Sigma Journal of Engineering and Natural Sciences, Vol. 39, No. 1, pp. 1-12, March, 2021
Yildiz Technical University Press, Istanbul, Turkey

METAHEURISTIC-BASED MACHİNE LEARNİNG MODELS FOR ACCURATE BATTERY REMAINING USEFUL LIFE PREDICTION IN RENEWABLE ENERGY STORAGE SYSTEM

Shravani Gandhi[1*], Ashlesha Chopane[2], Anushka Parkhi[3], Sakshi Deshmukh[4], Smita Kulkarni[5]

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. A precise Remaining Useful Life prediction optimizes battery consumption, safety, and efficiency while affecting energy storage devices' economic and environmental sustainability. 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. Each model is optimised for prediction accuracy through feature significance analysis, data preparation, and hyperparameter adjustment.

Comparative measures like RMSE, MAE, and R^2 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 R^2 score of 99.61. These findings show that ensemble approaches like Random Forest and XGBoost may accurately forecast Remaining Useful Life for battery management and repair. 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. Unlike previous studies, which often focused on single-model approaches, this research explores an ensemble of machine learning models, emphasizing optimization and feature importance for better prediction accuracy.

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 paper presents a rigorous, data-driven methodology that integrates many strategies for improving Remaining Useful Life forecasts, making it a valuable tool for real-world battery management applications.

Keywords: *Renewable Energy; Battery Remaining Useful Life; Machine Learning; Regression.*

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 Remaining Useful Life prediction, many studies rely on traditional algorithms that often lack the precision needed for real-world applications. For example, support vector machines 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.

[1] MIT Academy of Engineering, Department of E&TC Eng., Pune, India
[2] MIT Academy of Engineering, Department of E&TC Eng., Pune, India
[3] MIT Academy of Engineering, Department of E&TC Eng., Pune, India
[4] MIT Academy of Engineering, Department of E&TC Eng., Pune, India
[5] MIT Academy of Engineering, Department of E&TC Eng., Pune, India
**E-mail address: shravani.gandhi@mitaoe.ac.in, ashlesha.chopane@mitaoe.ac.in, anushka.parkhi@mitaoe.ac.in, deshmukh.sakshi@mitaoe.ac.in, sskulkarni@mitaoe.ac.in*
Orcid id: https://orcid.org/0009-0003-3701-8415, 0009-0004-8044-5372, 0009-0006-9236-9949, 0009-0005-1029-3657, 0000-0001-6378-1738

Lithium-ion batteries (LIBs), known for their high energy density and durability, are the primary choice for powering electric vehicles. Despite advancements, gaps remain in the research landscape. Many predictive models do not fully leverage modern machine learning techniques, particularly metaheuristic optimization strategies, which could enhance performance. Additionally, most studies inadequately address the real-time implementation of predictive models, limiting their applicability in dynamic environments where timely decision-making is essential. Remaining Useful Life predictions are vital for assessing battery safety, enhancing performance, and optimizing energy storage systems' efficiency and sustainability. Efficient battery management supports sustainable development goals, such as affordable clean energy and climate action [1], and promotes renewable energy adoption and innovation. Remaining Useful Life predictions, expressed in various metrics, provide early failure alerts essential for system prognostics and health management. Techniques for simulating complex system degradation include model-based and information-driven approaches. 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]. Machine learning plays a critical role in optimizing battery performance, as analyzed in a bibliometric study on ML methods for battery prediction [7]. Other studies have validated models like PNN, LSTM, and SVR in NASA's battery dataset to optimize real-world performance [8]. Moreover, [9] emphasizes the significance of SOH estimates in electric vehicles and highlights advancements in filtering techniques and co-estimation methods. Researchers explore the degradation factors of lead-acid batteries to enhance battery management and reliability, while a novel ensemble learning method demonstrates improved prediction accuracy for Remaining Useful Life compared to traditional approaches [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. Researchers have employed approaches such as hidden Markov models (HMM) and unscented Kalman filters (UKF) to address uncertainties and dynamic conditions in battery operation [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 findings suggest that integrating machine learning with model-based approaches can enhance RUL predictions and optimise battery usage, contributing to sustainability goals [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 Remaining Useful Life (RUL) of batteries, aiming to improve operational efficiency in renewable energy systems [27]. The study on predicting natural frequencies of beam structures showcases the effectiveness of machine learning models in engineering applications. This parallels the research on metaheuristic-based models for accurately predicting the Remaining Useful Life (RUL) of batteries in renewable energy systems [28]. The COVID-19 outbreak utilized machine learning for accurate predictions of patient numbers and deaths, paralleling research on battery Remaining Useful Life prediction [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. The study highlights the detailed extraction of critical battery performance features, including discharge times, voltage variations during charging and discharging, and cycle index. Emphasizing the importance of preprocessing steps like data standardization, feature selection, and handling missing values; these techniques ensure accurate Remaining Useful Life predictions by refining input data and enhancing model reliability. The study gives a thorough look at feature importance and error metrics,

showing that these models are more accurate at predicting Remaining Useful Life than traditional methods. 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 Remaining Useful Life 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. Furthermore, this research advances sustainable energy practices by providing reliable and accurate battery lifespan predictions, optimizing resource utilization, and extending battery life.

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 novelty of this work lies in its holistic approach to integrating advanced machine learning techniques into battery management systems, which enhances the reliability and sustainability of energy storage solutions. 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. For longitudinal research, the cycle index is essential since it makes it possible to correlate degradation patterns with certain cycle counts.

Discharge Time (s): The Discharge Time column records the duration, measured in seconds, required for the battery to discharge completely in each cycle. This temporal metric is crucial for assessing the rate at which the battery depletes its stored energy. 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.

Decrement 3.6-3.4V (s): This column measures how long it takes for the battery's voltage to drop from 3.6 volts to 3.4 volts during discharge, measured in seconds. 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 column records the maximum voltage observed during the discharge phase of each cycle, expressed in volts. The peak discharge voltage is indicative of the battery’s voltage response under load, reflecting its instantaneous power delivery capability. Monitoring this parameter helps understand the voltage's behavior and identify any anomalies that could indicate potential issues.

Min. Voltage Charg. (V): The Min. Voltage Charg. column captures the minimum voltage observed during the charging phase of each cycle. This metric is essential for understanding the lowest voltage level reached during recharging, which can provide insights into the charging efficiency and recovery characteristics of the battery. 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 duration at this voltage plateau reveals the battery's saturation behavior and its ability to accept charges close to its maximum capacity. This metric is essential for evaluating the stability and efficiency of the charging protocol.

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. This parameter, often denoted by the symbol τ (tau), reflects the dynamic response of the battery's electrochemical processes. It is critical to understand how quickly the current changes in response to a voltage alteration, influencing both charging and discharging profiles.

Charging time (s): This column records the total duration, measured in seconds, required to complete the charging process for each cycle. It encompasses the entire period from the initiation to the termination of charging, thus providing a holistic view of the battery's charge acceptance rate and efficiency. Variations in charging time can signal changes in the battery's capacity and internal resistance, making it a vital parameter for lifecycle assessment.
RUL (Remaining Useful Life): The Remaining Useful Life column contains estimations of the remaining life of the battery, based on its current performance characteristics. This predictive metric forecasts the expected operational lifespan of the battery before it reaches a predefined end-of-life criterion, such as a significant capacity loss or increased internal resistance. 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.

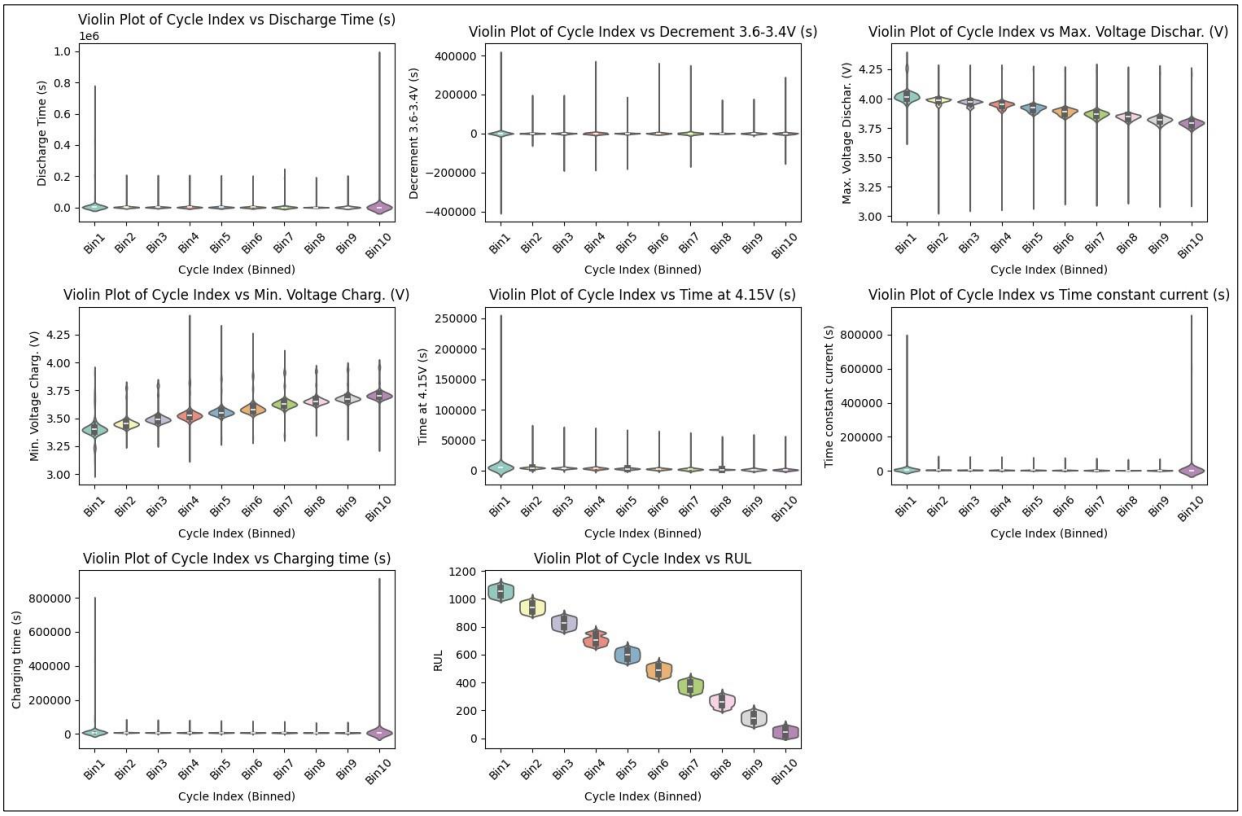

Figure 1. Violin Plots of Battery Performance Metrics Across Cycle Index Bins

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 first plot, which displays the discharge time (s), demonstrates a consistent distribution of discharge time across various cycle bins, albeit with a significant spread that highlights the variability in discharge times. 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 Time at 4.15V (s)

plot exhibits high variability and significant outliers, indicating inconsistent durations at this voltage level. 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. Similarly, the Charging time (s) plot shows that charging time is mostly concentrated around lower values, with significant outliers indicating variability in charging durations.

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. The output data is standardized and structured to facilitate accurate input into models like Random Forest and XGBoost, allowing them to predict Remaining Useful Life effectively across varied battery conditions. The dataset undergoes a structured testing process to validate model accuracy and ensure practical applicability. The steps include:

1. Splitting the dataset into training and testing sets, maintaining consistency in feature representation.
2. Performing k-fold cross-validation on the training data to prevent overfitting.
3. Evaluating models on the testing set using metrics like 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.

Heatmaps enhance the interpretability of statistical information by using colour gradients to represent data values, making them accessible even to non-experts. 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.

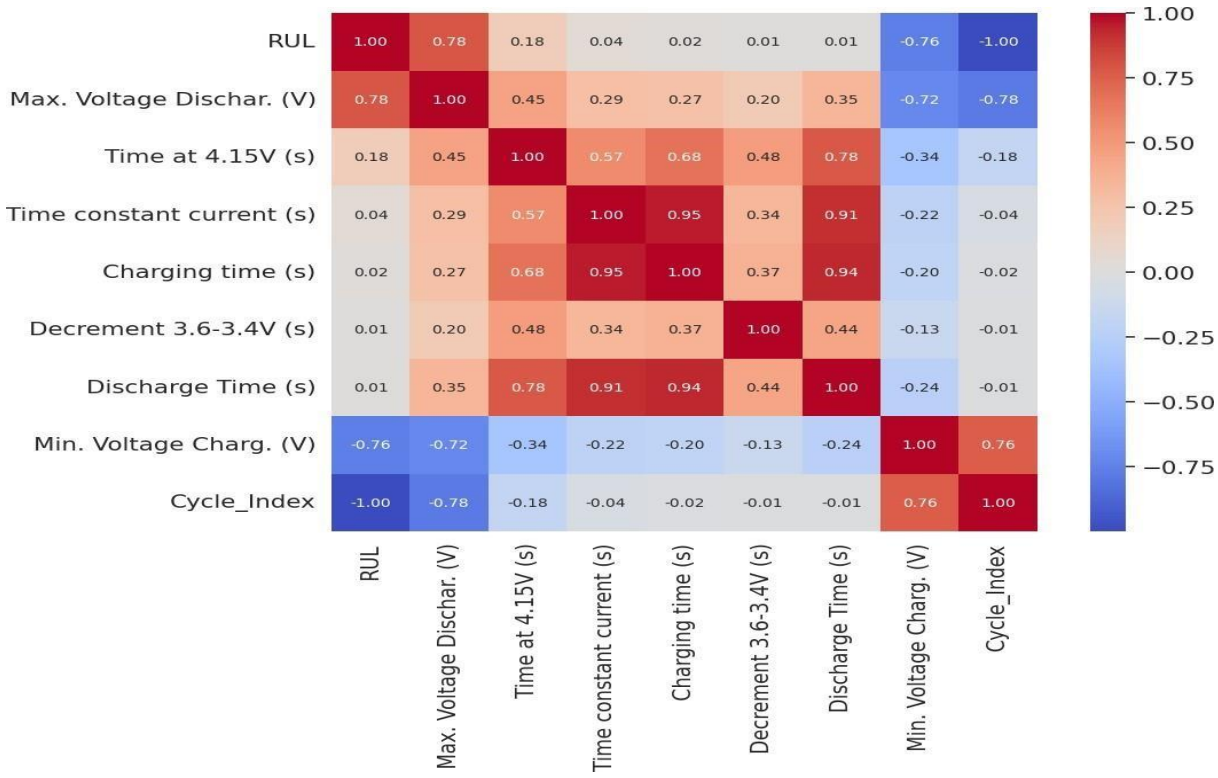

Figure 2. Correlation Heatmap of Battery Performance Metrics

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 correlation values are between -1 and 1, with red meaning there is a positive correlation and blue meaning there is a negative correlation. The heat map shows that there is a strong negative correlation (-1.00) between the Remaining Useful Life (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. Additionally, a strong positive correlation (0.76) exists between Remaining Useful Life and Minimum Voltage Charge, indicating a link between longer Remaining Useful Lifes and higher minimum charging voltages. The maximum voltage discharge also shows a moderately positive correlation (0.78) with Remaining Useful Life.

Time-based metrics such as constant current, charging time, and discharge time exhibit high correlations with each other, highlighting their interdependencies. The minimum voltage charge has a strong negative correlation with Remaining Useful Life and maximum voltage discharge, with values 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. Therefore, the heat map provides a comprehensive view of the connections between various battery performance metrics, aiding in battery management and future maintenance planning.

FRAMEWORK FOR REMAINING USEFUL LIFE PREDICTION

In the field of prognostics and health management (PHM), accurately predicting the Remaining Useful Life (RUL) of systems is essential. This paper introduces a novel framework integrating metaheuristic optimization techniques and machine learning models to enhance prediction. Metaheuristic algorithms optimize the parameters and structures of machine learning models, improving accuracy and robustness. This approach addresses limitations, such as overfitting and poor generalization in existing methodologies. Validated with a comprehensive dataset, the proposed framework demonstrates significant performance improvements, offering a more reliable tool for predictive maintenance. The diagram that illustrate the machine learning project pipeline and the proposed predictive model for Remaining Useful Life prediction are also given.

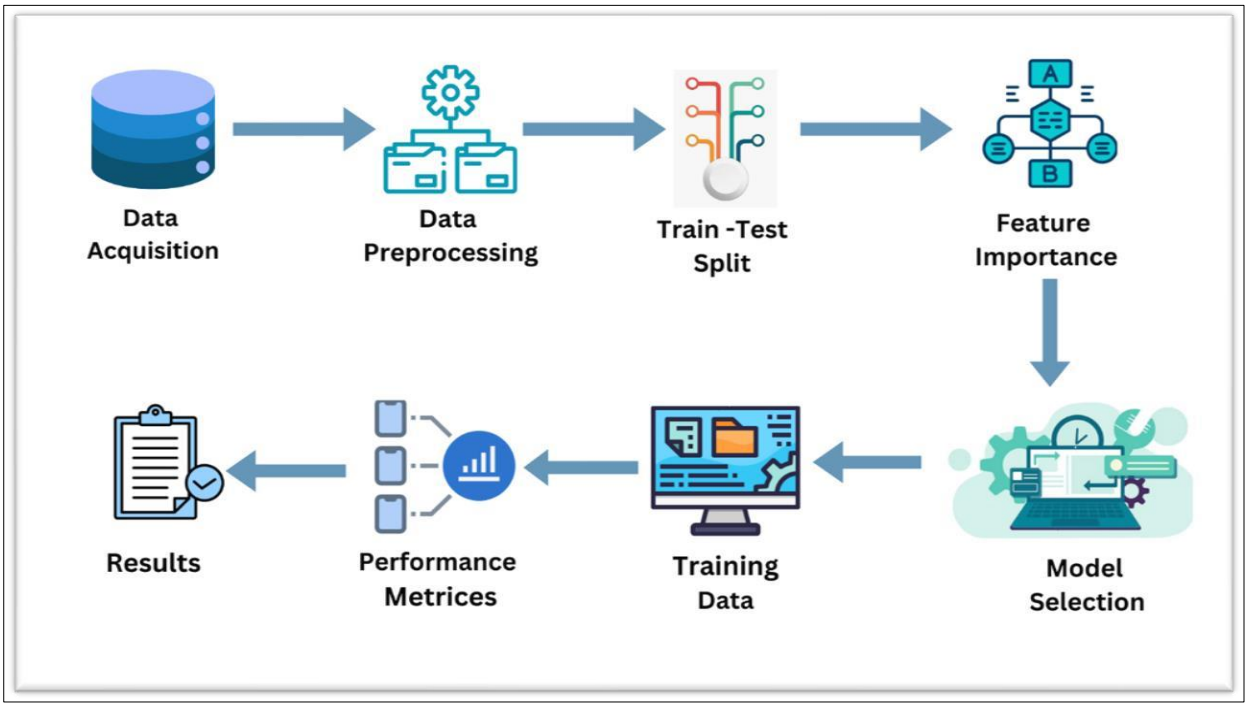

Figure 3. Machine Learning Project Pipeline

Figure 3 illustrates the typical workflow of a machine learning project. It begins with data acquisition, gathering raw data from various sources. Next, data preprocessing involves cleaning and preparing the data. 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

An outline of a process for predicting the Remaining Useful Life (RUL) using machine learning is shown in Figure 4. It starts with data collection and preprocessing, including data representation, standardization, and handling null values. Feature importance is determined using a correlation matrix, mutual information, and feature selection algorithms. The dataset is then split into training and testing sets. 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.

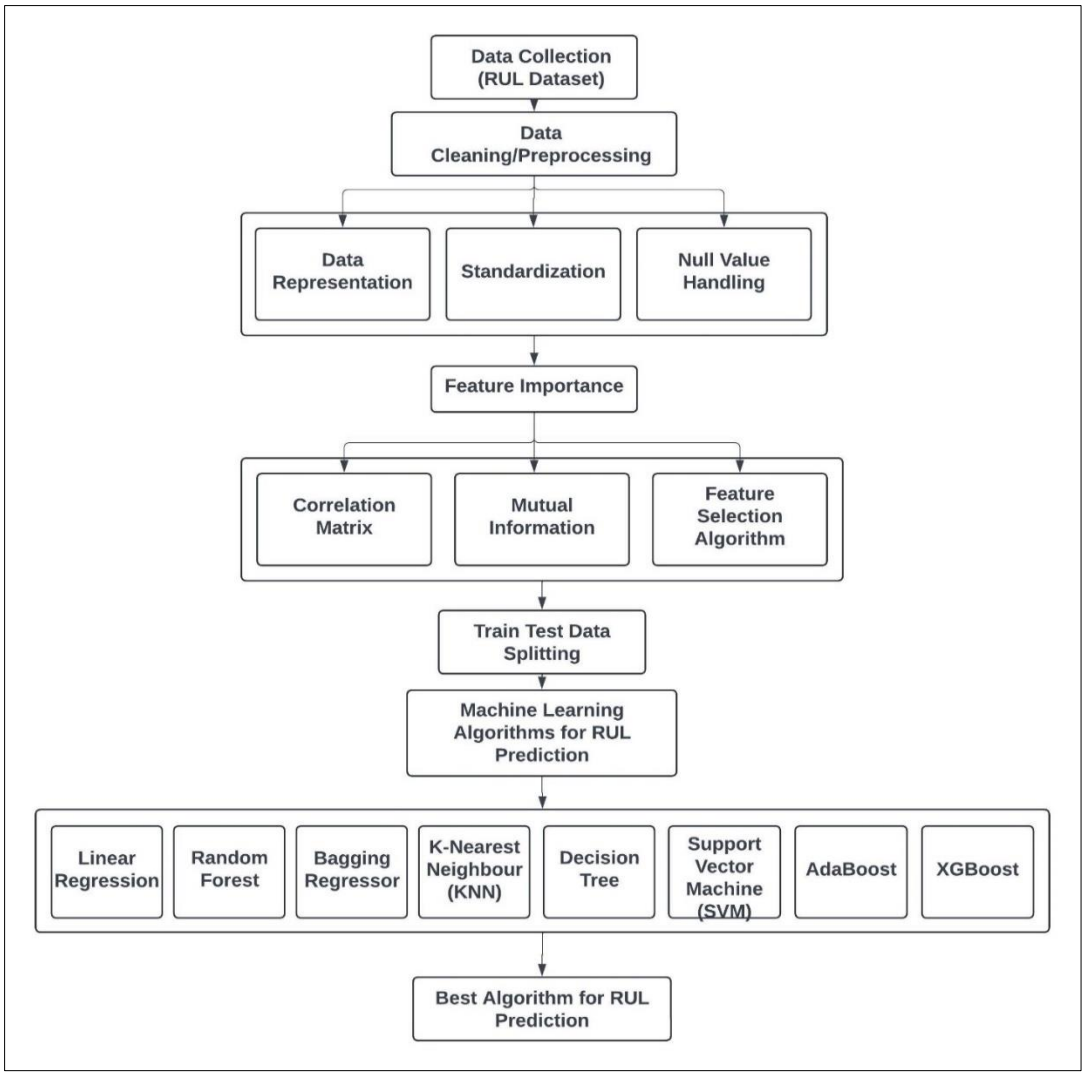

Figure 4. The proposed predictive model for RUL prediction

METHODOLOGY

The study tested metaheuristic-based models against traditional models like Linear Regression, SVM, etc to quantify improvements in predictive accuracy and error reduction. Unlike traditional models, metaheuristic techniques apply global optimization, capturing complex, non-linear degradation patterns in battery data more effectively. This study used statistical tests, including t-tests and ANOVA, to validate that observed gains, such as reduced RMSE and increased R^2, were statistically significant rather than random. These tests confirm the enhanced accuracy and robustness that metaheuristic approaches bring to RUL prediction.

1. 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 uncover 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 serves as a cornerstone in predictive analytics, delivering interpretable results that illuminate linear relationships within the dataset, guiding decisions on battery maintenance and replacement strategies effectively.

From Simple Linear Regression:

$$y = \beta 0 + \beta 1 x + \epsilon \tag{1}$$

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

$\beta 1$ $\beta 0$: weights

ϵ : error term

From Multiple Linear Regression:

$$y = \beta 0 + \beta 1 x1 + \beta 2 x2 \ \ldots + \beta p x p + \epsilon \tag{2}$$

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

2. SUPPORT VECTOR MACHİNE 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 (ϵ\epsilonϵ). This margin establishes a tolerance tube around the predicted function, accommodating acceptable deviations. 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 $K\ (x_i, x)$ the kernel function.

3. BOOSTİNG

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.

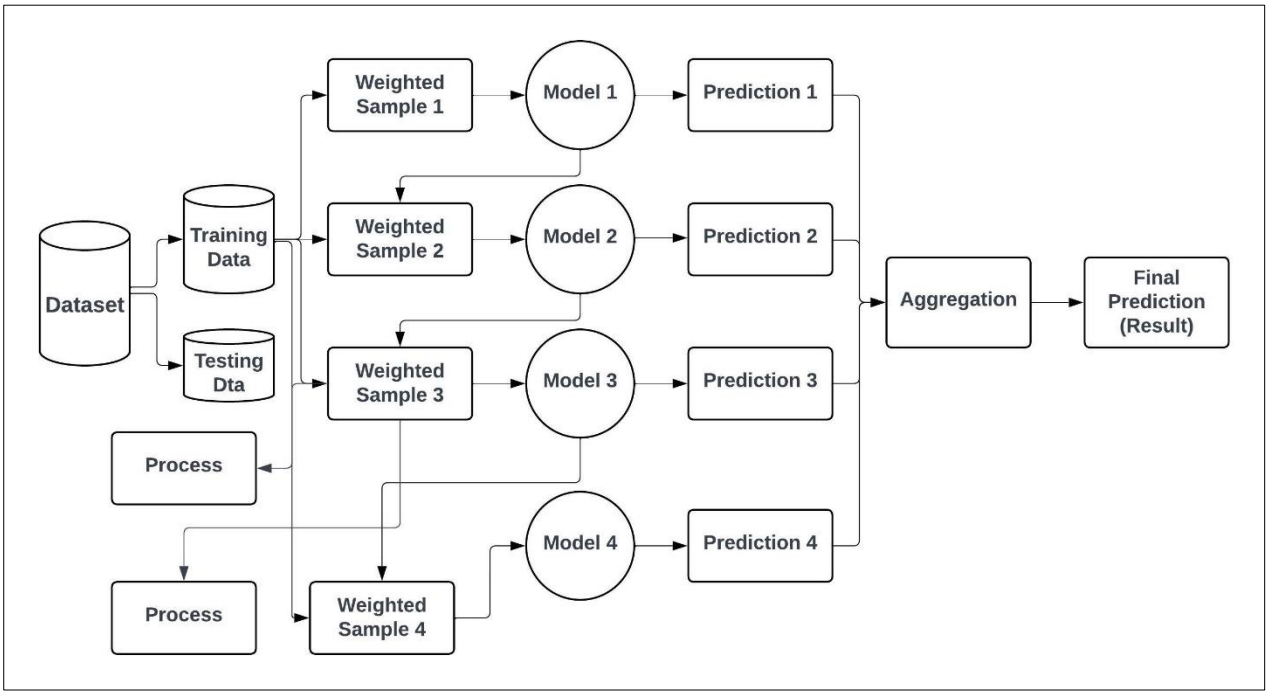

Figure 5. Boosting Ensemble Regressor Process Flow

Boosting, is a powerful ensemble learning technique used to improve the predictive performance of machine learning models is shown in Fig. 5. The dataset is initially divided into training and testing data. Multiple weighted samples are drawn from the training data, with the weights adjusted iteratively based on the performance of previous models. These weighted samples are then used to train individual models (Model 1, Model 2, Model 3, and Model 4). Each model generates its respective prediction (Prediction 1, Prediction 2, Prediction 3, and Prediction 4). These predictions are subsequently aggregated to produce a final prediction or result. During this iterative process, incorrectly predicted instances become more apparent, allowing subsequent models to focus on difficult cases, improving overall accuracy and robustness. The following are two important types of boosting:

3.1. 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 no (4):

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

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

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

3.2. XGBOOST ENSEMBLE REGRESSOR:

XGBoost excels at predicting RUL using features like Discharge Time, voltage dynamics, and cycle durations. It minimizes mean squared error by iteratively optimizing decision trees to refine predictions and handle residuals effectively. 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. With its ability to produce accurate continuous predictions, XGBoost stands out in forecasting RUL, crucial for optimizing battery lifecycle management and ensuring reliable operational planning based on comprehensive cycle data.

4. 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:

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

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:

$$H(S) = \sum_{i=1}^{n} p_i \log_2(p_i) \tag{7}$$

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

5. 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.

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

6. 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 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.

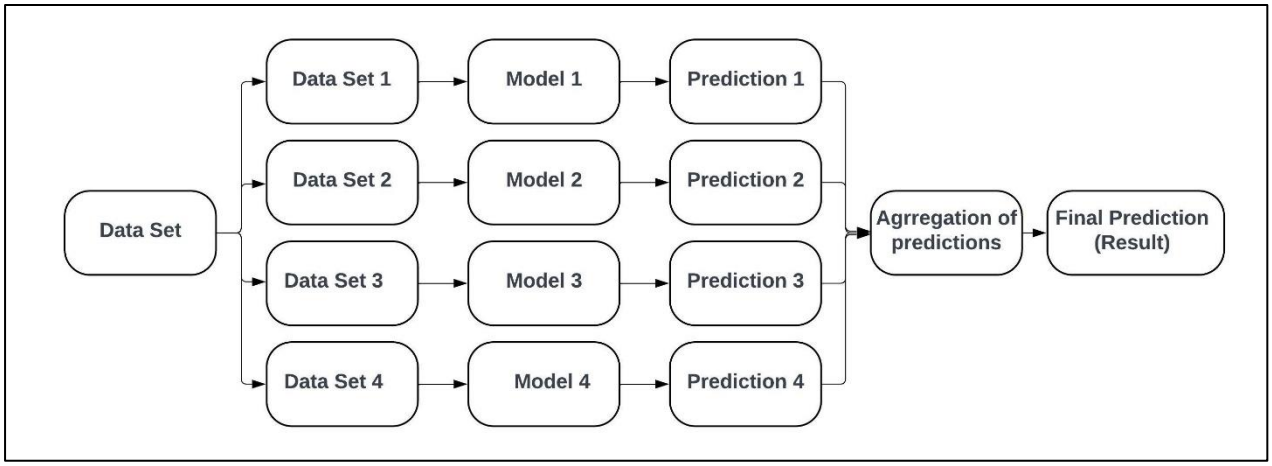

Figure 6. Bagging Regressor Workflow Diagram

Fig. 6 illustrates the bagging process, a robust ensemble learning technique in machine learning. It begins by generating multiple overlapping subsets from the original data set through random sampling with replacement. Each subset is used to train a separate model independently, resulting in diverse models (Model 1, Model 2, Model 3, and Model 4) tailored to different data distributions. Each model then produces its own prediction. These individual predictions are subsequently aggregated using methods like averaging (for regression tasks) or majority voting (for classification tasks) to form a final, consolidated prediction. Through this process of aggregation, the overall prediction accuracy and robustness are enhanced by reducing variance and mitigating overfitting.

7. RANDOM FOREST

Random Forest is an ensemble learning method known as bagging, employs multiple decision trees to enhance prediction accuracy and mitigate overfitting. It constructs various trees by randomly sampling the training data with replacement (bagging) and selecting random features at each split. For instance, in a dataset with battery performance metrics, Random Forest can predict the RUL by aggregating predictions from individual trees trained on different data subsets. This method leverages diverse combinations of features related to discharge times, voltage levels, and charging durations, resulting in a more robust and precise model than a single decision tree.

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

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.

MAJOR PERFORMANCE MATRICES

Major performance matrices in regression include Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and R-squared (R^2), each quantifying prediction accuracy, error magnitude, and model fit.

1. MEAN ABSOLUTE ERROR (MAE)

The Mean Absolute Error (MAE) is a metric used to assess prediction accuracy by averaging the absolute differences between predicted and observed values across the test sample. MAE would measure the average magnitude of errors in predicting RUL based on cycle attributes such as Discharge Time, voltage changes during discharge and charge, and various time durations. The ideal value of Mean Absolute Error (MAE) is typically close to zero.

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

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

2. MEAN SQUARE ERROR (MSE)

The Mean Squared Error (MSE), as shown in equation (11) is a metric that calculates the average of the squared differences between predicted and actual values in a dataset. MSE quantifies the average magnitude of squared errors when predicting RUL based on cycle attributes such as Discharge Time, voltage changes during discharge and charge, and various time durations. The idle value of Mean Squared Error (MSE) is typically considered to be 0.

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

3. ROOT MEAN SQUARED ERROR (RMSE)

The Root Mean Squared Error (RMSE), as shown in equation (12) is calculated as the square root of the Mean Squared Error (MSE), evaluates prediction accuracy in the battery cycle dataset by measuring the average magnitude of errors between predicted and actual RUL. mIt serves as a critical metric for assessing how closely models predict RUL values based on cycle characteristics like Discharge Time and voltage profiles. RMSE provides insights into the effectiveness of predictive models, essential for optimizing battery lifecycle management strategies. The idle value of Root Mean Squared Error (RMSE) is also 0.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{12}$$

4. R^2 SCORE

R^2 score, in equation (13) assesses how well a regression model predicts Remaining Useful Life (RUL) based on features like Discharge Time, voltage changes, and cycle durations. It quantifies the proportion of RUL variance explained by these variables. A higher R^2 score indicates a better fit, suggesting the model effectively captures variability in RUL. Conversely, a lower R^2 implies poorer model fit to the data. This statistical measure is crucial for evaluating the regression model's ability to accurately predict and understand the relationship between cycle attributes and battery longevity. The idle value of the R^2 (R-squared) score is 1.

(13)

$$R^2 = 1 - \frac{ss_r}{ss_t}$$

Where,

ss_r : Residual sum of squares

ss_t : Total sum of squares

RESULT AND DİSCUSSİON

This section presents the results and discussion of proposed framework for Remaining Useful Life prediction, employing rigorous evaluation metrics including RMSE and R^2 scores. The RMSE assesses the precision of predictions by quantifying the deviation between predicted and actual Remaining Useful Life values. At the same time, the R^2 score evaluates the model's goodness-of-fit relative to established benchmarks. 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. The output data is standardized and structured to facilitate accurate input into models like Random Forest and XGBoost, enabling them to predict Remaining Useful Life effectively across varied battery conditions The analysis reveals substantial improvements in prediction accuracy and model robustness across diverse datasets compared to conventional methods. The discussion elaborates on these findings, emphasizing implications for optimizing maintenance strategies and suggesting avenues for further research to enhance predictive performance in practical applications.The testing results provide essential insights into the models best suited for real-world applications in battery management, highlighting factors like accuracy, speed, and robustness. By following this structured testing process, the dataset is rigorously evaluated, ensuring that the machine learning models can accurately predict the Remaining Useful Life of batteries under practical conditions. These insights confirm that models like Random Forest and XGBoost demonstrate high reliability and efficiency, making them ideal choices for accurate battery life estimation in dynamic, real-world environments.

Table 1. Performance Comparison of Machine Learning Algorithms

Algorithm Name	Testing Score	Training Score	RMSE	R2_Score	Execution Time
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

Table 1 compares various regression algorithms based on performance metrics, including Training Score, Testing Score, RMSE (Root Mean Square Error), R^2 Score, and Execution Time. Linear Regression exhibits

moderate performance with a Training Score of 79.73 and a Testing Score of 78.14, resulting in the highest RMSE of 148.52 and an R^2 Score of 80.44. Despite its lower accuracy, it boasts the fastest Execution Time of 0.038 seconds, making it efficient in terms of computational speed. In contrast, Random Forest demonstrates outstanding performance, with Training and Testing Scores of 99.92 and 99.64, respectively. It achieves a very low RMSE of 23.26 and a near-perfect R^2 Score of 99.68, though it requires a longer Execution Time of 8.319 seconds. The Bagging Regressor also performs exceptionally well, with Training and Testing Scores of 99.88 and 99.42, an RMSE of 24.05, and an R^2 Score of 99.62. It balances accuracy and speed, with an Execution Time of 0.774 seconds. KNN (K-Nearest Neighbors) achieves high Training and Testing Scores of 99.67 and 99.08, respectively, an RMSE of 30.34, and an R^2 Score of 98.70, while maintaining a relatively quick Execution Time of 0.170 seconds.

The Decision Tree algorithm performs excellently with Training and Testing Scores of 99.99 and 98.85, respectively, an RMSE of 33.99, and an R^2 Score of 99.33, with an Execution Time of 0.245 seconds. SVM (Support Vector Machine), although achieving a good Training Score of 90.09 and a Testing Score of 89.46, has a higher RMSE of 103.14 and a lower R^2 Score of 89.70, and the longest Execution Time of 20.086 seconds, indicating it may not be as efficient for this dataset. AdaBoost provides a balanced performance with Training and Testing Scores of 94.40 and 94.03, an RMSE of 77.62, an R^2 Score of 93.94, and a moderate Execution Time of 2.178 seconds. Finally, XGBoost excels with Training and Testing Scores of 99.95 and 99.57, respectively, the lowest RMSE of 20.64, and an R^2 Score of 99.61, though it requires an Execution Time of 6.871 seconds. Overall, the table highlights the trade-offs between accuracy and execution time across different regression algorithms, with Random Forest and XGBoost emerging as the top performers regarding predictive accuracy. The best-performing models, such as Random Forest and XGBoost, are identified based on their high prediction accuracy and minimal error rates on the testing set. Testing results suggest that these models are well-suited for real-world battery management applications, considering their accuracy, speed, and robustness. By rigorously testing the dataset and evaluating performance, these models provide reliable RUL predictions, essential for proactive battery maintenance and resource optimization in practical scenarios.

Figure 7. Training Scores of Various ML Algorithms for RUL Prediction

Figure 7 depicts the "Training Score" for various machine learning algorithms, with the data points sorted in ascending order of their training scores. The algorithms included are Linear Regression, SVM, AdaBoost, KNN, Bagging Regressor, Random Forest, XGBoost, and Decision Tree. Each algorithm is represented along the x-axis in the plot, while their corresponding training scores are plotted along the yaxis. A smooth line curve is generated using cubic spline interpolation to provide a clearer trend visualization. The red markers indicate the actual data points. The graph reveals a steady increase in the training score from Linear Regression, which has the lowest score, to Decision Tree, which has the highest score. Notably, the training scores for most algorithms, especially from KNN onwards, are very close to 100, indicating very high training performance. This suggests that these algorithms fit the training data extremely well, which is often an indication of potential overfitting, especially in the absence of corresponding test scores. The SVM and Linear Regression algorithms, on the other hand, show comparatively lower training scores, suggesting a lesser degree of fitting to the training data compared to the other algorithms.

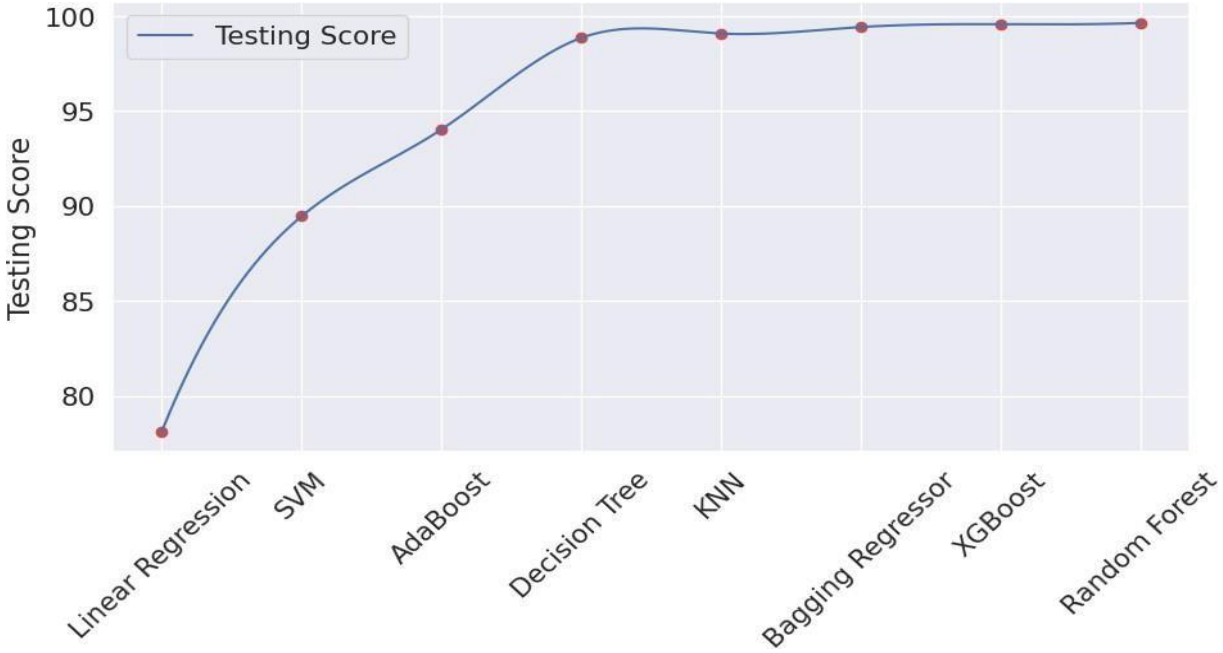

Figure 8. Testing Score of Various ML Algorithms for RUL Prediction

Figure 8 presents the "Testing Score" for various machine learning algorithms, with data points sorted in ascending order of their testing scores. The algorithms included in this plot are Linear Regression, SVM, AdaBoost, Decision Tree, KNN, Bagging Regressor, XGBoost, and Random Forest. In this plot, each algorithm is represented along the x-axis, while their corresponding testing scores are plotted along the y-axis. A smooth line curve is generated using cubic spline interpolation to provide a clearer visualization of the trend. The red markers indicate the actual data points. The graph shows a clear trend where the testing scores increase from Linear Regression, which has the lowest testing score, to Random Forest, which has the highest testing score. The increase is steep from Linear Regression to SVM, followed by a steady climb. From KNN onwards, the testing scores are very close to 100, indicating that these algorithms perform exceptionally well on the test data. This suggests that these algorithms generalize well to new data. However, given the high scores, it's essential to ensure these results are not due to overfitting, especially if training scores are similarly high. The Linear Regression and SVM models, with relatively lower testing scores, indicate a more moderate fit to the data compared to the other models, potentially offering better generalization in some scenarios.

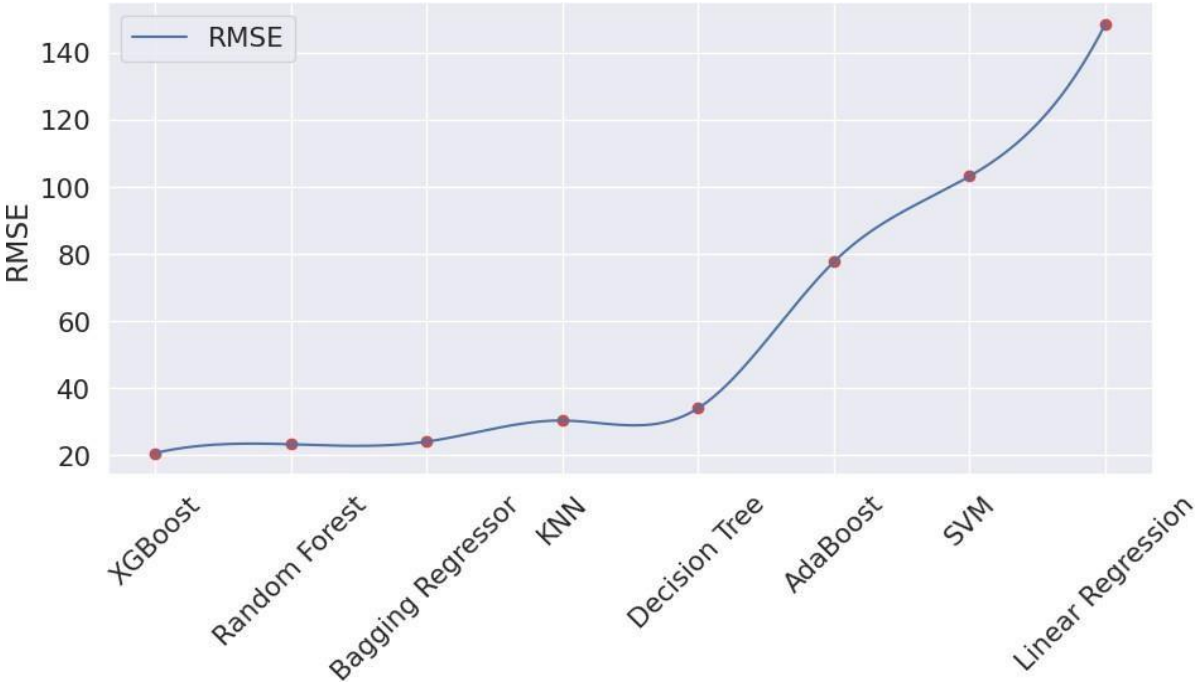

Figure 9. RMSE of Various ML Algorithms for RUL Prediction

The provided Figure 9 is a smooth line plot that illustrates the Root Mean Square Error (RMSE) values for various regression algorithms, arranged in ascending order of their RMSE values. The x-axis lists the algorithms, specifically XGBoost, Random Forest, Bagging Regressor, KNN, Decision Tree, AdaBoost, SVM, and Linear Regression, with the labels rotated for better readability. The y-axis displays the RMSE values, ranging from 20 to 150. The smooth curve, created using spline interpolation, connects these values, with red markers indicating the actual data points. This visualization effectively highlights the performance of each algorithm, where a lower RMSE value signifies better predictive accuracy. The title "RMSE Values for Different Regression Algorithms (Ascending Order)" succinctly encapsulates the graph's purpose, providing a clear comparison of the algorithms based on their RMSE metrics.

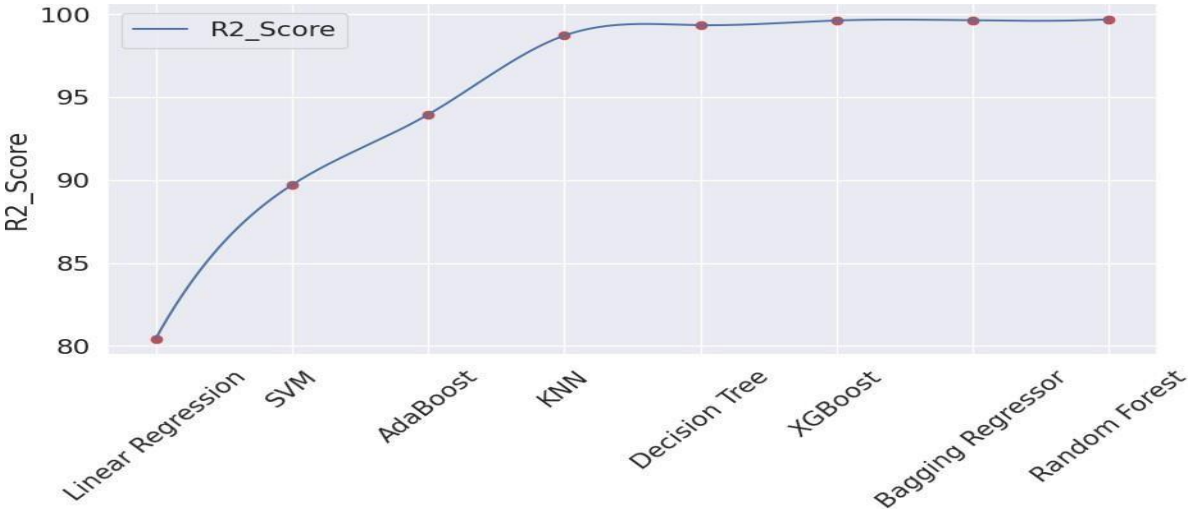

Figure 10. R^2 Score of Various ML Algorithms for RUL Prediction

Figure 10 shows a smooth line plot depicting the R^2 scores for various regression algorithms, arranged in ascending order. The x-axis represents the different algorithms, specifically Linear Regression, SVM, AdaBoost, KNN, Decision Tree, XGBoost, Bagging Regressor, and Random Forest, with the labels rotated for better readability. The y-axis shows the R^2 scores, ranging from 80 to 100. The smooth curve, generated using spline interpolation, connects these values, with red markers indicating the actual data points. This visualization highlights the performance of each algorithm, where a higher R^2 score indicates better predictive accuracy. The plot title, "R^2 Scores for Different Regression Algorithms (Ascending Order)," succinctly captures the graph's purpose, providing a clear comparison of the algorithms based on their R^2 metrics. This plot effectively conveys how each algorithm performs in terms of explaining the variability of the response variable, with higher values denoting superior performance.

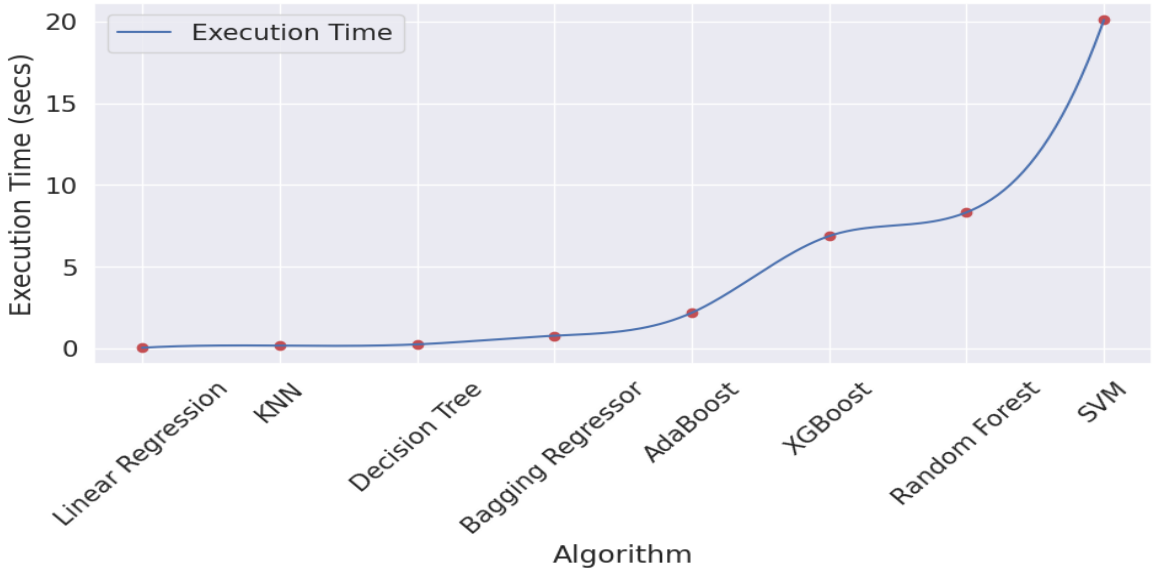

Figure 11. Execution Time of Various ML Algorithms for RUL Prediction

Figure 11 is a smooth line plot that illustrates the execution times for various regression algorithms, sorted in ascending order of their execution times. The x-axis features the algorithms, specifically Linear Regression, KNN, Decision Tree, Bagging Regressor, AdaBoost, XGBoost, Random Forest, and SVM, with labels rotated for improved readability. The y-axis quantifies the execution times, ranging from 0 to 20. The plot uses spline interpolation to generate a smooth curve connecting the data points, which are highlighted with red markers. This visualization effectively demonstrates the computational efficiency of each algorithm, where shorter execution times indicate faster performance. The plot reveals that Linear Regression and KNN are the quickest, while SVM has the longest execution time. By comparing these algorithms, the graph provides valuable insights into their practical applicability, especially in scenarios involving large datasets or time-sensitive tasks. The title "Execution Time" succinctly captures the essence of the graph, emphasizing the focus on the time each algorithm takes to execute.

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.

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

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. Additionally, Table 1 provides a broad selection of models with varying execution times, from very fast algorithms like KNN and SVM to more computationally intensive ones like XGBoost. This flexibility allows researchers to choose models based on their specific requirements, whether for speed or accuracy. 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. Metaheuristic optimization in model training improves computational efficiency, making these models suitable for deployment in time-sensitive applications. These optimizations are critical for battery management systems that need real-time analysis because they enable quick, accurate forecasts while controlling computing loads.

CONCLUSİON

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. By enhancing battery lifespan and minimizing maintenance costs, reliable Remaining Useful Life predictions contribute to both economic efficiency and environmental sustainability. The study employed various machine learning models, such as Random Forest, Decision Tree, Linear Regression, K-Nearest Neighbor, Support Vector Machine, Adaptive Boosting (AdaBoost), and Extreme Gradient Boosting (XGBoost). Metaheuristic optimization techniques were integrated with these models to refine their parameters and improve accuracy, thus enhancing predictive performance. The Hawaii Natural Energy Institute (HNEI) provided key features from a battery lifecycle dataset for 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. These results emphasize the effectiveness of advanced machine learning in improving RUL prediction. Employing a metaheuristic optimization approach combined with multiple machine learning models, particularly ensemble methods, which achieved substantial improvements in predictive accuracy. These improvements align with broader goals of sustainability and climate resilience, making reliable Remaining Useful Life prediction vital for the long-term management of renewable energy resources.

REFERENCES

[1] Shrihari Sankarasubramanian, Balaji Krishnamurthy, A capacity fade model for lithium-ion batteries including diffusion and kinetics, Electrochimica Acta, Volume 70, 2012, Pages 248254, ISSN 0013-4686.

[2] Yujie Wang, Duo Yang, Xu Zhang, Zonghai Chen, Probability based remaining capacity estimation using data-driven and neural network model, Journal of Power Sources, Volume 315, 2016, Pages 199-208, ISSN 0378-7753.

[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(3):124.

[4] Yujie Wang, Rui Pan, Duo Yang, Xiaopeng Tang, Zonghai Chen, Remaining Useful Life Prediction of Lithium-ion Battery Based on Discrete Wavelet Transform, Energy Procedia, Volume 105, 2017, Pages 2053-2058, ISSN 1876-6102. 15.

[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 Jan 23;8:20786-98.

[6] Chao Hu, Gaurav Jain, Puqiang Zhang, Craig Schmidt, Parthasarathy Gomadam, Tom Gorka, Data-driven method based on particle swarm optimization and k-nearest neighbor regression for estimating capacity of lithium-ion battery, Applied Energy, Volume 129, 2014, Pages 4955, ISSN 0306-2619.

[7] Kamul, Azure & Hananda, Nico & Harito, Christian & Djuana, Endang & Elwirehardja, Gregorius & Pardamean, Bens & Gunawan, Fergyanto & Budiman, Arief & Asrol, Muhammad & Redi, Anak Agung Ngurah Perwira & Pasang, Tim. (2022). Machine Learning Application in Battery Prediction: A Systematic Literature Review and Bibliometric Study.

[8] P. Venugopal et al., "Analysis of Optimal Machine Learning Approach for Battery Life Estimation of Li-Ion Cell," in IEEE Access, vol. 9, pp. 159616-159626, 2021, doi: 10.1109/ACCESS.2021.3130994.

[9] X. Feng et al., "Online State-of-Health Estimation for Li-Ion Battery Using Partial Charging Segment Based on Support Vector Machine," in IEEE Transactions on Vehicular Technology, vol. 68, no. 9, pp. 8583-8592, Sept. 2019, doi: 10.1109/TVT.2019.2927120.

[10] Robert R. Richardson, Michael A. Osborne, David A. Howey, Battery health prediction under generalized conditions using a Gaussian process transition model, Journal of Energy Storage, Volume 23, 2019, Pages 320-328, ISSN 2352-152X.

[11] Jiaju Wu, Linggang Kong, Zheng Cheng, Yonghui Yang, Hongfu Zuo, RUL Prediction for Lithium Batteries Using a Novel Ensemble Learning Method, Energy Reports, Volume 8, Supplement 12, 2022, Pages 313-326, ISSN 2352-4847.

[12] Shaheer Ansari, Afida Ayob, M.S. Hossain Lipu, Aini Hussain, Mohamad Hanif Md Saad, Remaining useful life prediction for lithium-ion battery storage system: A comprehensive review of methods, key factors, issues and future outlook, Energy Reports, Volume 8, 2022, Pages 12153-12185, ISSN 2352-4847.

[13] A. Wang, H. Chen, P. Jin, J. Huang, D. Feng and M. Zheng, "RUL Estimation of Lithium-Ion Power Battery Based on DEKF Algorithm," 2019 14th IEEE Conference on Industrial Electronics and Applications (ICIEA), Xi'an, China, 2019, pp. 1851-1856.

[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(3):124.

[15] Robert R. Richardson, Michael A. Osborne, David A. Howey, Battery health prediction under generalized conditions using a Gaussian process transition model, Journal of Energy Storag, Volume 23, 2019, Pages 320-328, ISSN 2352-152X.

[16] Que, Zijun & Xu, Zhengguo. (2019). A Data-Driven Health Prognostics Approach for Steam Turbines Based on Xgboost and DTW. IEEE Access. PP. 1-1. 10.1109/ACCESS.2019.2927488.

[17] M.S. Hossain Lipu, M.A. Hannan, Aini Hussain, M.M. Hoque, Pin J. Ker, M.H.M. Saad, Afida Ayob, A review of state of health and remaining useful life estimation methods for lithium-ion battery in electric vehicles: Challenges and recommendations, Journal of Cleaner Production, Volume 205, 2018, Pages 115-133, ISSN 0959-6526

[18] B. Xu, A. Oudalov, A. Ulbig, G. Andersson and D. S. Kirschen, "Modeling of Lithium-Ion Battery Degradation for Cell Life Assessment," in IEEE Transactions on Smart Grid, vol. 9, no. 2, pp. 1131-1140, March 2018, doi: 10.1109/TSG.2016.2578950.

[19] Magdy Abdullah Eissa, Pingen Chen, Machine Learning-based Electric Vehicle Battery State of Charge Prediction and Driving Range Estimation for Rural Applications, IFACPapersOnLine, Volume 56, Issue 3

[20] Ning Wang, Yelin Lyu, Yongjia Zhou, Jie Luan, Yuan Li, Chaojun Zheng, A hybrid framework for remaining driving range prediction of electric taxis, Sustainable Energy Technologies and Assessments, Volume 67,2024, 103832,ISSN 2213-1388

[21] Sekhar JNC, Domathoti B, Santibanez Gonzalez EDR. Prediction of Battery Remaining Useful Life Using Machine Learning Algorithms. Sustainability. 2023; 15(21):15283.

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

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

[24] Maheshwari, A. and Nageswari, S., 2022. 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 (pp. 139-151). Singapore: Springer Nature Singapore.

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

[26] Chandra Sekhar, J N & Domathoti, Bullarao & Santibanez Gonzalez, Ernesto. (2023). Prediction of Battery Remaining Useful Life Using Machine Learning Algorithms. Sustainability. 15. 15283. 10.3390/su15211528.

[27] Badik ŞT, Akar M. Machine learning classification models for the patients who have heart failure. Sigma J Eng Nat Sci 2024;42(1):235−244.

[28] Das O. Prediction of the natural frequencies of various beams using regression machine learning models. Sigma J Eng Nat Sci 2023;41(2):302−321.

[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(1):85–94.

[30] P. P. Chavan, K. Makhar, M. Kulkarni and S. S. Kulkarni, "Flight Price Prediction for Enhanced Recommendations via Machine Learning Web Application," 2024 2nd International Conference on Sustainable Computing and Smart Systems (ICSCSS), Coimbatore, India, 2024, pp. 1031-1037.