Journal of Electronics, Electromedical Engineering, and Medical Informatics
Multidisciplinary: Rapid Review: Open Access Journal **Vol. 7, No. 2, April 2025, pp: 380-390; eISSN: 2656-8632**

RESEARCH ARTICLE **OPEN ACCESS**

Manuscript received November 24, 2024; Revised February 10, 2025; Accepted March 1, 2025; date of publication March 25, 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v7i2.635

How to cite: S.Sathishkumar, R.Yogesh Rajkumar, "Energy Efficient Battery Optimization Model (EE-BOM) using Machine Learning Algorithms and Harris Hawks Optimization", Journal of Electronics, Electromedical Engineering, and Medical Informatics, vol. 7, no. 2, pp. 380-390, April 2025.

Energy Efficient Battery Optimization Model (EE-BOM) using Machine Learning Algorithms and Harris Hawks Optimization

S. Sathishkumar [1], R. Yogesh Rajkumar [2]

[1] Department of Computer Science and Engineering, Bharath Institute of Higher Education and Research, Chennai, Tamilnadu, India.
[2] Department of Information Technology, Bharath Institute of Higher Education and Research, Chennai, Tamilnadu, India.

Corresponding author: S. Sathishkumar (e-mail: skumarscse@gmail.com)

ABSTRACT Electric vehicles (EVs) are gaining popularity because of their cheap running costs and positive environmental impacts. However, EVs' limited battery life is one of their biggest drawbacks. Accurately controlling lithium-ion battery (LiB) capacity improves energy storage systems' economic viability, especially in large-scale applications. Long-term cost reductions are achieved by replacing or maintaining LiBs less often. This paper presents the Energy Efficient Battery Optimisation Model (EE-BOM), a novel model for early battery life detection. The information used in this research comes from the Hawaii Natural Energy Institute and comprises 14 different batteries that underwent over 1000 cycles in a controlled setting. Feature selection follows data collection and preparation with data normalisation in a multi-step procedure. The XGBoost Approach, which combines Artificial Neural Networks (ANN) with Harris Hawk Optimisation (HHO), is also used for early RUL prediction. Feature significance analysis makes it simpler to identify critical elements influencing battery lifetime and health. The evaluations are carried out based on Mean Squared Error (MSE), Mean Absolute Error (MAE), Root Mean Squared Error (RMSE) and R-Squared rate. The proposed model attains 23.68%, 36.41% and 24.67% of average MSE, MAE and RMSE, respectively, which is minimal than other compared works. Outlier reduction improves model accuracy, and statistical analyses show no missing or redundant data. Notably, with almost flawless predictions, XGBoost proved to be the most successful algorithm. This study emphasises how important RUL prediction is for improving battery lifetime management, especially in applications like electric cars, guaranteeing the best possible use of resources, economic viability, and environmental sustainability in the future.

INDEX TERMS Energy Efficient, Battery Optimization, Machine Learning, HHO Optimization, Battery Lifetime, XGBoost.

I. INTRODUCTION

The majority of lithium-ion (Li-ion) batteries used in electric vehicles (EVs) are selected due to their remarkable qualities, such as high energy density, lack of memory effect, long lifetime, and adaptability in terms of charging and discharging [1]. Notwithstanding these benefits, the automobile sector still has to contend with unpredictable renewable energy supply chains, changing weather patterns, and rising air pollution from car emissions [2]. EV batteries' energy storage provides a viable way to address environmental uncertainties and concerns. The development and broad use of EVs with improved range, safety, and dependability are essential to decarbonising the transportation industry. However, there are drawbacks to using Li-ion batteries, including as capacity loss, environmental effects, and difficulties managing end-of-life [3]. RUL, which stands for the predicted time or use before a component, device, or system is expected to fail or no longer satisfy its operational characteristics, is a key term in predictive maintenance and reliability engineering [4]. Machine learning techniques are used to forecast RUL in the context of EV batteries, taking into account a number of variables.

Given the impact of cycling, deterioration, and environmental factors, predicting the remaining usable life (RUL) of lithium-ion (Li-ion) batteries is crucial to guaranteeing the dependable and effective functioning of power systems. This deterioration may result in power outages, safety risks, and financial losses by lowering capacity, raising internal resistance, and raising the chance of failure [5]. Power system operators may take preventative

action to avoid catastrophic failures, optimise maintenance schedules, and enhance overall system dependability by accurately determining the RUL of Li-ion batteries [6]. RUL projections help utilities use predictive maintenance techniques, such planning battery replacements before they run out of life, lowering the chance of unplanned outages, and enhancing grid stability [7]. Significant improvements in battery manufacture, use, and optimisation are possible when Li-ion battery lifespan predictions are made using early-cycle data. Manufacturers are able to classify new cells according to their expected lifetime, verify innovative production techniques, and accelerate cell development. Additionally, end customers may choose how long their batteries will last. Despite regulated operating conditions, nonlinear deterioration and unpredictability make it difficult to forecast battery lifespan reliably [8].

An important difficulty is that an EV battery's capacity usually drops by roughly 10% after about 6.5 years of continuous use [9]. Given that Li-ion batteries gradually lose capacity during cycles of charging and discharging, predicting RUL and tracking capacity deterioration are challenging problems [10]. Battery management systems are responsible for these responsibilities (BMSs). It is crucial to accurately predict the complex and non-linear course of battery capacity loss. When it comes to forecasting EV battery life, machine learning (ML) provides significant benefits. This helps owners plan their trips more effectively and helps manufacturers create batteries with longer lifespans and better charging techniques [11] [12].

ML approaches are very helpful in tackling the technical problems associated with battery deterioration because of the non-linear and complex elements that impact battery performance. By overcoming time constraints and scalability issues, machine learning algorithms provide accurate, non-invasive solutions. Infrastructure for transport must be electrified in order to meet the dual demands of affordable mobility and sustainable energy. In order to assist EV users and producers alike and promote sustainable development worldwide, this research aims to create a reliable and precise approach for predicting EV battery life. This study's main goal is to predict the RUL of lithium-ion (Li-ion) batteries, which is a crucial undertaking with important practical implications. For companies that significantly depend on Li-ion batteries, predicting RUL is essential since it enables preemptive maintenance plans and effective resource allocation. The dataset includes crucial characteristics like the cycle index, discharge period, and maximum voltage discharge in order to do this. The development of a strong and accurate prediction model incorporating these qualities is made possible by the goal variable, RUL, which represents the battery's remaining operational lifetime. This study focusses on showing how artificial intelligence (AI) may significantly increase the precision and effectiveness of electric car battery diagnostics while also showcasing real-world uses for these technologies. The use of these technologies enhances battery performance management, prolongs battery life, lowers maintenance costs, and boosts overall operating efficiency—all of which support the sustainable growth of electric cars Considering those things, the contribution of the proposed model is given below.

1. Obtaining data from benchmark dataset and perform data pre-processing
2. Integrating HHO with ANN for parametric optimization based on filters for enhancing the result efficacy of model.
3. In order to improve the CNN-XGBoost model's performance for battery RUL prediction, a new set of features is proposed by concatenating features that have been retrieved from the CNN model with another set of features that were obtained from measured data to feed with XGBoost.
4. Results are evaluated using evaluation metrics.

The remainder of this work is organized as follows, Section 2 deliberates the related works in the domain of RUL predictions and performance efficacy. The working procedure of proposed model is presented in Section 3 with system flow. The results and discussions are given in Section 4. Finally, section 5 completes the paper with conclusions and future work.

II. RELATED WORKS

The difficult problem of forecasting the RUL of lithium-ion batteries is taken on by J.-H. Chou et al. [13]. They suggest a hybrid approach based on transfer learning that combines models of bidirectional long short-term memory with attention mechanisms, support vector regression, and empirical mode decomposition. With relative error levels of 6.96%, 0.6%, and 6.25% for various charging policy target batteries, this method dramatically improved RUL forecast accuracy, especially for batteries with higher cycle numbers. The difficulties in forecasting battery capacity for electric vehicles (EVs) are discussed by J. Zhao et al. [14]. They create a two-step noise reduction technique and use a stacking ensemble learning strategy using feature-based machine learning on a dataset that includes 420 cells and 9 battery packs. In the complex setting of EV battery systems, The work helps to provide predictions that are both accurate and physically consistent.

Scholars often use DDB models to study a battery's degrading performance in order to estimate its state-of-health (SOH) [15], end-of-life (EOL) [16], SOC [17], and RUL [18]. To ensure safe and effective operation, the battery management system (BMS) must accurately estimate the battery's RUL and EOL under various operating situations. This forecast helps to maximise longevity, ensure optimum battery use, and prevent unplanned failures. In order to estimate battery health variations and accurately forecast EOL, a DDB framework is provided in [19] that use an automated feature selection to create customised inputs for a Gaussian Process Regression (GPR) model. This methodology's feature selection process prioritises elements that have a major influence on battery deterioration and exhibits adaptability in response to a variety of inputs.

In order to improve SOH estimation using the NASA dataset, a hybrid strategy including variation mode decomposition, multi-kernel support vector regression, and the sparrow search algorithm is described in [20]. However,

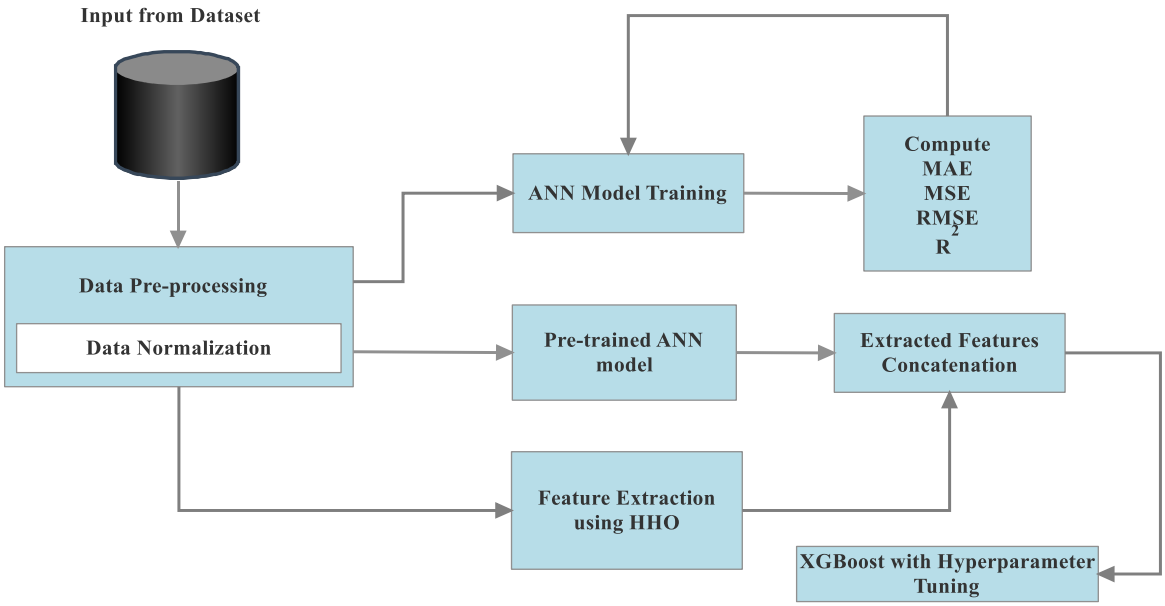

FIGURE 1. Workflow of RUL Prediction model

the proposed methodology is incompatible with RUL prediction. In [21], an analogous circuit model of a battery is evaluated on 65 batteries in order to extract physics-informed characteristics and use unsupervised learning to propose a combination detection and prediction model. The proposed model has a life prediction performance RMSE of 53.56% and detects deterioration stages with above 90% accuracy. A moving window-based approach for in-situ battery life categorisation and prediction using machine learning methods is provided in [22]. This method produces EOL predictions with RMSE and MAPE rates of 100 cycles and 10%, respectively, by using GPR and SVM to extract features from incomplete charging data FIGURE 1.

Lithium plating in lithium-ion batteries during rapid charging in embedded systems is the main topic of C. Zoerr et al. [23]. They successfully reduce the dangers of lithium plating by introducing a unique charging process that is based on the relationship between anode potential and negative electrode polarisation. This proven method, which is used in a variety of settings, uses an anode potential control that is developed from a Newman-type P2D modelling framework and shows a significant decrease in the dangers associated with lithium plating. A novel end-to-end deep learning architecture for quick lithium-ion battery RUL prediction is presented by D. A. Najera-Flores et al. [24]. The method improves the mean absolute error rate by a significant 10.6% and makes predictions 25 times quicker by focussing on temporal patterns and cross-data correlations from raw data, such as terminal voltage, current, and cell temperature.

A thorough approach to battery RUL prediction is put out in [25], which includes an improved PSO method for figuring out the ideal degradation parameter values as well as an information entropy-based technique. In order to handle noise and capacity degradation problems in the experimental data, a Moving Average Filter (MAF) is also used. The datasets from Maryland University and NASA are used to illustrate the efficacy of the proposed technique. According to the results, the method outperforms DDB alternatives in terms of prediction accuracy while using less training data. To conduct RUL prediction, the authors of [26] provide a DDB model that strengthens the relationship between features and the battery's ageing condition by combining the PSO approach with feature improvement via box-cox transformation. PSO is used in the model parameter optimisation procedure. Experimental findings show that this approach is successful, and its efficacy is confirmed using real Li-ion battery deterioration data. In order to forecast the RUL of Li-ion batteries, the authors in [27] use the NASA dataset to create a model of battery deterioration by integrating Support Vector Regression (SVR) with the Artificial Bee Colony (ABC) method. The SVR kernel parameters are optimised using ABC. The findings show that when it comes to parameter optimisation, the ABC method performs better than the PSO approach. The study in [28] emphasizes how crucial lithium-ion battery health is becoming to the electrification of transportation. In order to forecast RUL, they propose a novel method that combines transfer learning, Gaussian process regression, and gated recurrent neural network approaches. This approach improves accuracy beyond conventional techniques by optimising health indicators, introducing a self-correction mechanism, and implementing online model correction—all of which are critical for future maintenance in battery management. Examine the end-of-life prediction and degrading behaviour of lithium titanate oxide (LTO) batteries in [29]. Using a feed-forward neural network model for precise health condition and end-of-life predictions, the study investigates the effects of temperature, current rate, and cycle depth on capacity deterioration and cycle life. Their study highlights variables influencing the lifespan and performance of LTO batteries, showing that high temperatures increase deterioration. Using

a machine learning technique, the authors in [30] investigate how coolant flows and road gradient affect the electrical components of electric vehicles that are powered by batteries. While acknowledging the processing resources needed for training bigger datasets, the study highlights the critical impact that data quantity plays in improving prediction accuracy for artificial neural networks (ANNs), implying that sufficient data is essential for maximum performance.

III. PROPOSED MODEL

The working procedure of the proposed model is comprehensively detailed in this section to illustrate its robustness and effectiveness. The model initiates with data acquisition, where the necessary data is collected from a relevant dataset, ensuring it encompasses all critical parameters for battery health monitoring. Following this, data pre-processing is performed, which includes cleaning the data to remove inconsistencies and normalizing it to standardize the range of features for uniformity and to enhance the model's performance. Subsequently, the Harris Hawks Optimization (HHO) algorithm is employed for feature extraction, which identifies and selects the most relevant features contributing to accurate predictions. These features are then utilized by machine learning techniques, specifically XGBoost and Artificial Neural Networks (ANN), to predict the Remaining Useful Life (RUL) of the batteries based on their current levels and operational conditions. The prediction accuracy and efficacy of the model are validated using regression performance metrics such as RMSE and MAE. Lastly, a feature importance analysis is carried out to determine which characteristics in the dataset exert the most influence on the predictions, providing valuable insights into the critical factors affecting battery life. The workflow, illustrated through the block diagram in FIGURE 1 highlights the seamless integration of these stages, showcasing the structured approach adopted by the proposed RUL prediction model.

A.DATA ACQUISITION

The data comes from a research by the Hawaii Natural Energy Institute that examined 14 distinct 18,650 Nickel Manganese Cobalt-Lithium Cobalt Oxide (NMCLCO) batteries, each of which had a 2.8 Ah nominal capacity. "18,650" refers to a certain battery size standard that is around 18 mm in diameter and 65 mm in length. These batteries were subjected to a rigorous cycling schedule, lasting more than 1000 cycles at a regulated 25 °C. Using a constant current–constant voltage (CC-CV) charge rate set at a C/2 rate—that is, charging at half the battery's capacity per hour—the charging and discharging procedures were standardised. They were also discharged at a rate of 1.5 C, which means that the battery was being charged at a rate that was 1.5 times its capacity per hour. The dataset is openly available in (https://www.kaggle.com/datasets/ignaciovinuales/battery-remaining-useful-life-rul)

Specific characteristics were extracted from the original dataset in order to assist predictive modelling and extract valuable insights. These characteristics draw attention to the patterns of voltage and current seen throughout each battery cycle [21]. The carefully selected properties provide crucial information with the goal of accurately predicting these batteries' RUL, a crucial indicator for evaluating battery health. Electric cars are among the many applications that often employ NMC-LCO batteries. The behaviour and chemistry of those batteries in the dataset may provide light on the lifespans and patterns of deterioration of comparable batteries used in EVs. The dataset's batteries experienced more than 1000 cycles, which is comparable to the kind of cycle life testing that EV batteries would experience. A comprehensive dataset for comprehending how these batteries deteriorate over time and cycles is provided by this thorough cycling.

B. DATA PRE-PROCESSING

Duplicate and missing values were checked in the data. Following preprocessing, the dataset was found to be free of duplicate or missing occurrences. Maintaining the dataset's integrity involves making sure there are no duplicate or missing entries. Particularly in machine learning models, any missing data might result in skewed analysis or incorrect predictions. Every record in the dataset is guaranteed to be unique by eliminating duplicates. The quality and efficiency of statistical analyses and machine learning model training may be impacted by duplicated data, therefore this is very crucial. There were no missing or duplicate occurrences in this dataset. Here, Min-Max normalization is used here for normalizing data, which are specifically in the range of (0, 1) or (-1, 1). The mathematical expression can be given in Eq. (1),

$$d_{norm} = \frac{d - min\,(d)}{max(d) - min\,(d)} \qquad (1)$$

Here, 'd' is the obtained data and 'd_{norm}' is the data after normalization, $\max(d)$ and $\min\,(d)$ are the maximal and minimal rates of 'd' in the dataset, considerably.

C. FEATURE SELECTION

Feature selection makes it easier to select a subset of characteristics, which significantly improves ML models' capacity for prediction. HHO is used to pick features in order to achieve this. Feature extraction is the procedure of choosing and removing a subset of important features from a larger redundancy rate and undesired data in order to achieve effective training. To improve training results for learning performance detection and design project duration, feature selection is a technique for removing redundant and unnecessary information. Certain computations may be eliminated with the use of feature selection rather than replica complexity. In this instance, HHO is used for feature selection. The freshly developed HHO meta-heuristic algorithm is based on the concept of swarm intelligence. To mimic their behaviour in the wild, Harris Hawks' characteristic chasing techniques are used to capture their prey. In HHO, many hawks work together to chase prey in various ways, making it a population-based algorithm. The exploration phase and the exploitation phase are the two main phases of the Harris

Hawk. Hawks utilise a range of attack strategies, make a quick jump, and get ready to find their target throughout these stages. The HHO method may be used to any optimisation problem. The primary working exploitation and exploration phases are shown based on the prey energy level (EL) and activity chances (a *and* b).

1)EXPLORATION PHASE:

The hawks are said to have extraordinary vision for following and finding their prey in HHO. Thus, the region is set up to keep an eye on the prey. As a result, the hawks use the two unofficial surveillance locations to view their prey and pole at the region. When $a < 0.5$, the hawks rest on the same tree as their prey; when $a \geq 0.5$, however, they rest unpredictably on any tree of any size. Every monitoring strategy has the same likelihood rate. The HHO model may go from the exploration phase to the exploitation phase based on the energy rate that the prey is exerting in Eq. (2).

$$X = 2X_0 \frac{1-n}{N} \quad (2)$$

With 'Nstanding for the total number of iterations, 'n' for the current iteration, and X_0 for the prey's beginning energy, the above equation represents the prey's energy level, which may decrease as the number of iterations increases. The prey is regarded as influential, and the initial energy level, X_0, for each cycle begins with $(-1,-1)$, albeit it may range from $(0,1)$.

2) EXPLOITATION PHASE

Prey can often dodge dangerous situations with ease. The hawks use a variety of chase techniques as a result. During the exploitation stage, the hawks' plan asks for the use of four main tactics. Here, 'b' stands for the prey's likelihood under two conditions: $b \geq 0.5$ for being unable to escape, or $b < 0.5$ for being able to do so. Hard besieges (HB) and soft besieges (SB) are processed for encircling the prey. The hawks may circle the prey at various locations depending on its EL. The hawks attack their target in unison to increase their chances of catching it. As soon as the victim releases energy, the hawks increase the intensity of their siege to capture it. SB is processed when $a \geq 0.5$, while HB is used when $|X| < 0.5$. In ALGORITHM 1, the HHO pseudo code is introduced. The prey may begin moving quickly and the SB can be processed before a sudden assault occurs when $(a < 0.5)$ and $X < 0.5$. However, when $(b < 0.5)$ and $|X| < 0.5$, the HB is used, which prevents the prey from running. Here, the TABLE 1 presents the RUL feature dataset for the proposed model.

ALGORITHM 1: **HHO Algorithm for Feature Selection**

1.$P \leftarrow population_{size}$,
$N \leftarrow total\ number(iterations), n \leftarrow present\ iteration$
2. $P_i\ (i = 1,2,3, \dots, K)$
3. WHILE $(\boldsymbol{end})$
4.COMPUTE
$(new\ fitness\ rate\ (F), ideal\ location\ (L), prey's\ location$
5. FOR $\boldsymbol{everyhawk(P_i)}$
6. UPDATE X_0
7. CALCULATE $energy(prey)(X) using(1)$
8. IF $(|\boldsymbol{X}| \geq \boldsymbol{1})$ THEN
9. $(RunExplorationPhase)$
10. IF $(|\boldsymbol{X}| < \boldsymbol{1})$ THEN
11. $(Run\ Exploration\ Phase)$
12. ELSE IF$(\boldsymbol{a \geq 0.5\ and\ |X| < 0.5})$ THEN
13. EXECUTE HB
14.UPDATE $Position(Hawk)$
15. ELSE IF (A$< \boldsymbol{0.5\ and\ |X| \geq 0.5}$) THEN
16. EXECUTE SB
17. COMPUTE $F(hawk)$
18. result← $\boldsymbol{best\ feature\ subset}$
19. return result

TABLE 1
RUL feature Dataset

Features	Depiction
Cycle Index	Represents the chronological number of battery cycle
Battery Discharge Time (s)	Denotes the time of discharge period for each cycle
Max. Voltage Charging (v)	Denotes the maximal battery discharge
Min. Voltage Charging (v)	Denotes the minimal battery discharge
Time Constant Current (t)	Represents the time constant during the battery cycle
Charging Time	Denotes the time taken for the battery to be charged fully
RUL	Denotes Remaining Useful Life

D. TRAINING AND RUL PREDICTION

This section introduces the proposed methods for estimating Li-ion battery RUL. We have created an ANN that effectively pulls features from training data in order to solve this difficulty. 20% of this data is used for validation, while the remaining 80% is used for training. Following training and validation, the CNN model is adjusted by removing the last two layers and obtaining a collection of features from the layer before it. The chosen features from the measured data and the newly proposed feature taken from the battery charging regulations are then concatenated with the ANN-derived feature set. For early RUL prediction, this combined feature set is then put into an XGBoost model. The ANN model and feature extraction process hyperparameters are fine-tuned using the COM to increase the accuracy of the proposed technique. The best feature subset values for the ANN model are found with the help of COM. In-depth mathematical illustrations and descriptions of the used techniques—ANN, XGBoost, COM, hyperparameter tweaking, and evaluation metrics—are given in the following subsections.

1) ANN IN RUL PREDICTION

A subtype of deep learning models called ANN models are ideal for handling grid-like data, including pictures. ANNs are often used for tasks involving a lot of visual input, such as object identification and picture categorisation. Convolution is the main method used in the construction of ANNs to find patterns in data. The input data is filtered before the convolution procedure is carried out. Once the filter, which is a compact matrix of weights, has been moved, the dot product between the input data and the filter is calculated at each location. Upon completion of the convolution process, a

feature map is generated. The feature map is then subjected to a nonlinear activation function, such as the rectified linear unit (ReLU). The activation function adds nonlinearity to the model, enabling it to recognise intricate patterns. Filters that increase in complexity with each new layer are often produced by repeatedly iterating the activation function and convolution process. In general, the final element of ANN is the fully connected layer, computing the classification of processed data. The input data matrix is considered as, M of size $\times\, y$, 'H' is considered as filter matrix with size $q \times r$, hence, the convolution matrix can be defined as Eq. (3),

$$C_{i,j} = \sum_{m=0}^{q=1} \sum_{n=0}^{r=1} H_{mn} M_{i+m,j+n} \quad (3)$$

Here, the output feature map is '$C_{i,j}$', where, i, j are the indices of output element. One way to conceptualise the convolution procedure is as a sliding window applied to the incoming data. The dot product of the input data and the filter is computed when the filter is positioned in the upper-left corner of the data. The dot product is then computed once again after moving the filter one pixel to the right. Until the filter has been applied to all of the input data, this procedure is repeated. The features that have been retrieved from the input data are represented in the feature map, which may subsequently be used to categorise the input data.

2) BOOSTING WITH XGBOOST:

XGBoost is a type of gradient-boosting machine that is designed for speed and accuracy, and is commonly used for regression and classification tasks. XGBoost works by building a series of decision trees. Each decision tree is built on a subset of the data, and the predictions from the individual trees are combined to make a final prediction. XGBoost uses a gradient-boosting algorithm to train the decision trees. Gradient boosting is an iterative algorithm that builds a model by repeatedly adding new decision trees to the model. Each new decision tree is built to correct the errors of the previous trees. The objective function of the model can be defined as Eq. (4)

$$L(c, \hat{c}) = \sum_{i=1}^{n} r(c_i, \hat{c}_i) + \varphi(\widehat{pf}) \quad (4)$$

Here, $'c'$ denotes the true label, '$\hat{c}$' denotes the predicted label, and the loss function is 'n' with 'φ' is the regularization term, for managing the model complexity. The computation can be given as Eq. (5),

$$\varphi(\widehat{pf}) = \gamma N + \frac{1}{2} l \sum_{j=1}^{m} g_j^2 \quad (5)$$

Here, $\widehat{pf}$ is the predicted function, 'γ and l' are regularization factors, g_j^2, is the weight factor.

3)HYPERPARAMETER TUNING

Optimising hyperparameters becomes a crucial step in building an ML model in order to increase the model's efficiency. Finding the ideal values for the parameters that control the model's behaviour is the process of hyperparameter optimisation. Aspects such as learning rate, number of trees, estimators, maximum depth, minimum child weight, and regularisation parameters are among the hyperparameters in consideration. There are many ways to optimise hyperparameters, but the most often used ones include grid search, random search, and Bayesian optimisation. These approaches may provide computational challenges, especially when dealing with large search areas. Grid search, which evaluates all possible combinations of hyperparameters within preset limits, was used in this work. By developing a hyperparameter-tuned version of XGBoost, which is achieved by Grid Search CV, its effectiveness is increased. Learning rate, number of trees, maximum depth of each tree, and number of samples and features in each tree are among the XGBoost model hyperparameters optimised in this study.

IV. RESULTS

The performance of the proposed model is evaluated with some factors such as, Mean Squared Error (MSE), Mean Absolute Error (MAE), Root Mean Squared Error (RMSE) and R-Squared rate. The mathematical computations of these factors are given with x_a denotes actual rate and x_p denotes the predicted rate for the set of 'm' samples. The R-squared value, obtained from the coefficient of computation, determines the variation proportion in the dependent variable that is conventional from self-determining variables. The formulae are given in Eq. (6), Eq. (7), Eq. (8), and Eq. (9).

$$Mean\ Squared\ Error\ (MSE) = \frac{|(x_a - x_p)|}{m} \quad (6)$$

$$Mean\ Absolute\ Error\ (MAE) = \frac{\sum (x_a - x_p)^2}{m} \quad (7)$$

$$Root\ Mean\ Squared\ Error\ (RMSE) = \sqrt{\frac{\sum (x_a - x_p)^2}{m}} \quad (8)$$

$$R^2 = 1 - \frac{\sum (x_a - x_p)^2}{\sum (x_a - \widehat{x_a})^2} \quad (9)$$

Here, '$\widehat{x_a}$' is the average of all the actual rates. The unit for the above metrics is volts (V).

Here, the performance measures of many machine learning methods are compared in order to predict the RUL values of batteries. For both the training and test datasets, the MAE, MSE, RMSE, and R-Squared metrics are assessed. In both the training and test sets, the XGBoost algorithm has the lowest MAE and RMSE values, demonstrating its better RUL prediction accuracy. Furthermore, the R-Squared values are quite high, indicating that the model fits the data almost exactly. The other regression models has somewhat higher MAE and RMSE values than XGBoost of EE-BOM, despite its strong performance, particularly in the training set. Its R-Squared values, however, are always high, suggesting a high degree of predictive power. Existing models like Support Vector Regression (SVR) and Gaussian Process Regression (GPR) are used to assess the outcomes.

V. DISCUSSIONS

TABLE 2 presents the overall findings, and the related graphs for comparative assessments are provided below. The FIGURE 2 portrays the comparison graph for MSE among models. It can be observed from the Figure that the proposed model with XGBoost technique attains minimal rate of MSE

and MAE, showing the results in FIGURE 3. Further, the results for RMSE and R^2 are demonstrated in FIGURE 4 and FIGURE 5. Both the graphs evidence that the proposed EE-BOM model with HHO and ANN with XGBoost produce better results than compared models. MSE tells the average squared difference between the actually measured and the predicted value of battery performances; thereby making it a suitable measure for estimating the accuracy of a model. As it can be evidenced, based on the MSE values depicted in FIGURE 2, the EE-BOM outperforms GPR, and SVR in all the datasets used in the study. For instance, in dataset 1, the value of MSE for EE-BOM is 21:1 while GPR and SVR=27:3 MSE and 43:1 MSE respectively. The same trend is observed in all the datasets; however, EE-BOM has smaller error bounds. It is also noticeable in the figure that the highest MSE is depicted by SVR implying that the measure of mean squared error denotes a higher level of prediction error than the rest of the models. Therefore, it can be concluded that in cases of energy variations, EE-BOM can be considered as having the lowest prediction errors which are important in battery management. MAE offers a rather simple average of absolute errors and is effective to check the models' resilience to real-life conditions. As it can be observed in the FIGURE 3, EE-BOM has the lowest MAE values for all the datasets. In dataset 1, MAE of proposed EE-BOM is 22.3 while for GPR it is 44.6 and SVR 59.3. This format is repeated in all the data sets to show that EE-BOM is capable of handling the reduction of the extent of the prediction errors. Comparing the values, it can be conclusion that SVR has higher MAE that may signify that it has low accuracy in predicting the optimal battery level. MAE and RMSE are outlined as follows: RMSE is more severe type of error measure as compared to MAE used in battery capability applications as they punish large errors in battery efficiency. It is worthwhile to mention that in all the scenarios defined in the experiment, EE-BOM consistently yields the lowest RMSE values as illustrated in FIGURE 4. For instance, in dataset 3, the RMSE score is recorded as 13 for EE-BOM while the GPR and SVR give out 28.5 and 44 respectively. The same applies to other sets; these statistics support the effectiveness of EE-BOM in efficiency of battery optimization, not in large oscillations in the forecast.

R^2 defines the proportion of variance in battery performance that is predictable by the model; values closer to 1 are the best. In contrast to the general observations made in terms of the error measures indicated above, the following trends of R^2 as depicted in FIGURE 5 can be deduced, EE – BOM has lesser R^2 values as compared to GPR as well as SVR. For example, in dataset 1, EE-BOM results of R^2 is 0.294 while in the case of GPR it is 0.787 and for the SVR it is 0.654. In every data case, SVR has both higher R^2 values and coefficient of determination than the other algorithms, which indicate that SVR presents higher variance in battery performance though it has higher error rates. These are the tradeoffs: EE-BOM is good at preventing errors overshadowing them, but SVR may be superior in capturing varied behavior of batteries.

The analysis of state-of-health (SOH) and remaining useful life (RUL) prediction for batteries has gained significant traction due to advancements in methodologies and the availability of diverse datasets. These studies play a critical role in enhancing battery performance, extending battery life, and ensuring reliability across a variety of applications, such as electric vehicles (EVs), renewable energy storage, and portable electronic devices. Researchers have explored innovative techniques, combining physics-based models, data-driven methods, and hybrid approaches to achieve remarkable improvements in prediction accuracy and computational efficiency.

Chou et al. [13] introduced a groundbreaking hybrid approach that leverages transfer learning, Bi-LSTM, attention mechanisms, support vector regression (SVR), and empirical mode decomposition (EMD). This comprehensive method significantly reduced relative errors to as low as 6.96%, showcasing its robustness across diverse charging policies and battery datasets. Zhao et al. [14], on the other hand, utilized a feature-based machine learning approach with a stacking ensemble technique to analyze a dataset comprising 420 cells and 9 battery packs. By employing a two-step noise reduction process tailored for EV batteries, they achieved highly accurate and physically consistent predictions.

Greenbank et al. [19] presented a data-driven Bayesian (DDB) approach, integrated with automated feature selection and Gaussian process regression (GPR), to predict the end-of-life (EOL) of batteries. Their model demonstrated remarkable adaptability, handling diverse input datasets with precision. Meanwhile, Chen et al. [20] developed a hybrid framework combining variational mode decomposition (VMD), multi-kernel SVR, and the salp swarm algorithm (SSA) for SOH estimation. Although effective for SOH prediction, this method faced challenges when applied to RUL estimation.

Further extending these efforts, Zhao et al. [21] combined equivalent circuit models with unsupervised learning to identify physics-informed features, achieving a high deterioration detection rate exceeding 90% with an RMSE of 53.56% on data from 65 batteries. Zhang et al. [22] tackled battery life classification using a moving window technique, GPR, and SVM, achieving impressive results with an RMSE of 100 cycles and a mean absolute percentage error (MAPE) of 10% from incomplete charging data. Ali et al. [23] addressed lithium plating risks during fast charging through anode potential control and P2D modeling, demonstrating a significant reduction in degradation risks during rapid charging.

Montesinos López et al. [24] developed an end-to-end deep learning framework capable of capturing temporal and cross-data correlations from raw voltage, current, and temperature data. Their approach achieved a 10.6% improvement in mean absolute error (MAE) while operating 25 times faster than traditional techniques. Similarly, Long et al. [25] employed particle swarm optimization (PSO) alongside information entropy and moving average filtering (MAF) to address noise and capacity degradation, surpassing

the performance of DDB models with minimal training data requirements.

Peng et al. [26] harnessed the power of PSO and Box-Cox transformations for optimizing feature selection in battery degradation datasets, while Wang et al. [27] applied artificial bee colony (ABC) optimization for SVR kernel parameter tuning, showcasing superior results over PSO on NASA datasets. Zhao et al. [28] proposed a novel method combining transfer learning with GPR and gated recurrent units (GRU). Their self-correction mechanism enabled online RUL prediction for EV batteries with enhanced accuracy and adaptability to dynamic conditions. Soltani et al. [29] utilized feed-forward neural networks (NNs) to investigate the effects of environmental and operational factors, such as temperature, current rate, and cycle depth, on the degradation of lithium titanate oxide (LTO) batteries. Their findings revealed that elevated temperatures significantly accelerated battery degradation. Finally, Çolak et al. [30] employed machine learning techniques to analyze the impact of coolant flow and road gradients on EV battery performance. They demonstrated that artificial neural network (ANN) accuracy improved significantly when trained on larger datasets.

These diverse studies collectively highlight the transformative potential of hybrid methods, feature-based models, and deep learning frameworks in advancing battery SOH and RUL predictions. By combining physics-informed models with data-driven insights, researchers can enhance prediction accuracy while accounting for real-world variability in battery performance. As the field evolves, integrating interdisciplinary approaches and leveraging emerging technologies such as IoT, edge computing, and advanced sensors will further optimize battery management systems. These advancements not only support the development of efficient energy storage solutions but also contribute to the scalability of renewable energy systems. Ultimately, these innovations pave the way for sustainable energy transitions and improved reliability in critical applications like electric vehicles and smart grids.

TABLE 2
Performance of different models' comparison

Authors	Methodology	Dataset	Key Features	Performance Metrics
Chou, J et al. [13]	Transfer learning + Bi-LSTM + Attention + SVR + EMD	Various charging policy batteries	Hybrid approach for improved RUL prediction	Relative error: 6.96%, 0.6%, 6.25%
Zhao, J et al. [14]	Feature-based ML + Stacking ensemble	420 cells, 9 battery packs	Two-step noise reduction for EV batteries	Accurate, physically consistent predictions
S. Greenbank et al. [19]	DDB + Feature selection + GPR	Custom dataset	Automated feature selection for EOL prediction	Adaptability in diverse inputs
Y Chen et al. [20]	VMD + Multi-kernel SVR + SSA	NASA dataset	Hybrid method for SOH estimation	Not applicable for RUL prediction
Zhao, M et al. [21]	Equivalent circuit model + Unsupervised learning	65 batteries	Physics-informed features	RMSE: 53.56%, Deterioration detection: >90%
Zhang, Y et al. [22]	Moving window + GPR + SVM	Incomplete charging data	In-situ battery life classification	RMSE: 100 cycles, MAPE: 10%
Ali et al. [23]	Anode potential control + P2D model	Embedded systems	Mitigates lithium plating during fast charging	Significant reduction in plating risks
Montesinos López et al. [24]	End-to-end deep learning	Raw voltage, current, temp data	Captures temporal & cross-data correlations	10.6% MAE improvement, 25× faster
B. Long et al. [25]	PSO + Information entropy + MAF	NASA, Maryland University datasets	Handles noise, capacity degradation	Outperforms DDB models with less training data
Peng J et al. [26]	PSO + Box-Cox transformation	Real battery degradation data	Feature improvement for ageing analysis	Successful parameter optimization
Y. Wang et al. [27]	SVR + ABC	NASA dataset	ABC for SVR kernel parameter optimization	ABC outperforms PSO in optimization
Zhao, G et al. [28]	Transfer learning + GPR + GRU	EV battery datasets	Self-correction, online model correction	Improved RUL prediction accuracy
Soltani, M et al. [29]	Feed-forward NN	LTO batteries	Impact of temp, current rate, cycle depth	High temp accelerates degradation
Çolak, B et al. [30]	ML for coolant flow & road gradient effects	EV battery systems	Evaluates impact on electrical components	ANN accuracy improves with more data

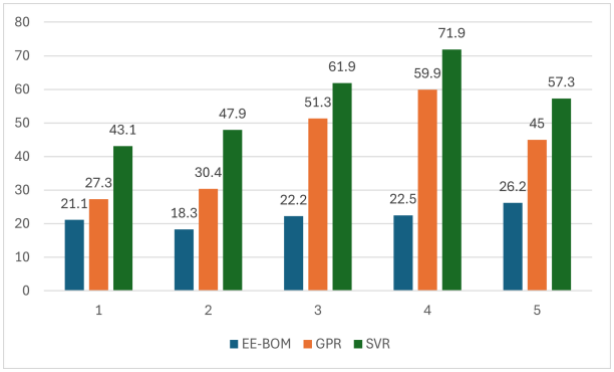

FIGURE 2. (a) MSE among Models Vs Dataset

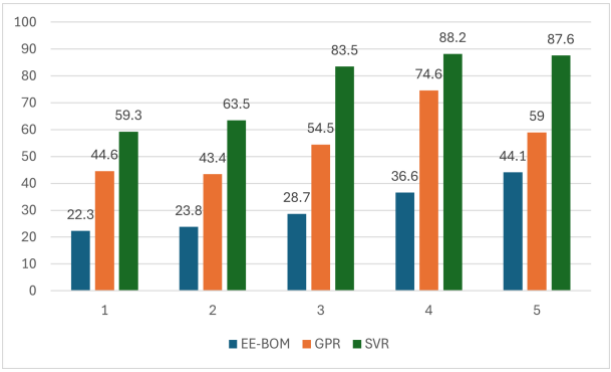

FIGURE 2. (b) MAE among Models Vs Dataset

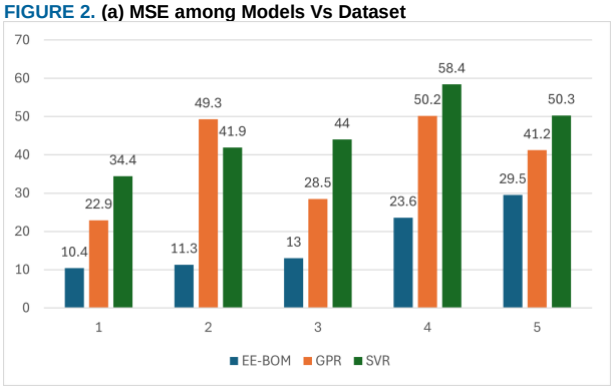

FIGURE 2. (c) RMSE among Models Vs Dataset

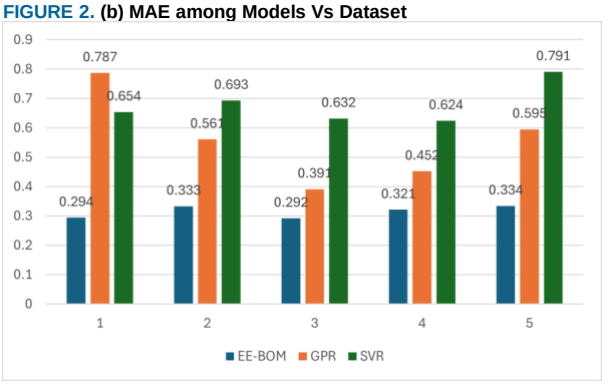

FIGURE 2. (d) R2 among Models Vs Dataset

Figure. (a) MSE among Models Vs Dataset (b) MAE among Models Vs Dataset, (c) RMSE among Models Vs Dataset, (d) R2 amond Models Vs Dataset

TABLE 3
Observations of MSE, MAE, RMSE, R2

Model/Training data	10	20	30	40	50	**Model/Training data**	10	20	30	40	50
EE-BOM	21.1	18.3	22.2	22.5	26.2	**EE-BOM**	22.3	23.8	28.7	36.6	44.1
GPR	27.3	30.4	51.3	59.9	45.0	**GPR**	44.6	43.4	54.5	74.6	59.0
SVR	43.1	47.9	61.9	71.9	57.3	**SVR**	59.3	63.5	83.5	88.2	87.6
A. MSE						**B.MAE**					
Model/Training data	10	20	30	40	50	**Model/Training data**	10	20	30	40	50
EE-BOM	10.4	11.3	13.0	23.6	29.5	**EE-BOM**	0.294	0.333	0.292	0.321	0.334
GPR	22.9	49.3	28.5	50.2	41.2	**GPR**	0.787	0.561	0.391	0.452	0.595
SVR	34.4	41.9	44.0	58.4	50.3	**SVR**	0.654	0.693	0.632	0.624	0.791
C.RMSE						**D.R2**					

High-importance features are useful, but it's crucial to make sure they don't cause overfitting or add biases into the model. Even if a feature is very predictive, relying too much on it might weaken the model's resilience or limit its generalisability. As a result, it is essential to take into account all qualities fairly, especially those with lower ratings. The research's findings will contribute to long-term improvements in the infrastructure for electric vehicles. According to this proposed work, enterprises that depend on Li-ion batteries may save money and improve operational efficiency by using preventative maintenance and resource optimisation made possible by precise RUL prediction. Li-ion batteries are essential parts of many different sectors, but especially of EVs. It is essential to forecast these batteries' RUL values in order to carry out preventative maintenance plans. Because it allows companies or other applications to accurately plan maintenance as required, accurate RUL prediction promotes effective resource utilisation. Cost reductions and improved operational effectiveness may arise from this. Li-ion battery longevity may be increased by comprehending and forecasting RUL. Longer-lasting batteries lessen the environmental effect of frequent replacements and disposal, which is essential for sustainable growth. Accurate RUL forecasts let electric car owners plan their trips more effectively. This model is in line

with the continuous changes occurring at the nexus of industry and technology. Comparative evaluation of different machine learning methods offers important information on how well they predict RUL. Finding better algorithms, such as XG Boost in this instance EE-BOM, sets standards for further study and use in related fields. The environmental effect of Li-ion batteries is acknowledged in this paper, and it is proposed that precise RUL projections might allay worries about battery disposal. Technologically driven sustainable practices are essential for reducing environmental impacts.

V. CONCLUSIONS

The paper devices a model for RUL prediction of EVs using machine learning approaches. Initially, data acquisition is for obtaining input dataset, following data-pre-processing is employed for enhancing the input data quality. When a variety of machine learning methods are tested, XGBoost performs better in RUL prediction. By employing HHO to acquire the extracted features, the results demonstrate how well the XGBoost algorithm minimises errors and predicts RUL. Mean Squared Error (MSE), Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and R-Squared rate are used to conduct the assessments. In comparison to previous examined works, the suggested model achieves a minimum average MSE, MAE, and RMSE of 23.68%, 36.41%, and 24.67%, respectively. The proposed work promotes effective travel planning and makes it easier batteries to create with longer lifespans.

In future, the research can be enhanced by removing the data quality and availability limitations. Additionally, this study could not have taken into consideration changes in the surrounding environment that might affect battery performance. Although they weren't specifically mentioned, variables including temperature, humidity, and use patterns may have an impact on RUL.

REFERENCES

[1] Khan, F. N. U., Rasul, M. G., Sayem, A., & Mandal, N. Maximizing energy density of lithium-ion batteries for electric vehicles: A critical review. Energy Reports,2023, 9, 11-21.

[2] Alanazi, F. Electric Vehicles: Benefits, Challenges, and Potential Solutions for Widespread Adaptation. Applied Sciences, 2022, 13(10), 6016.

[3] Kumar, M., Panda, K. P., Naayagi, R. T., Thakur, R., & Panda, G. Comprehensive Review of Electric Vehicle Technology and Its Impacts: Detailed Investigation of Charging Infrastructure, Power Management, and Control Techniques. Applied Sciences, 2022, 13(15), 8919.

[4] Kang, Z., Catal, C., &Tekinerdogan, B. Remaining Useful Life (RUL) Prediction of Equipment in Production Lines Using Artificial Neural Networks. Sensors, 2020 21(3), 932.

[5] Y. Zheng et al., "Sensorless State of Temperature Estimation for Smart Battery based on Electrochemical Impedance," 2023 25th European Conference on Power Electronics and Applications (EPE'23 ECCE Europe), Aalborg, Denmark, 2023, pp. 1-8.

[6] L. Ren, L. Zhao, S. Hong, S. Zhao, H. Wang and L. Zhang, "Remaining Useful Life Prediction for Lithium-Ion Battery: A Deep Learning Approach," in IEEE Access, vol. 6, pp. 50587-50598, 2018.

[7] C. Zhang, S. Zhao and Y. He, "An Integrated Method of the Future Capacity and RUL Prediction for Lithium-Ion Battery Pack," in IEEE Transactions on Vehicular Technology, vol. 71, no. 3, pp. 2601-2613.

[8] Schuster, S. F., Bach, T., Fleder, E., Müller, J., Brand, M., Sextl, G., &Jossen, A. Nonlinear aging characteristics of lithium-ion cells under different operational conditions. Journal of Energy Storage, 2015, 1, 44-53.

[9] Uzair, M., Abbas, G., & Hosain, S. Characteristics of Battery Management Systems of Electric Vehicles with Consideration of the Active and Passive Cell Balancing Process. World Electric Vehicle Journal, 2021, 12(3), 120.

[10] Wang X, Huang R, Wen J, Shi Y, Jia J, Zeng J. A Lithium-ion Battery RUL Prediction Method Considering the Capacity Regeneration Phenomenon. *Energies*. 2019, 12(12):2247.

[11] Wang, Y., Zhao, Y., &Addepalli, S.Remaining Useful Life Prediction using Deep Learning Approaches: A Review. Procedia Manufacturing, 2019, 49, 81-88.

[12] Wu, J., Kong, L., Cheng, Z., Yang, Y., & Zuo, H. RUL Prediction for Lithium Batteries Using a Novel Ensemble Learning Method. Energy Reports, 2022, 8, 313-326.

[13] Chou, J., Wang, F., & Lo, S. Predicting future capacity of lithium-ion batteries using transfer learning method. Journal of Energy Storage, 2023, 71, 108120.

[14] Zhao, J., Ling, H., Liu, J., Wang, J., Burke, A.F., Lian, Y. Machine learning for predicting battery capacity for electric vehicles. eTransportation 2023, 15, 100214. https://doi.org/10.1016/j.etran.2022.100214.

[15] Chang, C., Wu, Y., Jiang, J., Jiang, Y., Tian, A., Li, T., & Gao, Y. Prognostics of the state of health for lithium-ion battery packs in energy storage applications. Energy, 2022, 239, 122189.

[16] B. Gou, Y. Xu and X. Feng, "State-of-Health Estimation and Remaining-Useful-Life Prediction for Lithium-Ion Battery Using a Hybrid Data-Driven Method," in IEEE Transactions on Vehicular Technology, 2020, vol. 69, no. 10, pp. 10854-10867.

[17] H Meng, H., & Li, Y. A review on prognostics and health management (PHM) methods of lithium-ion batteries. Renewable and Sustainable Energy Reviews, 2019, 116, 109405.

[18] S. Jafari, Y. -C. Byun and S. Ko, "A Novel Approach for Predicting Remaining Useful Life and Capacity Fade in Lithium-Ion Batteries Using Hybrid Machine Learning," in IEEE Access, 2023, vol. 11, pp. 131950-131963.

[19] S. Greenbank and D. Howey, "Automated Feature Extraction and Selection for Data-Driven Models of Rapid Battery Capacity Fade and End of Life," in IEEE Transactions on Industrial Informatics, 2022, vol. 18, no. 5, pp. 2965-2973.

[20] Y Chen, Y., Duan, W., He, Y., Wang, S., & Fernandez, C.A hybrid data driven framework considering feature extraction for battery state of health estimation and remaining useful life prediction. Green Energy and Intelligent Transportation, 2024, 3(2), 100160.

[21] Zhao, M., Zhang, Y., & Wang, H. Battery degradation stage detection and life prediction without accessing historical operating data. Energy Storage Materials, 2024, 69, 103441.

[22] Zhang, Y., Zhao, M., &Xiong, R.Online data-driven battery life prediction and quick classification based on partial charging data within 10 min. Journal of Power Sources, 2024, 594, 234007.

[23] Ali, A., Emran, N. A., &Asmai, S. A. Missing values compensation in duplicates detection using hot deck method. Journal of Big Data, 2021, 8(1), 1-19.

[24] Montesinos López, O.A., Montesinos López, A., Crossa, J. Overfitting, Model Tuning, and Evaluation of Prediction Performance. In Multivariate Statistical Machine Learning Methods for Genomic Prediction, Springer: Cham, Switzerland, 2022.

[25] B. Long, X. Gao, P. Li and Z. Liu, "Multi-Parameter Optimization Method for Remaining Useful Life Prediction of Lithium-Ion Batteries," in IEEE Access, 2020, vol. 8, pp. 142557-142570.

[26] Peng J, Zheng Z, Zhang X, Deng K, Gao K, Li H, Chen B, Yang Y, Huang Z. A Data-Driven Method with Feature Enhancement and Adaptive Optimization for Lithium-Ion Battery Remaining Useful Life Prediction. *Energies*. 2020, 13(3), 752.

[27] Y. Wang, Y. Ni, S. Lu, J. Wang and X. Zhang, "Remaining Useful Life Prediction of Lithium-Ion Batteries Using Support Vector Regression Optimized by Artificial Bee Colony," in IEEE Transactions on Vehicular Technology, 2019, vol. 68, no. 10, pp. 9543-9553.

[28] Zhao, G.; Kang, Y.; Huang, P.; Duan, B.; Zhang, C. Battery health prognostic using efficient and robust aging trajectory matching with ensemble deep transfer learning. Energy 2023, 282, 128228. https://doi.org/10.1016/j.energy.2023.128228.

[29] Soltani, M.; Vilsen, S.B.; Stroe, A.I.; Knap, V.; Stroe, D.I. Degradation behaviour analysis and end-of-life prediction of lithium titanate oxide batteries. J. Energy Storage 2023, 68, 107745. https://doi.org/10.1016/j.est.2023.107745.
[30] Çolak, B. A new study on the prediction of the effects of road gradient and coolant flow on electric vehicle battery power electronics components using machine learning approach. J. Energy Storage 2023, 70, 108101.

AUTHOR BIOGRAPHY

SATHISHKUMAR is currently pursuing his Doctor of Philosophy (Ph.D.) in Computer Science and Engineering at Bharath Institute of Higher Education and Research (Bharath University), Chennai, India. He earned his Master of Engineering (M.E.) in Computer Science and Engineering from Sri Krishna Engineering College, Chennai, and his Bachelor of Engineering (B.E.) in Computer Science and Engineering from K.S.R. College of Engineering, Tiruchengode, India. His research interests include Generative AI, Prompt Engineering, Machine Learning, Data Analytics, and Data Science. With a strong passion for innovation, he is focused on creating AI-driven solutions to address complex challenges and is actively exploring emerging technologies in automation, data-driven decision-making, and advanced computing techniques.

Dr. YOGESH RAJKUMAR is an accomplished Assistant Professor in the Department of Information Technology at Bharath Institute of Higher Education and Research (Bharath University), Chennai, India. Renowned for his dedication to academic excellence, he has guided and mentored numerous students in their research endeavors, particularly in the dynamic field of Information Technology. Dr. Rajkumar's exceptional contributions to both teaching and research have been recognized with the esteemed Best Faculty Award, highlighting his commitment to shaping future professionals. His expertise encompasses a wide range of areas, including Data Mining, Cloud Computing, and Networking. Through his unwavering dedication to innovation and knowledge dissemination, Dr. Rajkumar continues to inspire and empower students and researchers, fostering advancements in cutting-edge IT technologies.