Machine learning-based prediction of lithium-ion battery life cycle for capacity degradation modelling

S. Kumarappa 1 and Manjunatha H M 2, *

1 Department of Mechanical Engineering, Bapuji Institute of Engineering & Technology, Davangere-577 004, Karnataka, India.
2 Department of Electrical and Electronics Engineering, Bapuji Institute of Engineering & Technology, Davangere-577 004, Karnataka, India.

World Journal of Advanced Research and Reviews, 2024, 21(02), 1299–1309

Publication history: Received on 08 January 2024; revised on 18 February 2024; accepted on 20 February 2024

Article DOI: https://doi.org/10.30574/wjarr.2024.21.2.0553

Abstract

With a particular focus on capacity degradation modeling, this paper offers a ground-breaking examination of the use of machine learning approaches for the precise prediction of lithium-ion battery life cycles. Because lithium-ion batteries are essential to many technological applications, it’s critical to comprehend and anticipate their life cycles in order to maximize performance and guarantee sustainable energy solutions. The study starts with an extensive examination of the literature, assessing current approaches critically and setting the stage for the introduction of models based on machine learning. The process entails the methodical collecting of data across a range of operational settings, environmental variables, and charging-discharging cycles. Thorough preprocessing guarantees the dataset’s consistency and quality for further machine learning model training. Predictive models are created using a variety of machine learning algorithms, including regression models, support vector machines, and deep neural networks. In order to improve prediction accuracy, the paper focuses on the reasoning behind model selection, parameter tuning, and the incorporation of ensemble approaches. In order to uncover important elements influencing the life cycles of lithium-ion batteries and provide important insights into degradation mechanisms, feature selection approaches are used. Using cross-validation techniques and real-world lithium-ion battery datasets, the built machine learning models go through rigorous evaluation and validation processes to determine their robustness, capacity for generalization, and performance metrics. Comparing machine learning-based predictions with conventional models, the results are presented and discussed, offering insights into the interpretability of the models and the identification of important affecting elements. In order to promote proactive maintenance and optimize battery usage, predictive models are integrated into real-time monitoring systems. The consequences for battery management systems are examined. The paper continues by discussing the challenges that come with using machine learning to estimate the life cycle of batteries and outlining possible directions for further research and development, such as scalability, interpretability, and the incorporation of emerging technologies. This research contributes to the ongoing efforts to increase the reliability and sustainability of lithium-ion battery technologies by highlighting the potential impact of machine learning on energy storage system optimization.

Keywords: Machine Learning; Lithium-ion Battery; Capacity Degradation; Life Cycle Prediction; Battery Management System.

1. Introduction

Lithium-ion batteries (LIBs) are becoming increasingly important in today’s energy landscape because they power a wide range of electronic devices, electric cars, and renewable energy storage systems. The need to improve battery longevity and performance has led to a thorough investigation of predictive models to prevent and manage capacity
decline, which is the main factor affecting a battery’s life cycle. In order to fully capture the intricate dynamics of degradation, traditional methods frequently fall short, requiring creative solutions. With an emphasis on capacity degradation modeling, this research sets out on a revolutionary investigation into the integration of machine learning approaches to transform the forecast of LIB life cycles.

The current energy transition to renewable and sustainable sources increases the demand for precise LIB performance forecasts. The phrase “capacity degradation” here refers to the progressive deterioration of a battery’s capability to hold and deliver charge over time, a phenomena impacted by a wide range of variables including temperature changes, operating conditions, and charge-discharge cycles. Precise prediction of this deterioration is essential for enhancing battery management tactics, prolonging battery longevity, and guaranteeing energy storage system dependability.

This work makes use of machine learning, a subject that has shown unmatched performance in a variety of predictive tasks, to address the shortcomings of traditional modeling approaches. Machine learning models which can be anything from complex neural networks to traditional regression techniques have the ability to identify complex patterns in large datasets, which can lead to a more detailed understanding of the variables causing capacity degradation. Through the utilization of these models, this research seeks to increase forecast accuracy and dependability and promote improvements in the long-term, sustainable deployment of LIBs in a variety of fields.

Incorporating machine learning into LIB remaining usable life (RUL) cycle prediction is in line with the worldwide desire to develop cleaner and more efficient energy solutions, as well as offering potential for improving current technology. The RUL prediction result describes when the battery will fail (i.e., it will no longer meet the requirement of the application). The battery RUL can be expressed as

$$\text{RUL} = T_{\text{EOL}} - T_{\text{C}}$$ (1)

where TEOL represents the battery life obtained from the battery life experiment. TC is the current usage time of the battery. Equation (1) considers calendar aging and cycle aging at the same time. Most research usually defines the RUL based on cycle aging only. Another definition that can reflect RUL is expressed as

$$\text{RUL}_i = \frac{C_i - C_{\text{EOL}}}{C_{\text{nominal}} - C_{\text{EOL}}}$$ (2)

where $C_i$, $C_{\text{nominal}}$ and $C_{\text{EOL}}$ represent the present capacity, nominal capacity, and end-of-life capacity respectively.

The integration of machine learning into LIB RUL cycle prediction appears as a revolutionary step towards a future where energy storage systems are not only powerful but also environmentally conscious and commercially viable, as we stand at the crossroads of technological innovation and sustainable energy practices. The approach, outcomes, and consequences of using machine learning to estimate the life cycle of a lithium-ion battery will be covered in detail in the parts that follow. This research will add significant knowledge to the ongoing discussion on energy storage technology.

1.1. Related Work

The growing popularity of electric vehicles (EVs) in recent years has brought attention to how crucial it is to enhance the technology of lithium-ion batteries (LIBs), with a particular emphasis on cost, energy density, and other important performance metrics. Schmuch et al. [1] stress that for EV adoption to be widely adopted, a minimum 500 km range must be achieved at a reasonable cost. The review explores Lithium-ion battery materials for automotive applications, describing the progress and difficulties encountered along the battery value chain. In the context of electric vehicle propulsion, factors including energy density, cost, and performance metrics are thoroughly analyzed to give a thorough picture of the situation now and going forward.

By addressing the critical component of comprehending the cost dynamics of battery packs, particularly for battery electric vehicles (BEVs), Nykvist and Nilsson [2] add to the body of literature. According to their analysis, industry-wide cost estimates have decreased, with an annual drop of over US$1,000 per kWh to about US$410 per kWh between 2007 and 2014. Modeling future energy and transportation systems will be significantly impacted by this trend of lowering costs, which contradicts earlier assumptions. The perspectives offered by Nykvist and Nilsson lay the groundwork for positive projections about the role BEVs will play in low-carbon transportation.

Zhang et al. [3] address the accuracy of lithium-ion battery remaining usable life (RUL) prediction, recognizing the significance of extremely accurate predictions in the energy storage and automotive industries. The suggested model
attempts to address reliability and uncertainty quantification in remaining usable life (RUL) prediction by integrating Monte Carlo Dropout with gated recurrent unit. The work offers a novel strategy for preventing over-fitting and getting the probability distribution of prediction outcomes, improving the uncertainty quantification of the prediction model by merging dropout approaches with the gated recurrent unit model.

Understanding lithium-ion rechargeable battery degradation is aided by Saha and Goebel [4], who specifically concentrate on the prognostic algorithmic scheme for predicting the State-of-Life (SOL). Their research offers an adaptable and real-time forecast of battery capacity without the requirement for physics-based models by utilizing feed-forward neural networks and particle filters. This method exhibits a promising path for adaptively predicting battery performance based on past observations, as it is straightforward and flexible.

A innovative approach for estimating the remaining usable life (RUL) of Li-ion batteries in real-time has been proposed by Patil et al. [5]. Their method combines the features of regression and classification from Support Vector (SV) based machine learning algorithms. The study uses Support Vector Machine (SVM) and Support Vector Regression (SVR) to create models for gross estimation and precise RUL prediction by evaluating cycling data under different operating conditions. The suggested method is generic and appropriate for real-time onboard RUL estimation in electric vehicle battery packs because of the multistage approach's improvement in accuracy and processing efficiency.

Liu et al. [6] offer a perceptive analysis of the development of deep learning methods in AI and their uses in a range of fields, such as computer vision, pattern recognition, and speech recognition. Popular deep learning architectures such limited Boltzmann machines, autoen coders, convolutional neural networks, and deep belief networks are reviewed in this paper. In the context of LIBs and energy storage systems, this overview not only provides a useful resource for the state-of-the-art in deep learning, but it also identifies future research paths and prospective applications.

The understanding and progress of lithium-ion battery technologies is enhanced by the collection of these chosen papers, which address topics including materials, cost dynamics, accuracy in predicting remaining usable life, adaptive prognostics, and the use of deep learning techniques.

1.2. Contribution
This study highlights significant contributions to battery management systems and sustainable technology, including the integration of machine learning methods like Support Vector Regression (SVR) for accurate life cycle prediction of lithium-ion batteries. The study's integrated approach, practical usefulness, and interdisciplinary nature position it at the forefront of developments shaping the future of energy storage. Key outcomes include the implementation of a linear regression model with elastic net regularization, tailored feature extraction, hyperparameter optimization, and model performance assessment, offering a clear framework for predicting battery cycle life and assessing battery conditions.

2. Predictive Modeling Process
In the first phase of the process of predictive modeling, the collection of data is an essential component that plays a significant part in determining the effectiveness of lithium-ion battery life cycle forecasts. It is vital to have a dataset that is fully complete and includes a wide variety of information pertaining to the behavior of batteries. The complete charge and discharge patterns, temperature conditions during operation, voltage characteristics displayed during cycles, and exact capacity measurements are all included in this. The variety and complexity of these features make it possible for the model to represent the nuanced intricacies of the battery's performance throughout the course of different time periods. As a result, the dataset becomes a store of useful insights into the intricate interaction of elements that influence battery health, so setting the groundwork for accurate forecasts.

When the data gathering stage is complete, the attention then switches to the preparation step, which is the second stage. For the purpose of ensuring that the dataset is trustworthy, a painstaking cleaning procedure is undertaken, during which missing values, outliers, and noise are addressed. The numerical characteristics are normalized or standardized, which aligns them on a similar scale and reduces the impact of variables that have, by their very nature, greater magnitudes. In order to facilitate their incorporation into machine learning models, the categorical variables belonging to the dataset have been encoded. It is essential to complete this phase in order to refine the dataset and make certain that it is free of any anomalies that might potentially undermine the effectiveness of the future predictive modeling. The steps of data collecting and preprocessing, when combined, produce a solid and well-prepared basis for the following creation of machine learning models that are aimed at estimating the remaining usable life of lithium-ion batteries as well as the capacity deterioration of these batteries.
3. Model Selection

One of the most important decisions that must be made in order to accurately anticipate the life cycle and capacity deterioration of lithium-ion batteries is to choose an appropriate machine learning model. Within the framework of this discussion, models like as Random Forests, Gradient Boosting Machines, Neural Networks, Support Vector Machines (SVM), and Support Vector Regression (SVR) each have their own unique set of benefits and factors to take into account. Neural Networks give flexibility in capturing intricate patterns, but they may lack interpretability. Random Forests and Gradient Boosting Machines are recognized for their ability to handle complex connections and provide feature importance. Neural Networks are also known for their ability to provide feature importance. The support vector machine (SVM) is distinguished by its ability to perform well in high-dimensional spaces, its resistance to over fitting, and its adaptability with a variety of kernel functions. Because of its ability to handle non-linear patterns, support vector machines (SVM) are an excellent option for forecasting the behavior of lithium-ion batteries. SVM is particularly well-suited for situations in which the correlations between features and battery life are not strictly linear [7].

When it comes to projecting the RUL of lithium-ion batteries and modeling capacity decline, Support Vector Machines (SVM) appear as the best option among the numerous models that have been taken into consideration. The Support Vector Machine (SVM) is a well-known algorithm that is capable of constructing a robust decision boundary in high-dimensional feature spaces. This capacity makes it excellent for capturing subtle connections inside large datasets. Because of its flexibility to a wide range of kernel functions, it is able to efficiently handle linear as well as non-linear patterns when applied. Furthermore, support vector machines (SVM) provide an excellent mix between accuracy and generalization, which makes them a great competitor for applications in which the prediction of battery life is considered to be of considerable importance. The capability of the model to handle a wide range of operating circumstances, temperature changes, and voltage characteristics fits well with the multifarious nature of the behavior of lithium-ion batteries. As a result, the model is an effective instrument for producing accurate and trustworthy predictions in the field of battery management systems [8].

4. Capacity Degradation Modelling

Support Vector Regression (SVR) modeling is a strategic use of machine learning techniques that is used to anticipate the remaining usable life of lithium-ion batteries. This modeling methodology is considered to be capacity degradation modeling. The first step in the process is called "data preparation," and it involves the collection of a complete dataset that includes essential parameters such as charging and discharging patterns, temperature conditions, and voltage characteristics. After the data has been preprocessed and there has been feature engineering done in order to extract information that is pertinent, the dataset is then divided into training and testing sets. In the subsequent step, the SVR model is trained on the training set, with hyperparameters like the kernel type and regularization parameters being taken into consideration. Evaluation of the model on the testing set using regression metrics such as Mean Squared Error guarantees that the model is accurate and can be generalized to data that has not yet been observed.

Because of its capability to manage non-linear interactions and to capture complex patterns, SVR is ideally suited for forecasting the complicated process of capacity deterioration in lithium-ion batteries. This is because SVR is able to handle non-linear relationships successfully. In order to gain a comprehensive picture of the state of the battery, the selected characteristics, which include charging rates, temperature profiles, and voltage behaviors, are essential. Upon completion of the training process, the SVR model is able to provide predictions on the remaining capacity or degradation level, hence offering significant insights into the anticipated lifespan of batteries. Continuous monitoring and refining, in addition to validation in the actual world, are additional factors that contribute to the flexibility and dependability of the model over time [9].

In terms of practical application, capacity degradation modeling that makes use of SVR gives stakeholders the ability to make educated decisions on the maintenance and replacement of batteries as well as other aspects of system optimization. As a result of the utilization of this prediction tool, the efficiency and lifetime of lithium-ion batteries in a variety of applications, including electric cars and renewable energy storage systems, are significantly improved, therefore contributing to the development of energy solutions that are both sustainable and dependable [10].

Capacity degradation modeling involves capturing the changes in a battery's capacity over time, usually through mathematical equations or machine learning models. Below, I'll provide a conceptual overview using equations and a simplified flowchart to illustrate the process:
Mathematical Equation for Capacity Degradation Modeling:

A basic equation for capacity degradation might take the form:

\[ \text{Remaining Capacity} = \text{Initial Capacity} \times e^{-\beta t} \]  

(3)

Here:

- Remaining Capacity is the capacity at a given time.
- Initial Capacity is the battery's initial capacity.
- \( \beta \) is a degradation rate constant.
- \( t \) is time.

This exponential decay model suggests that the remaining capacity decreases over time, and the degradation rate is determined by the constant \( \beta \). More sophisticated models can involve multiple parameters and factors. The Flowchart for Capacity Degradation Modeling is shown in Figure 1.

![Flowchart for Capacity Degradation Modeling](image)

**Figure 1** Flowchart for Capacity Degradation Modeling

5. Design and validation

In the context of lithium-ion batteries, the validation of a capacity degradation model is an essential step that must be taken into consideration in order to guarantee that the model can be used to real-world scenarios. The subsequent phase entails verifying the performance of the model with data from lithium-ion batteries that are collected from the actual world. This comes after the model has been trained and fine-tuned using historical data. This method of validation...
evaluates how effectively the model generalizes to examples that have not been encountered before, so offering insights into the correctness and dependability of the model when it is applied to settings that are practical and daily.

As new data becomes available over time, it is vital that changes and improvements be made to the model on a consistently ongoing basis. This technique, which is iterative, guarantees that the model will continue to be adaptable to the shifting patterns and developing situations that are present in lithium-ion batteries. By incorporating new data, the model is able to maintain its relevance and accuracy, allowing it to account for any changes in the behavior of batteries, patterns of usage, or external variables that influence capacity degradation. The performance of the model is monitored against real-world data on a regular basis, which enables stakeholders to make educated decisions on battery maintenance, replacement methods, and overall system optimization. This, in turn, contributes to the lifetime and efficiency of applications that use lithium-ion batteries. Figure 1 and 2 shows the Schematic diagram and working model of proposed system.

The proposed system uses a feature-based approach in our early prediction model. Here, features are created as linear or nonlinear modifications of the raw data and then incorporated in an elastic net, a regularized linear model. By linearly integrating a subset of the proposed features, the final model predicts the logarithm of cycle life. Our choice of a regularized linear model allows us to suggest characteristics specific to a domain with different levels of complexity while keeping interpretability high. Moreover, online prediction with linear models only needs a single dot product operation following data preparation, making them computationally efficient. It's possible to train the model offline.

The proposed system uses the properties of lithium-ion batteries, including initial discharge capacity, charge time, and cell can temperature, whilst remaining neutral to chemistry and degradation mechanisms. To capture the electrochemical evolution of individual cells during cycling, a number of features derived from the discharge voltage curve are computed. Our attention is on Q (V), the discharge voltage curve as a function of voltage for a given cycle and how it changes over time. As the voltage range is fixed for every cycle, we consider capacity as a function of voltage rather than voltage as a function of capacity to provide a consistent basis for cycle comparisons. For instance, we may look at the equation ΔQ30-20(V) = Q30 (V) - Q20 (V), which represents the difference in discharge voltage curves between the 20th and 30th cycles. The cycle number is indicated by the subscript. This transformation, ΔQ (V), is particularly significant since voltage curves and their variations offer a rich data source helpful for diagnosing degradation.
To be more precise, the 100th and 10th cycles are used to depict the $\Delta Q(V)$ curves in our dataset as $\Delta Q_{100-10}(V)$. These cycle numbers have a deeper explanation that will come later. Summary data, including variance, mean, and minimum, are then presented for each cell’s $\Delta Q(V)$ curve. The difference in voltage curves between two cycles is represented by each summary statistic, which is a scalar variable. We select these summary statistics based on their prediction power rather than their literal interpretation in our data-driven methodology. When cycle life is compared with a summary statistic (variance) applied to $\Delta Q_{100-10}(V)$, a distinct pattern becomes evident right away.

Because of the high predictive capacity of features based on $\Delta Q_{100-10}(V)$, we investigate three different models using:

- only the variance of $\Delta Q_{100-10}(V)$,
- additional potential features acquired during discharge, and
- features from extra data streams such as temperature and internal resistance, due to the high predictive power of features based on $\Delta Q_{100-10}(V)$.

In all cases, data from only the first 100 cycles are used. These three models, each having progressively more candidate features, were chosen to evaluate the trade-off between the cost of collecting new data streams and prediction accuracy thresholds. The training data from 41 cells are used to select the model features and establish the coefficient values, and the primary testing data from 43 cells are used to evaluate the model’s effectiveness. After model creation, we further test the model using a secondary testing dataset of 40 cells. Our predictive accuracy is gauged using two metrics: average percentage error (APE) equation (4), which measures variance, and root-mean-square error (RMSE) equation (5), which is expressed in cycles. These metrics are described in the section on developing machine learning models. The cloud base is used to build the data sheet.

$$APE = \frac{1}{n} \sum_{i=0}^{n} \frac{y_i - \hat{y}_i}{y_i} \times 100 \quad (4)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=0}^{n} (y_i - \hat{y}_i)^2} \quad (5)$$

Where:

- $n$ is the number of observations.
- $y_i$ is the actual value for observation $i$.
- $\hat{y}_i$ is the predicted value for observation $i$.

In the proposed system:

- A battery, when connected to a load, effectively carries a nominal voltage of 3.7 volts.
- The battery has a capacity of 2500mAh, it would require 1 hour to discharge 1000mAh and 2 hours and 30 minutes to fully discharge 2500mAh at a steady current of 1A.
- The battery is considered fully discharged once the discharge voltage reaches 3 volts.

Figure 4 shows the voltage profile of the battery for various conditions. Once datasets for various voltages are prepared, the csv file is uploaded to the machine learning regression model. Since it is a regression model utilizing raw data, a prospective voltage range of 2.75 to 9.7 is provided during feature creation. At 3 volts, the capacity is 2000mAh after 2 hours, and at 3.7 volts, its 2500mAh. By forecasting the voltage, the expected capacity can be determined through linear interpolation.

The battery life cycle is estimated by linearly extrapolating the discharge capacity to the life cycle. At the 60th minute of an hour, the machine learning model is designed to forecast the anticipated battery voltage level. Below is a linear interpolation of the expected voltage to the discharge capacity:

$$X-axis: \ x_0 = 2.9, x_1 = 2.75$$

$$Y-axis: \ y_0 = 2500, y_1 = 2000$$
The life cycle can be estimated using the projected discharge capacity. Figure 5 shows the predicted remaining capacity for four batteries from proposed prediction algorithm.

Figure 4(a) The battery is connected for experiment with no load is applied

Figure 4(b) The battery is connected connected for external load i.e DC Motor

Figure 4(c) The battery is trying to regaining its energy after application of external load
The variation of discharge capacity with life cycles in machine learning-based prediction of lithium-ion battery life cycle for capacity degradation modeling refers to how a battery's ability to hold charge changes over time with repeated charging and discharging cycles. This phenomenon is analyzed using machine learning algorithms to develop predictive models for estimating a battery's remaining useful life based on its discharge capacity and cycle count. Understanding this variation is crucial for predicting battery performance and determining when maintenance or replacement is needed, ultimately improving the reliability of battery-powered systems.

6. Conclusion

This paper has examined a crucial and developing battery management system topic. A notable development in the subject is the incorporation of machine learning methods, such as Support Vector Regression (SVR), to forecast the life cycle.
The cycle of lithium-ion batteries. The thorough examination of numerous processes, such as feature engineering, data collecting, preprocessing, and model selection, highlights the integrated approach used to predict the complex behavior of these energy storage devices.

The strategic choice to employ SVR for capacity degradation modeling is justified by its capacity to handle non-linear connections and capture intricate patterns in the data. The proposed methodology's practical usefulness is further enhanced by the inclusion of real-world validation and ongoing model modification. With the advent of an era mostly dependent on renewable energy sources and electric vehicles, it is critical to accurately anticipate the life cycle and capacity deterioration of batteries. In addition to adding to our scientific knowledge of lithium-ion battery behavior, the insights offered in this study provide workable methods for enhancing battery management tactics and guaranteeing the longevity and effectiveness of energy storage systems in a range of applications. This research is interdisciplinary in nature, combining knowledge of machine learning and battery technology to put it at the forefront of developments that will shape the future of sustainable technology and energy storage.

The key outcomes of this work are:

- Implemented a linear regression model reinforced with elastic net regularization for battery cycle life prediction based on cycle metrics.
- Extracted tailored features from raw data and integrated them into a linear regression model fortified with elastic net regularization using training datasets.
- Optimized hyperparameters using a dedicated validation dataset.
- Assessed the model's performance using test data.
- Additionally, this model offers a clear view of a vehicle's battery conditions to both end-users and service providers.

Compliance with ethical standards

Disclosure of conflict of interest

The author discloses that there is no conflict of interest declare.

Reference

Author's short biography

Dr. S. Kumarappa, working as Professor in the Department of Mechanical Engineering & Dean of the Institute, Bapuji Institute of Engineering & Technology, Davangere-577 004, Karnataka, India. He received his Ph.D. from Kuvempu University, M. Tech degree from NITK, Surthkal and B.E. from Mysore University. He is having 30 years of experience in Teaching and 22 years of Research experience. He has served as Head of Department of Mechanical Engineering for five years, worked as Coordinator of IGNOU for four years, and PG Coordinator of Thermal Power Engineering for five years at BIET. His research interests are in the field of Alternative fuels for IC Engines, Engine Management system, CFD and he has published around 70 research papers in various International and National Journals, and conferences. He has guided 55 Under Graduate and 32 Post Graduate projects and actively involved in Research.

Manjunatha H.M. received the B.E. degree in electrical and electronics engineering in 2009, the M.Tech. degree in computer application in industrial drives in 2012 and Ph.D in 2022 from Visvesvaraya Technological University, Belagavi, Karnataka, India. He is currently working as Assistant Professor at Bapuji Institute of Engineering and Technology, Davanagere affiliated to Visvesvaraya Technological University, Belagavi, Karnataka, India. His main research interests are in Power System Operation and Control, Distribution Automation, Artificial Intelligence tools applied to Smart-Grid, Energy Trading, etc.