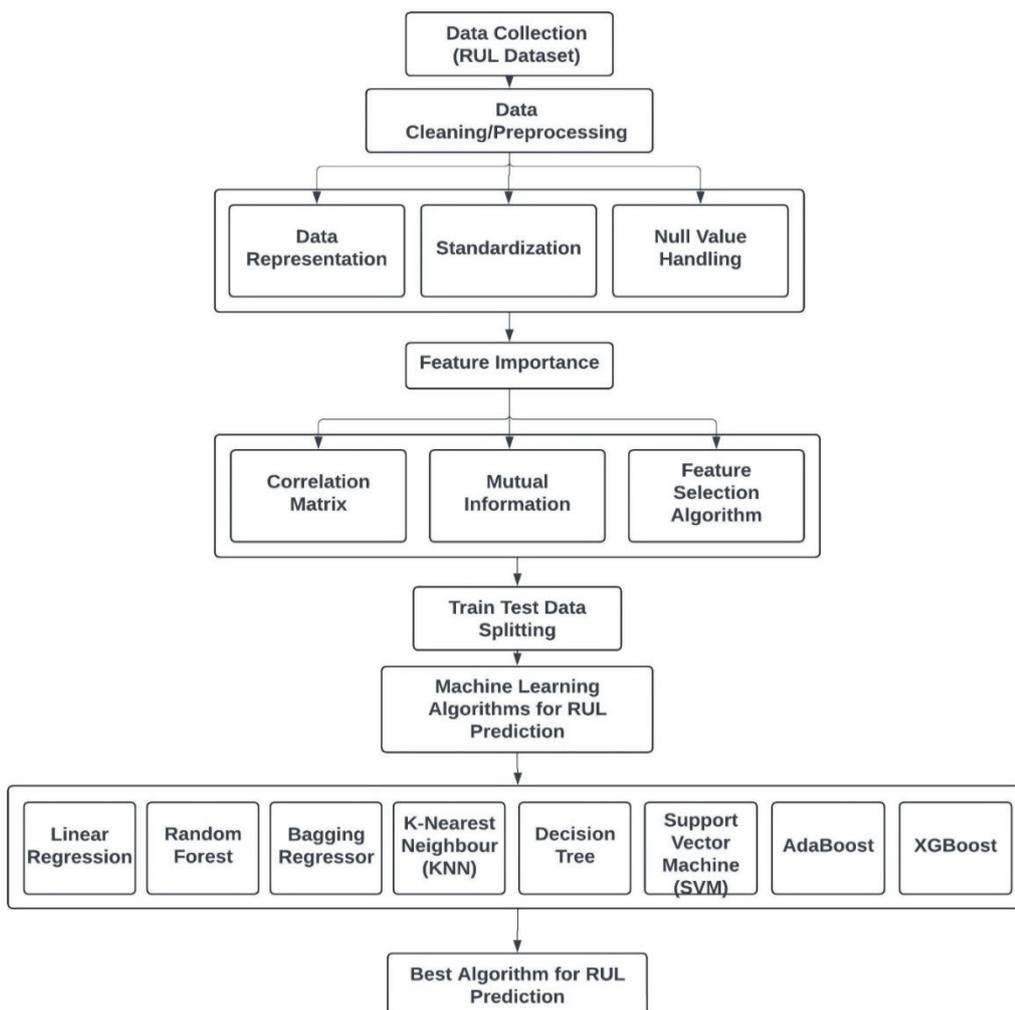

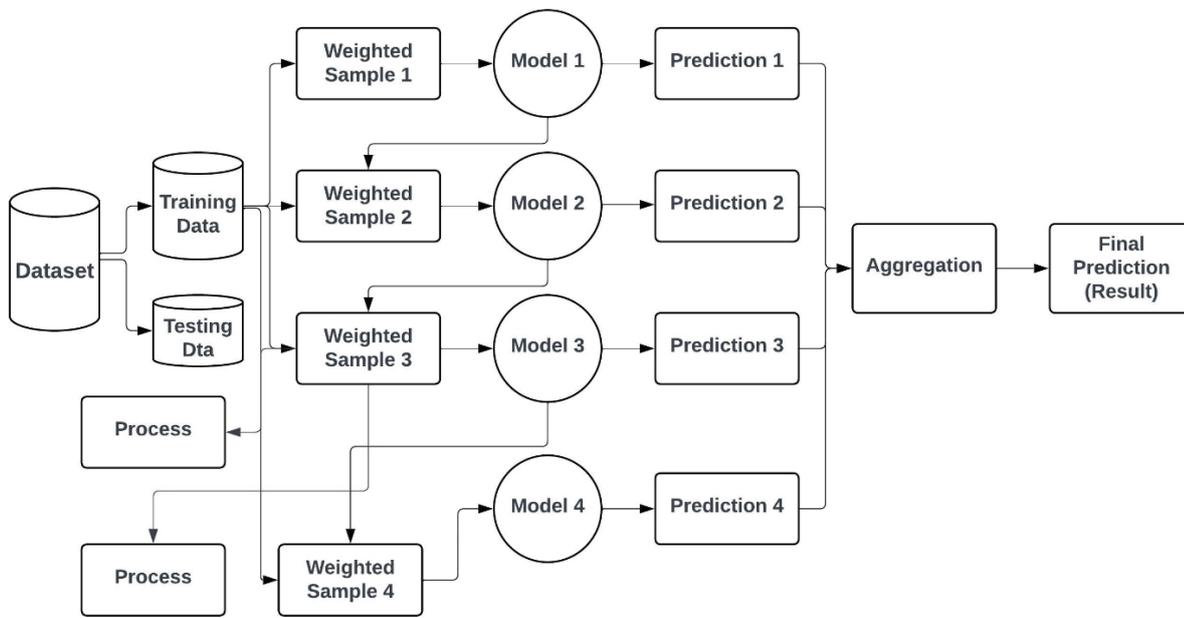**Figure 4.** The proposed predictive model for RUL prediction.

Figure 5. Boosting ensemble regressor process flow.

$$y = \beta0 + \beta1x + \epsilon \tag{1}$$

In (1), there are: y : dependent variable, x : independent variable

$\beta1\ \beta0$: weights

ϵ: error term

From Multiple Linear Regression:

$$y = \beta0 + \beta1x1 + \beta2x2 \ldots + \beta pxp + \epsilon \tag{2}$$

Here in equation (2), p is an independent variable.

Support Vector Machine Regressor

Support vector regression (SVR) could predict RUL based on various operational parameters. Each cycle's features include discharge time, voltage changes during discharge and charge, times at specific voltage levels, and other relevant metrics. SVR aims to identify the optimal hyperplane that fits the data while allowing for a margin of error defined by epsilon. This boundary defines a tolerance envelope surrounding the regressed function, permitting a specific range of permissible variations. The regularization parameter balances the flatness of the function with the allowance for deviations, ensuring robust performance in predicting continuous RUL values based on the dataset's detailed operational characteristics.

The final regression function is given as:

$$f(x) = \sum_{i=1}^{n} (\alpha_i - \alpha_i')K(x_i, x) + b \tag{3}$$

Equation (3) has α_i and α_i' are the Lagrange multipliers obtained from the optimization problem, and

$K(x_i, x)$ is the kernel function.

Boosting

Boosting theory is a machine learning framework aimed at improving the accuracy of predictions by combining multiple weak learners to form a strong learner. It involves sequentially training models, each focusing on correcting the errors of its predecessors. Boosting can handle a variety of data types and has shown great success in reducing bias and variance, thereby enhancing model performance. Insufficient regulation can lead to overfitting, however, since it is computationally intensive.

Adaboost ensemble regressor

AdaBoost can be employed to enhance the prediction of RUL based on various operational parameters. Each cycle is characterized by features such as Discharge Time, voltage changes during discharge and charge, times at specific voltage levels, and other relevant metrics. AdaBoost, as an ensemble method, iteratively improves regression models by focusing on the errors of each successive model. It adjusts the weights of training data points, giving more importance to instances where predictions were inaccurate. By combining weak learners effectively, AdaBoost can create a robust predictive model that accurately forecasts RUL using the dataset's detailed operational data.

Here, each data point has the same weights initially as represented in equation (4):

$$\omega_i = \frac{1}{N} \; for \; i = 1,2,3, \ldots \ldots, N \tag{4}$$

The weights that are in error should be increased through the formula (5):

$$\alpha_t = log \, log \, \left(\frac{1-\epsilon_t}{\epsilon_t} \right) \qquad (5)$$

XGBoost ensemble regressor

XGBoost excels at predicting RUL using features like Discharge Time, voltage dynamics, and cycle durations. The algorithm reduces the mean squared error through the iterative refinement of decision trees, which sequentially corrects prediction errors by focusing on the remaining residuals. XGBoost's advanced capabilities include regularization to prevent overfitting, methods for managing missing data, and efficient tree-learning techniques, making it adept at handling large datasets. Due to its proficiency in generating precise regression outputs, XGBoost excels at estimating RUL, which is vital for maximizing battery longevity and securing dependable operational strategies derived from extensive cycling datasets.

Decision Tree Regressor

Decision tree techniques can address both regression and classification tasks by forming a tree structure with internal nodes representing features, branches as decision rules, and leaf nodes as outcomes. Splits are based on criteria like Gini impurity or entropy to optimize data separation. For a dataset with battery performance metrics, decision trees predict the RUL by evaluating features such as cycle count, discharge times, and voltage levels. In regression tasks, this method reduces data variability, leading to precise and reliable predictions of continuous values like RUL, making it an effective tool for battery life estimation.

Gini Index Impurity formula:

For a dataset S with n classes, the Gini impurity G(S) is calculated as:

$$G(S) = 1 - \sum_{i=1}^{n} p_i^2(x) \qquad (6)$$

In equation (6), pi represents the number of samples in dataset S that belong to class i. For a dataset S with n classes, the entropy H(S) is calculated as:

$$H(S) = \sum_{i=1}^{n} p_i(p_i) \qquad (7)$$

In equation (7), pi denotes the proportion of samples belonging to class i within the dataset S.

K-Nearest Neighbor Regressor

In the dataset of battery cycles, KNN could predict RUL by comparing new cycle data with historical records based on discharge times, voltage changes, and other metrics. KNN identifies 'K' nearest neighbors using distance measures like Euclidean distance (8), then averages their RUL values to estimate when a battery may need replacement. This instance-based method leverages the dataset's specifics to provide localized RUL predictions, making it suitable for understanding and managing battery longevity based on past operational patterns.

$$Euclidean \; distance \; d(x,y) = \sqrt{\sum_{i=1}^{n} \left(x_i^2 - y_j^2 \right)} \qquad (8)$$

Bagging Regressor

The bagging regressor is an ensemble learning method that enhances the accuracy of regression models by combining predictions from multiple instances of the same base estimator, such as decision trees. It employs a technique called Bootstrap Aggregating, or Bagging, where multiple subsets of the training data are created through sampling with replacement. Each subset is used to train a separate
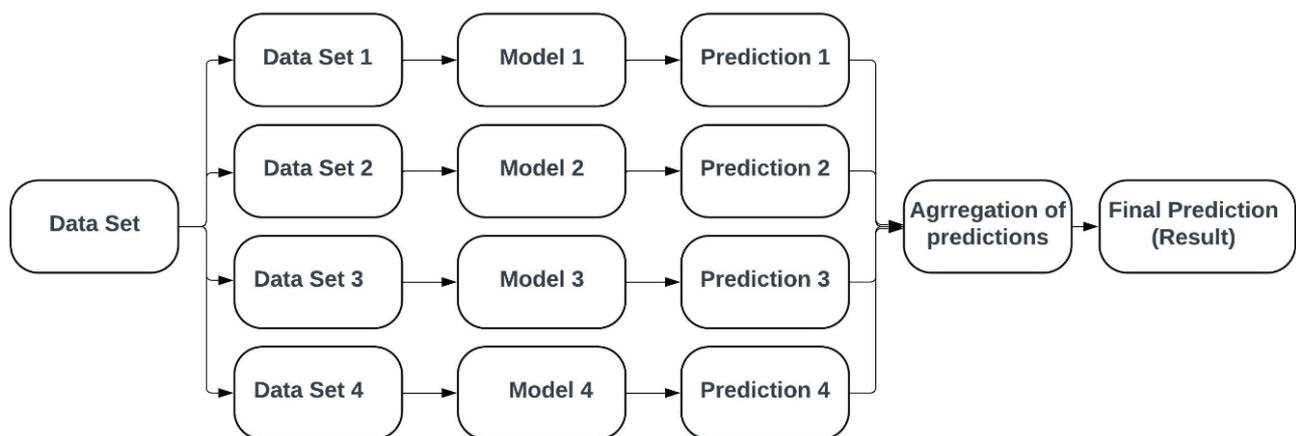

Figure 6. Bagging regressor workflow diagram.

regressor, and the predictions from all models are averaged to produce the final output. This approach reduces overfitting by lowering variance, making it especially effective for high-variance models. Additionally, since each model is trained independently, the method benefits from computational efficiency through parallelization. Bagging Regressors are particularly useful in regression tasks where the base estimator might overfit the data.

Figure 6 depicts the bagging method, an effective ensemble approach in machine learning. The process starts by arbitrarily sampling the novel dataset with replacement to create different overlapping subsets. Each of these subsets trains a distinct model autonomously, leading to a set of disparate models (Model 1, Model 2, Model 3, and Model 4), each shaped by different data samples. Each model generates its own prediction, and these predictions are then combined—using techniques such as averaging for regression or majority voting for classification—to produce a unified final prediction. By aggregating these results, the method increases overall accuracy and robustness, reduces variance, and helps prevent overfitting.

Random Forest

The random forest algorithm leverages an ensemble approach based on bagging, where a set of decision trees operates together to ameliorate predictive performance and mitigate the likelihood of overfitting. It accomplishes this by creating multiple trees, each built from randomly resampled training data and by opting for random features for every split within the tree. For example, when working with data on battery performance, Random Forest estimates remaining useful life (RUL) by consolidating the outputs of individual trees, each trained on different subsets of the data By consolidating a diverse array of variables—including discharge intervals, voltage measurements, and charging periods—this methodology produces a framework that demonstrates superior precision and consistency over a solitary decision tree.

$$\hat{y}(x) = \frac{1}{n}\sum_{i=1}^{n} p_i(x) \qquad (9)$$

Let,

p1, p2 ….., pn: number of decision trees in the random forest

pi(x): forecast of the i-th tree of x, given an input x.

$\hat{y}(x)$: aggregated prediction of the Random Forest for the input x.

Equation (1) represents the averaging mechanism used in Random Forest regression, where the final prediction is the mean of the predictions made by all the trees.

FUNDAMENTAL PERFORMANCE CRITERIA

For regression problems, the principal evaluation metrics include mean absolute error (MAE), mean squared error (MSE), RMSE, and the coefficient of determination (R^2). These metrics measure the forecasting fidelity of the model, the scale of its deviations, and its capacity to identify the fundamental trends within the dataset.

Mean Absolute Error (MAE)

The mean absolute error (MAE) serves as an indicator for evaluating the precision of predictions by calculating the mean of the absolute deviations between forecasted and actual observations over the test dataset. In the context of Remaining Useful Life (RUL) estimation, MAE quantifies the typical size of prediction errors considering cycle-specific parameters such as Discharge Time, voltage variations during charging and discharging, and different operational durations. In an optimal scenario, the MAE converges toward zero, signifying a negligible variance between the forecasted values and the observed results.

$$MAE = \frac{1}{n}\sum_{i=1}^{n} |y_i - \hat{y}_i| \qquad (10)$$

In equation (10), y_i is the actual value, and $\hat{y}_i$ is the predicted value, and n is the number of observations.

Mean Square Error (MSE)

The mean squared error (MSE), defined in Equation (11), represents a measure that computes the mean of the squared deviations between predicted and observed values within a dataset. In the context of RUL estimation, the MSE quantifies the average quadratic deviation of the errors by accounting for cycle-dependent variables, including discharge durations, voltage instabilities during cycling, and different operational temporal windows. Ideally, the MSE value is close to zero, reflecting minimal divergence between predicted and actual results.

$$MSE = \frac{1}{n}\sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \qquad (11)$$

R² Score

The R^2 score, as defined in Equation (13), evaluates the extent to which a regression model can forecast Remaining Useful Life (RUL) using parameters such as Discharge Time, voltage fluctuations, and cycle durations. It measures the fraction of variance in RUL that is accounted for by these features. An elevated coefficient of determination (R^2) denotes a superior model fit, suggesting the system successfully accounts for fluctuations in RUL; conversely, a diminished value suggests the model fails to align closely with the empirical measurements.

This metric is essential for assessing the regression model's capability to reliably predict RUL and understand the influence of cycle-specific characteristics on battery lifespan. Ideally, the R^2 score approaches 1.

$$R^2 = 1 - \frac{ss_r}{ss_t} \qquad (12)$$

Where,

ss_r: Residual sum of squares

ss_t: Total sum of squares

RESULTS AND DISCUSSION

This portion of the study evaluates the results derived from the introduced architecture for estimating Remaining Useful Life (RUL). Model performance has been evaluated using established metrics, including RMSE and the R^2 coefficient. RMSE is employed to measure the accuracy of the predictions by calculating the extent of deviation between the estimated and actual RUL values. Simultaneously, the R^2 coefficient assesses the model's explanatory power by comparing its performance against standard baseline references.

The dataset includes key features such as discharge time, voltage variations, and cycle index, which serve as input for machine learning models to predict Remaining Useful Life. To facilitate precise inputs for predictive modeling, the data was normalized and organized for algorithms such as Random Forest and XGBoost, allowing accurate estimation of RUL under a variety of battery operating conditions. Comparative analysis shows notable gains in forecast accuracy and model resilience over conventional techniques. The consequences of these findings are examined, emphasizing their importance for streamlining maintenance scheduling and proposing avenues for subsequent research to bolster predictive performance in practical applications. The experimental findings offer valuable guidance on selecting models for practical battery management applications, taking into account metrics such as prediction accuracy, processing efficiency, and robustness. By systematically evaluating the dataset, it is confirmed that models like Random Forest and XGBoost can reliably predict battery RUL under operational conditions. These observations establish their effectiveness and reliability,

positioning them as optimal choices for precise and robust battery life assessment in dynamic, real-life environments.

Table 1 presents a comparative evaluation of several regression models using performance indicators such as Training Score, Testing Score, RMSE, R^2 Score, and computational Execution Time. Linear Regression shows moderate predictive capability, recording a Training Score of 79.73 and a Testing Score of 78.14, accompanied by the highest RMSE of 148.52 and an R^2 of 80.44. While it provides reduced forecasting precision, it exhibits the shortest processing time at a mere 0.038 seconds, underscoring its superior computational speed. On the other hand, the Random Forest model delivers exceptional results, with Training and Testing Scores of 99.92 and 99.64, respectively, a minimal RMSE of 23.26, and an almost perfect R^2 of 99.68, although it requires a longer runtime of 8.319 seconds. In a similar fashion, the Bagging Regressor demonstrates high-tier proficiency, yielding Training and Testing scores of 99.88 and 99.42, respectively, alongside an RMSE of 24.05 and an R^2 of 99.62, all while offering a competitive processing duration of 0.774 seconds. The K-Nearest Neighbors (KNN) model also performs strongly, attaining Training and Testing Scores of 99.67 and 99.08, an RMSE of 30.34, and an R^2 of 98.70, with a relatively short execution duration of 0.170 seconds, demonstrating a favorable trade-off between speed and accuracy.

Strong performance is demonstrated by the Decision Tree model, with Training and Testing Scores of 99.99 and 98.85, respectively. The model records an RMSE of 33.99 and an R^2 value of 99.33, completing execution in 0.245 seconds. While SVM (Support Vector Machine) achieves reasonable Training and Testing Scores of 90.09 and 89.46, its higher RMSE of 103.14, lower R^2 of 89.70, and the longest runtime of 20.086 seconds indicate reduced efficiency for this dataset. AdaBoost delivers a balanced outcome, with Training and Testing Scores of 94.40 and 94.03, an RMSE of 77.62, an R^2 of 93.94, and an intermediate execution duration of 2.178 seconds. XGBoost excels in predictive performance, attaining Training and Testing Scores of 99.95 and 99.57, the lowest RMSE of 20.64, and an R^2 of 99.61, albeit requiring 6.871 seconds for execution.

Table 1. Performance analysis of predictive models

| Algorithm name	Testing score	Training score	RMSE	R^2 Score	Execution time (s)
Random forest	99.64	99.92	23.26	0.9999	0.319
XGBoost	99.57	99.95	20.64	0.9961	6.871
Bagging regressor	99.42	99.88	24.05	0.9962	0.774
KNN	99.08	99.67	30.34	0.9870	0.170
Decision tree	98.85	99.99	33.99	0.9933	0.245
AdaBoost	94.03	94.40	36.62	0.9394	2.178
SVM	89.46	90.09	42.14	0.8970	0.086
Linear regression	78.14	79.73	48.52	0.8044	0.038
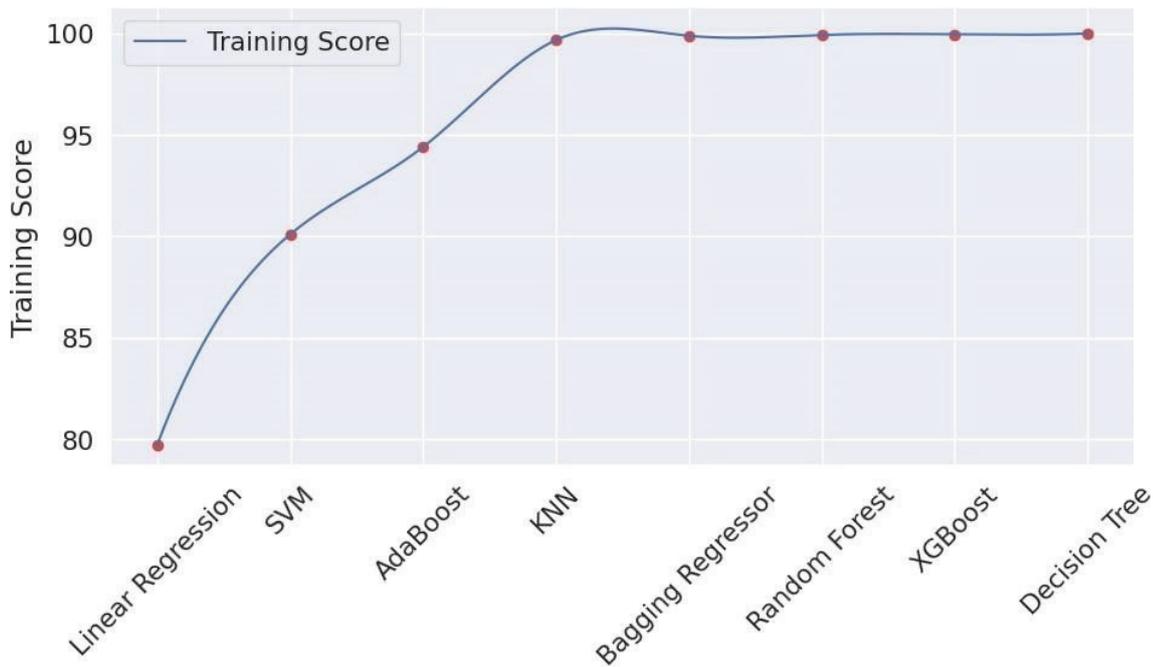
Figure 7. Comparative analysis of training scores for ML-based RUL prediction.

The results summarized in Table 1 emphasize the trade-offs between computational time and prediction accuracy among the evaluated regression algorithms. Random Forest and XGBoost are recognized as the superior frameworks, distinguished by their exceptional forecasting accuracy and their ability to minimize residuals when evaluated against the test data. Based on these findings, both models are considered highly suitable for practical battery management tasks, providing reliable RUL estimates while balancing speed, robustness, and accuracy. Through a methodical assessment of the data and evaluation metrics, vital intelligence is extracted to support preemptive battery servicing and the refinement of resource allocation within practical environments.
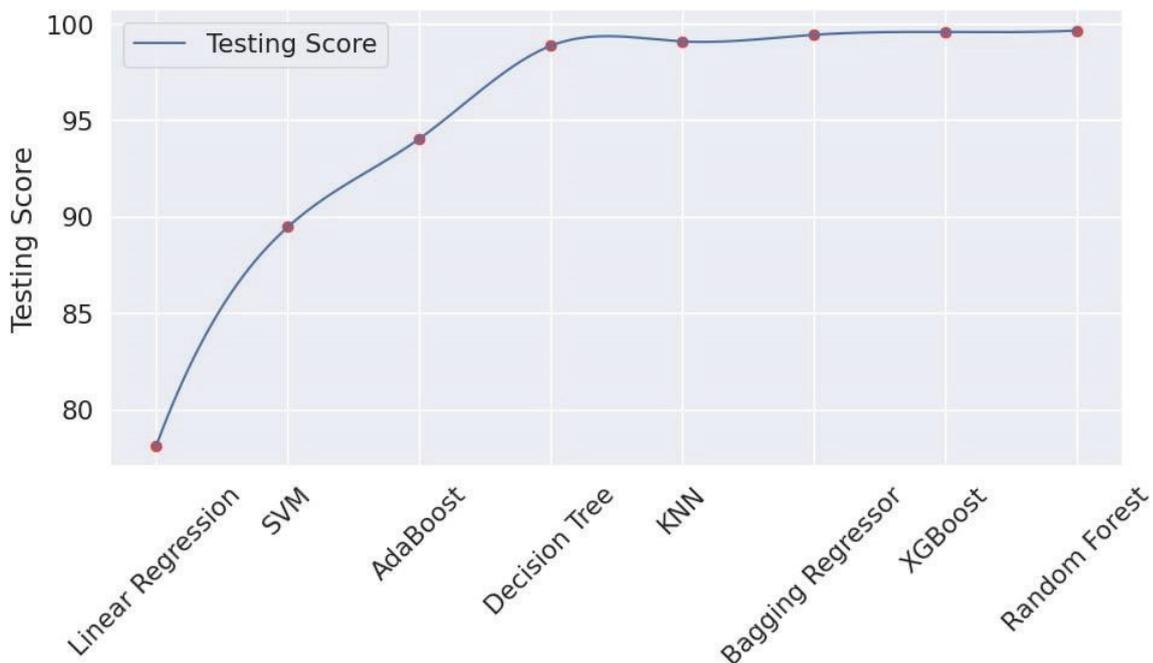

Figure 8. Assessment of testing performance in machine learning models for RUL forecasting.

The training scores of various machine learning algorithms are presented in Figure 7, with models organized in ascending order of their training performance. The evaluated algorithms comprise Linear regression, decision tree, random forest, XGBoost, bagging regressor, K-nearest neighbors (KNN), AdaBoost, and support vector machine (SVM). Each algorithm is represented along the x-axis in the plot, while its corresponding training scores are plotted along the y-axis. A smooth line curve is generated using cubic spline interpolation to provide a clearer trend
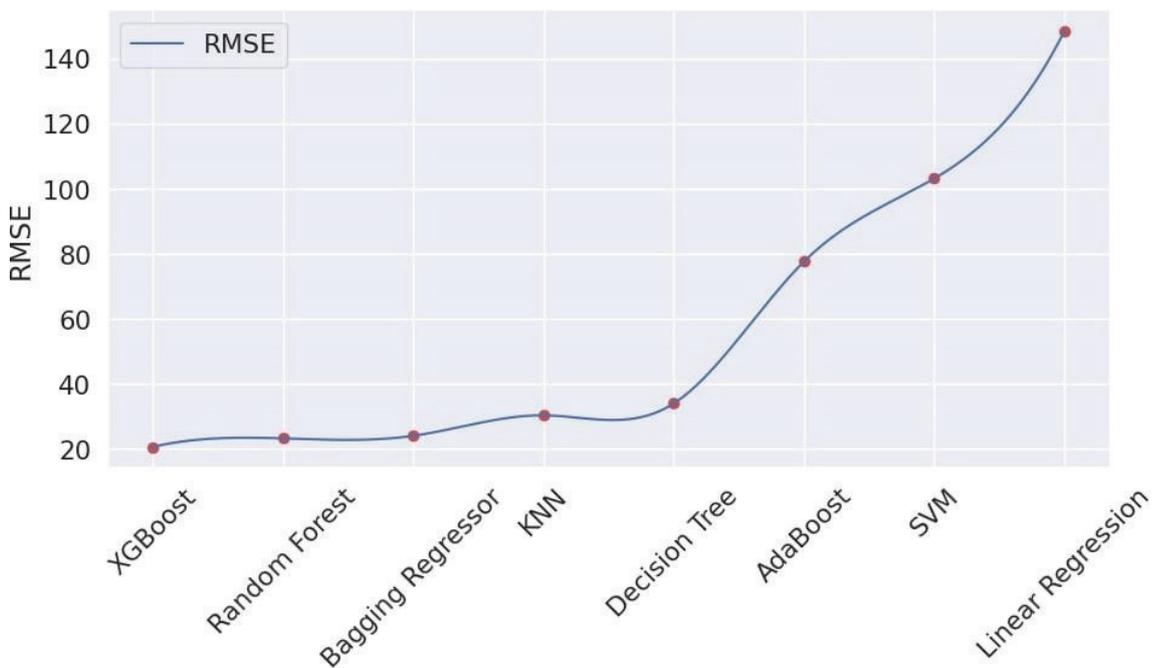

Figure 9. Root mean squared error comparison for ML-based RUL prediction.

Figure 10. Assessing the predictive accuracy of machine learning models using R^2 metrics.

visualization. The red markers indicate the actual data points. The figure demonstrates a gradual progression in training performance, starting from Linear Regression, which exhibits the lowest score, up to Decision Tree, which achieves the highest. Most algorithms, particularly from K-nearest neighbors (KNN) onward, attain training scores approaching 100, reflecting exceptionally strong learning on the training dataset. Such high performance may indicate overfitting, especially when not accompanied by a corresponding evaluation on test data. Conversely, the training scores for linear regression and support vector machine (SVM) are more modest. This indicates that these specific algorithms demonstrate a reduced capacity for capturing underlying training patterns when compared to the alternative models evaluated.

The testing performance of different machine learning algorithms is visualized in Figure 8, where each model is plotted along the x-axis and its corresponding testing score is shown on the y-axis. Based on their respective testing performance, the evaluated algorithms—ranging from Linear Regression and SVM to XGBoost and Random Forest—have been organized in ascending order. This list also includes AdaBoost, decision tree, KNN, and bagging regressor. To better illustrate the overall trend, a smooth curve generated through cubic spline interpolation is overlaid, while red markers denote the actual observed data points. Moving across the sequence of models, the data reveals a steady climb in testing outcomes. Random forest achieves the maximum value, contrasting sharply with Linear Regression's baseline. Notably, the most dramatic increase is situated between the first two models, giving way to a more consistent, gradual rise thereafter. Testing scores for models starting from KNN are nearly 100, reflecting excellent generalization capability on unseen data. Nevertheless, this sustained high performance requires rigorous cross-validation to confirm the results aren't skewed by overfitting. This is especially critical in cases where the training metrics are similarly high. In contrast, the comparatively lower testing scores of Linear Regression and SVM suggest a more conservative fit to the data, which may support better generalization in certain practical scenarios.

The RMSE for several regression models is illustrated in Figure 9 using a smooth line. Error values are ordered from lowest to highest to facilitate direct comparison among the algorithms. Along the x-axis, the evaluated models—XGBoost, Random Forest, Bagging Regressor, K-Nearest Neighbors (KNN), Decision Tree, AdaBoost, Support Vector Machine (SVM), and Linear Regression—are displayed with rotated labels for improved readability. The vertical axis corresponds to RMSE values, ranging from 20 to 150. For better visibility of the general trend, the data points are linked via spline interpolation. Distinct red markers are utilized to draw attention to the specific observed measurements. This visual arrangement clearly demonstrates differences in predictive accuracy, where lower RMSE values reflect stronger model performance. By prioritizing a comparative assessment of error, the title "Root Mean Squared Error Comparison for ML-Based RUL Prediction" effectively summarizes the graph's objective.
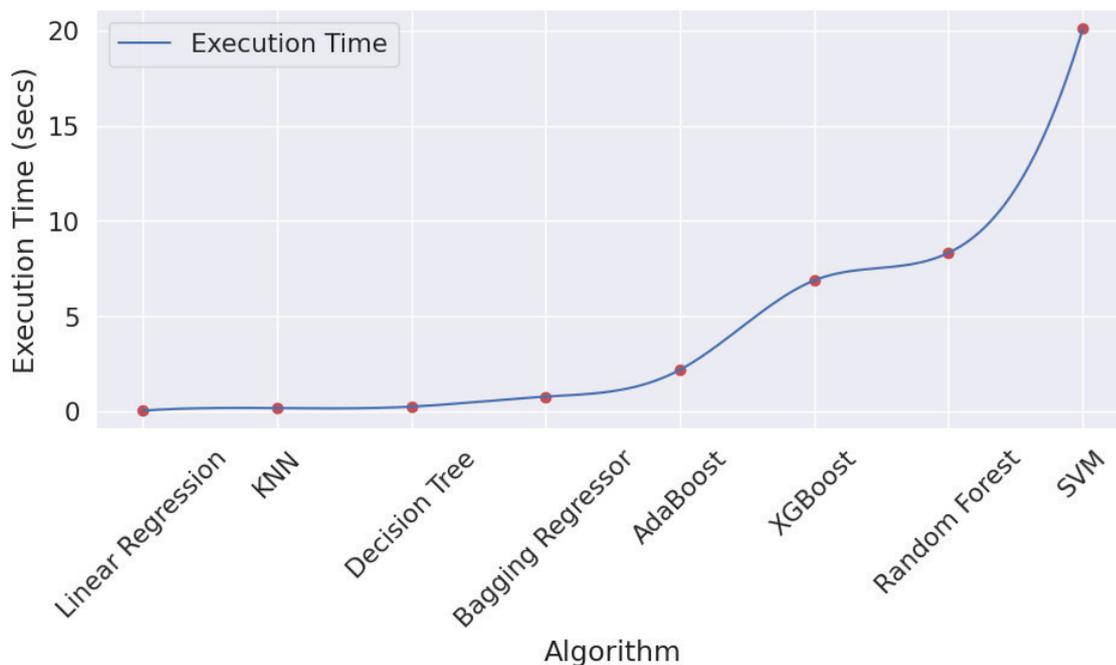

Figure 11. Computational time analysis of machine learning models for RUL prediction.

Table 2. Performance comparison of reference paper

ML algorithms	MSE	RMSE	MAE	R-Squared	Time
Random forest regression	14.1186	3.7574	2.0930	0.9998	0.0651
Decision tree regression	26.8489	5.1816	2.2296	0.9997	0.1510
Linear regression	54.5430	7.3853	4.6441	0.9994	0.0248
Bayesian network	54.5478	7.3856	4.6483	0.9993	1.2356
Gradient boosting regression	57.4476	7.5794	4.9842	0.9990	2.5640

Using a continuous curve, Figure 10 displays the coefficient of determination values achieved by different regression techniques. The models are sequenced from lower to higher R^2 performance, with Linear Regression, Support Vector Machine (SVM), AdaBoost, K-Nearest Neighbors (KNN), Decision Tree, XGBoost, Bagging Regressor, and Random Forest positioned along the horizontal axis, where label rotation improves visibility. The vertical axis captures R^2 magnitudes spanning from 80 to 100. Individual results are connected through spline-based smoothing, and the measured values are emphasized with red data markers. This presentation facilitates straightforward evaluation of predictive strength, as greater R^2 values indicate improved explanatory capability. Clearly stated in the title, "R^2-Based Performance Comparison of Machine Learning Algorithms," the figure underscores comparative assessment of model performance. This graph clearly illustrates the capacity of each model to explain the variance within the target variable. Higher values in this context represent a more precise and accurate modelling approach.

A continuous visualization of runtime performance is shown in Figure 11, where regression models are ranked based on their temporal requirements. This arrangement moves from the model with the minimum execution period to the one with the maximum. The algorithms displayed along the horizontal axis include Linear Regression, K-Nearest Neighbors (KNN), Decision Tree, Bagging Regressor, AdaBoost, XGBoost, Random Forest, and Support Vector Machine (SVM), with axis labels rotated to enhance clarity. Within a range of 0 to 20 on the vertical scale, the processing durations are visualized. A spline-based curve tracks the trend between measurements, while the specific observations are highlighted using red data points. This graphical representation enables a clear comparison of computational efficiency, where reduced execution time corresponds to faster model performance. Regarding execution speed, Linear Regression and KNN are the fastest-performing algorithms, whereas SVM sits at the opposite end of the spectrum with the maximum processing time. Through this comparative visualization, valuable insight is provided into the suitability of each algorithm for practical deployment, particularly in time-critical applications or large-scale datasets. By centering on the resource requirements of each model, the figure title "Computational Time Analysis of Machine Learning Models for RUL Prediction" accurately describes the graph's focus on efficiency.

Comparative Analysis

The Table 2 shows the result of similar research. By comparing results of this paper with recent research, it is shown that this paper's results are more effective and efficient for research applications, primarily due to its higher R-squared values and strong testing and training scores, which indicate a precise model fit and robust generalizability.

Models like Random Forest and XGBoost in Table 1 achieve R-squared values close to 0.9999, suggesting they capture nearly all data variability, which is critical for ensuring reliability in predictive tasks. As shown in Table 1, the evaluated models represent a broad spectrum of computational demands. While SVM and KNN offer rapid results, XGBoost serves as a more intensive alternative. Such a variety provides practitioners with the freedom to prioritize either low-latency performance or high-level accuracy based on their specific application needs. Although Table 1 has slightly higher RMSE values compared to Table 2, the extremely high R-squared values and consistency between testing and training scores offset this drawback, making Table 1's models more suited to applications where precision and generalizability are essential. Furthermore, ensemble approaches such as Bagging and Boosting improve computing efficiency by minimizing variance while focusing on prediction accuracy. The integration of metaheuristic techniques refines model training and boosts execution speed. Consequently, these optimized algorithms become highly effective for use cases that demand rapid response times and time-critical decision-making. These optimizations are critical for battery management systems that need real-time analysis because they enable quick, accurate forecasts while controlling computing loads.

CONCLUSION

The primary focus of this study is on accurately predicting the remaining useful life (RUL) of lithium-ion batteries, particularly within renewable energy storage systems. This area is significant because accurate RUL predictions are essential for battery management in electric vehicles (EVs) and renewable energy systems, which rely on efficient

energy storage to meet sustainability goals. The optimization of battery life cycles through dependable RUL estimates serves to minimize maintenance overhead while simultaneously supporting sustainability. This approach enhances the overall balance between commercial productivity and ecological stewardship. This research proposes a range of machine learning models, including Random Forest, Decision Tree, Linear Regression, K-Nearest Neighbours, Support Vector Machine, Adaptive Boosting (AdaBoost), and Extreme Gradient Boosting (XGBoost Metaheuristic-based parameter selection served as a catalyst for improving the RUL performance of the implemented ML models, ensuring more precise predictive outcomes.

This research was carried out through the Hawaii Natural Energy Institute (HNEI) battery lifecycle dataset for machine learning model training and validation. Random Forest and XGBoost emerged as top performers, achieving R^2 scores of 99.68% and 99.61%, respectively, and RMSEs of 23.26 and 20.64, respectively. The data highlights how advanced computational modeling significantly upgrades RUL prediction capabilities. Specifically, the fusion of metaheuristic parameter optimization and ensemble-based learning proved instrumental in reaching superior levels of predictive accuracy. This research improves the management of renewable energy resources by efficiently predicting the Remaining Useful Life to improve sustainability goals.

AUTHORSHIP CONTRIBUTIONS

Authors equally contributed to this work.

DATA AVAILABILITY STATEMENT

The authors confirm that the data that supports the findings of this study are available within the article. Raw data that support the finding of this study are available from the corresponding author, upon reasonable request.

CONFLICT OF INTEREST

The author declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

ETHICS

There are no ethical issues with the publication of this manuscript.

STATEMENT ON THE USE OF ARTIFICIAL INTELLIGENCE

Artificial intelligence was not used in the preparation of the article.

REFERENCES

[1] Sankarasubramanian S, Krishnamurthy B. A capacity fade model for lithium-ion batteries including diffusion and kinetics. Electrochim Acta 2012;70:248–254.[CrossRef]

[2] Wang Y, Yang D, Zhang X, Chen Z. Probability based remaining capacity estimation using data-driven and neural network model. J Power Sources 2016;315:199–208. [CrossRef]

[3] Safavi V, Mohammadi Vaniar A, Bazmohammadi N, Vasquez JC, Guerrero JM. Battery remaining useful life prediction using machine learning models: A comparative study. Information 2024;15:124. [CrossRef]

[4] Wang Y, Pan R, Yang D, Tang X, Chen Z. Remaining useful life prediction of lithium-ion battery based on discrete wavelet transform. Energy Procedia 2017;105:2053–2058. [CrossRef]

[5] Park K, Choi Y, Choi WJ, Ryu HY, Kim H. LSTM-based battery remaining useful life prediction with multi-channel charging profiles. IEEE Access 2020;8:20786–20798. [CrossRef]

[6] Hu C, Jain G, Zhang P, Schmidt C, Gomadam P, Gorka T. Data-driven method based on particle swarm optimization and k-nearest neighbor regression for estimating capacity of lithium-ion battery. Appl Energy 2014;129:49–55. [CrossRef]

[7] Kamul A, Hananda N, Harito C, Djuana E, Elwirehardja G, Pardamean B, et al. Machine learning application in battery prediction: A systematic literature review and bibliometric study. E3S Web Conf 2022;338:01021. [CrossRef]

[8] Venugopal P, et al. Analysis of optimal machine learning approach for battery life estimation of Li-ion cell. IEEE Access 2021;9:159616–159626. [CrossRef]

[9] Feng X, et al. Online state-of-health estimation for Li-ion battery using partial charging segment based on support vector machine. IEEE Trans Veh Technol 2019;68:8583–8592. [CrossRef]

[10] Richardson RR, Osborne MA, Howey DA. Battery health prediction under generalized conditions using a Gaussian process transition model. J Energy Storage 2019;23:320–328. [CrossRef]

[11] Wu J, Kong L, Cheng Z, Yang Y, Zuo H. RUL prediction for lithium batteries using a novel ensemble learning method. Energy Rep 2022;8:313–326. [CrossRef]

[12] Ansari S, Ayob A, Lipu MSH, Hussain A, Saad MHM. Remaining useful life prediction for lithium-ion battery storage system: A comprehensive review of methods, key factors, issues and future outlook. Energy Rep 2022;8:12153–12185. [CrossRef]

[13] Wang A, Chen H, Jin P, Huang J, Feng D, Zheng M. RUL estimation of lithium-ion power battery based

on DEKF algorithm. In: Proceedings of the 2019 14th IEEE Conference on Industrial Electronics and Applications (ICIEA); 2019. p. 1851–1856. [CrossRef]

[14] Safavi V, Mohammadi Vaniar A, Bazmohammadi N, Vasquez JC, Guerrero JM. Battery remaining useful life prediction using machine learning models: A comparative study. Information 2024;15:124. [CrossRef]

[15] Richardson RR, Osborne MA, Howey DA. Battery health prediction under generalized conditions using a Gaussian process transition model. J Energy Storage 2019;23:320–328. [CrossRef]

[16] Que Z, Xu Z. A data-driven health prognostics approach for steam turbines based on xgboost and DTW. IEEE Access 2019;7:1–1. [CrossRef]

[17] Lipu MSH, Hannan MA, Hussain A, Hoque MM, Ker PJ, Saad MHM, et al. A review of state of health and remaining useful life estimation methods for lithium-ion battery in electric vehicles: Challenges and recommendations. J Clean Prod 2018;205:115–133. [CrossRef]

[18] Xu B, Oudalov A, Ulbig A, Andersson G, Kirschen DS. Modeling of lithium-ion battery degradation for cell life assessment. IEEE Trans Smart Grid 2018;9:1131–1140. [CrossRef]

[19] Eissa MA, Chen P. Machine learning-based electric vehicle battery state of charge prediction and driving range estimation for rural applications. IFAC-PapersOnLine 2022;56(3). [CrossRef]

[20] Wang N, Lyu Y, Zhou Y, Luan J, Li Y, Zheng C. A hybrid framework for remaining driving range prediction of electric taxis. Sustain Energy Technol Assess 2024;67:103832. [CrossRef]

[21] Sekhar JNC, Domathoti B, Santibanez Gonzalez EDR. Prediction of battery remaining useful life using machine learning algorithms. Sustainability 2023;15:15283. [CrossRef]

[22] Babaeiyazdi I, Rezaei-Zare A, Shokrzadeh S. State of charge prediction of EV Li-ion batteries using EIS: A machine learning approach. Energy 2021;223:120116. [CrossRef]

[23] Lai X, Huang Y, Gu H, Han X, Feng X, Dai H, et al. Remaining discharge energy estimation for lithium-ion batteries based on future load prediction considering temperature and ageing effects. Energy 2022;238:121754. [CrossRef]

[24] Maheshwari A, Nageswari S. Modelling and State of Charge Estimation of Li-Ion Battery for Electric Vehicle. In: Power Electronics and High Voltage in Smart Grid: Select Proceedings of SGESC 2021. Singapore: Springer Nature Singapore; 2022. p. 139–151. [CrossRef]

[25] Wang L, Ma J, Zhao X, Li X, Zhang K, Jiao Z. Adaptive robust unscented Kalman filter-based state-of-charge estimation for lithium-ion batteries with multi-parameter updating. Electrochim Acta 2022;426:140760. [CrossRef]

[26] Sekhar JNC, Domathoti B, Santibanez Gonzalez EDR. Prediction of battery remaining useful life using machine learning algorithms. Sustainability 2023;15:15283. [CrossRef]

[27] Badik ŞT, Akar M. Machine learning classification models for the patients who have heart failure. Sigma J Eng Nat Sci 2024;42:235–244. [CrossRef]

[28] Das O. Prediction of the natural frequencies of various beams using regression machine learning models. Sigma J Eng Nat Sci 2023;41:302–321. [CrossRef]

[29] Pınar C. The machine learning approach for predicting the number of intensive care, intubated patients and death: The COVID-19 pandemic in Turkey. Sigma J Eng Nat Sci 2022;40:85–94.

[30] Chavan PP, Makhar K, Kulkarni M, Kulkarni SS. Flight Price Prediction for Enhanced Recommendations via Machine Learning Web Application. In: Proceedings of the 2024 2nd International Conference on Sustainable Computing and Smart Systems (ICSCSS); 2024. p. 1031–1037. [CrossRef]