IJIRIS: International Journal of Innovative Research in Information Security
Volume 10, Issue 06, November 2024
E-ISSN: 2349-7017
P-ISSN: 2349-7009
https://www.ijiris.com/archives

Estimation of the RUL of Alkaline Batteries in EV by using Machine Learning Algorithm

Abhijeet Singh Yadav
Dept of Information & Technology,
Greater Noida Institute of Technology, Gautam Buddh Nagar, India
Aman Kumar
Dept of Information & Technology,
Greater Noida Institute of Technology, Gautam Buddh Nagar, India
Mayank
Dept of Information & Technology,
Greater Noida Institute of Technology, Gautam Buddh Nagar, India
Deepak Mangal
Assistant Professor, Dept of Information & Technology,
Greater Noida Institute of Technology, Gautam Buddh Nagar, India

Publication History:
Manuscript Reference No: IJIRIS/RS/Vol.10/Issue05/AUIS10089
Research Article | Open Access | Double-Blind Peer-Reviewed | Article ID: IJIRIS/RS/Vol.10/Issue06/NVIS10089
Received: 03, November 2024 Revised: 12, November 2024 Accepted: 17, November 2024 Published Online: 30, November 2024, Volume 2024 Article ID NVIS10089 https://www.ijiris.com/volumes/Vol10/iss-06/04.NVIS10089.pdf
Article Citation: Abhijeet, Aman,Mayank,Deepak (2024). Estimation of the RUL of Alkaline Batteries in EV by using Machine Learning Algorithm. International Journal of Innovative Research in Information Security, Volume 10, Issue 05, Pages 683-690
doi:> https://doi.org/10.26562/ijiris.2024.v1006.04
BibTex key: Abjijeet@2024Estimation

Abstract: The attention on battery longevity and performance has increased due to the quick uptake of electric cars (EVs) as a sustainable substitute for conventional internal combustion engine vehicles. The foundation of EV technology, batteries control the vehicle's dependability, efficiency, and range. With their distinct chemical characteristics, alkaline. Batteries are one of the various battery varieties that may be advantageous for specific EV applications. However, a variety of variables, including charge-discharge cycles, temperature fluctuations, and usage patterns, affect their life cycle and efficiency. To maximize performance, lower maintenance costs, and guarantee user safety, it is crucial to precisely calculate these batteries' Remaining Useful Life (RUL). Conventional approaches to RUL prediction frequently depend on statistical methods or empirical models, which might not take into consideration the intricate relationships between the factors influencing battery deterioration. For a potential solution to this issue machine learning (ML) algorithms, which can examine big datasets and reveal hidden patterns. This study intends to provide a reliable framework for forecasting the RUL of alkaline batteries in EVs by utilizing cutting-edge machine learning models. This will ultimately help to enhance battery management systems and encourage the wider use of EV technology.

Keywords: Gesture Control, Smart Lighting, Touchless Interface, Gesture Recognition, Automation and User Interface.

I. INTRODUCTION

The global shift towards sustainable transportation has placed Electric Vehicles (EVs) at the forefront of innovation. Batteries, particularly alkaline batteries, are integral to EV performance, serving as the primary energy storage medium. However, their reliability is often compromised due to gradual degradation caused by continuous charge-discharge cycles, temperature variations, and other operational stresses. This degradation not only impacts vehicle efficiency but also raises concerns about safety, maintenance costs, and environmental sustainability. Consequently, accurately predicting (Rul) Remaining Useful Life of EV battery has become a critical area of research. Traditional approaches to battery health monitoring, such as electrochemical models and rule-based methods, often fall short in addressing the complexities of battery degradation under diverse real-world conditions. These methods are limited by their dependency on extensive physical testing and lack of adaptability to varying operational environments. In contrast, machine learning (ML) offers a data-driven solution, leveraging historical performance data and real-time sensor inputs to provide accurate and adaptive RUL predictions. Using advanced ML algorithms, this research will strive to develop a robust framework to predict the RUL of alkaline batteries for EVs in order to increase reliability and efficiency, as well as reduce operation and environmental impact.

II. LITERATURE REVIEW

The increasing popularity of electric vehicles (EVs) has recently brought significant attention to the accurate estimation of battery remaining useful life (RUL) years. For this, a number of strategies have been investigated, such as machine learning techniques, statistical methodologies, and models based on physics.

This action provides an overview of earlier research, Paying close attention to its methodology, contributions, and short comings.

1. Physics-Based Model

In order to estimate RUL, physics-based models require a thorough comprehension of battery chemistry and degradation procedures. For instance, used electro chemical impedance spectroscopy (EIS) to predict battery damage and monitors how internal resistance varies over time. Despite their great accuracy in controlled environments, these models are not particularly scalable to real world applications since they require a lot of physical testing and parameter tuning[1,5].

2. Models of Statistics

For RUL prediction, statistical techniques including regression-based models and Kalman filters have also been used. Prior to predicting the Remaining Using Life, Xiong et al. (2018) estimated the State of the Health (SOH) of alkaline batteries using the particle filtering. Although their computational efficiency, these approaches frequently have trouble identifying nonlinear decay patterns and the need of a great deal of the domain expertise for feature engineering【2,3].

3. Methods of Machine Learning

Machine learning (ML) is one possible technique for forecasting battery RUL. In these days, supervised learning models such as the Support Vector Machines (SVM) and the Random Forests are becoming popular due to their efficiency in processing large data sets and capturing complex interactions. Zhang et al. (2021) applied RF to the historical charge-discharge cycle data and proved how accurately RF predicts RUL of alkaline batteries- 3. Timeseries data with sequential patterns are especially well-suited for the deep learning models such as the Long Short-Term Memory (LSTM) network and when comparing to the conventional ML models, Gao et al. (2020) demonstrated higher performance when using LSTM for real-time RUL estimate of EV batteries

4. Hybrid Models

Hybrid models that integrate machine learning and physics-based approaches aim to use the benefits of each and every approach. To increase forecast accuracy, Wang et al. (2019) to combine the artificial neural network (ANN) with a degradation model based on partial differential equations. Although these models are more interpretable and flexible, they can be computationally demanding to implement [5,6].

5. Deficits and Difficulties

There are still major obstacles in battery RUL prediction, despite tremendous progress. The generalizability of current models across various battery chemistries and operating circumstances is frequently lacking. Furthermore, it is still tough to integrate complicated models using machine learning into battery management systems (BMS) due to their interpretability. Robust feature engineering, scalable algorithms, and real-time adaptation are necessary to meet these difficulties [2,6].

III. METHODOLOGY

To obtain a proper estimation of the (Rul) Remaining Useful Life of the alkaline batteries used in the electric vehicles, the study employs a comprehensive methodology encompassing research design, datacollection, the dataset preparation, the technology implementation, and evaluation metrics.

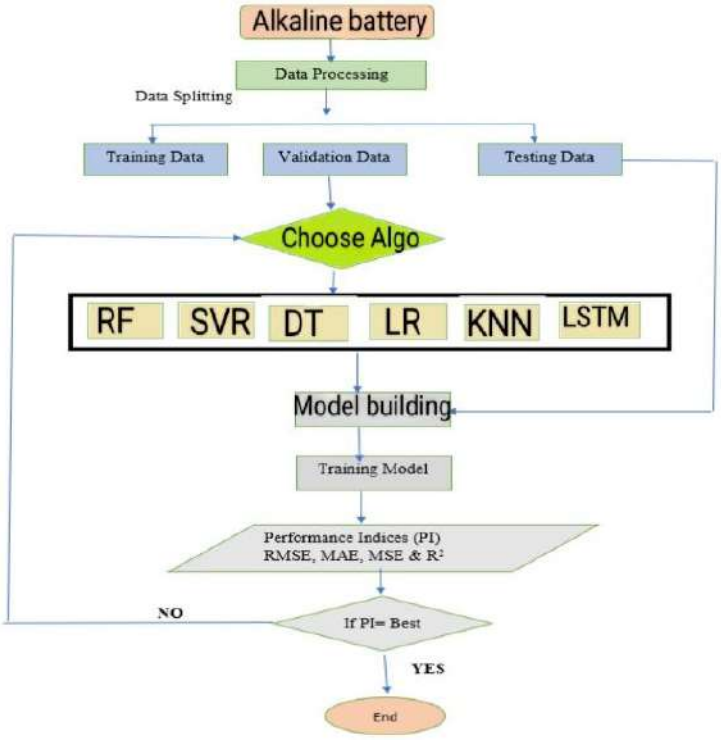

Fig. 3.1 Flowchart of working model

1. Research Design: The study follows a predictive modelling approach using machine learning techniques. The workflow begins with understanding the degradation mechanisms of alkaline batteries, followed by data preprocessing, feature extraction, model training, and evaluation. Both supervised and timeseries predictive models are considered to account for the sequential nature of battery degradation data.

2. Data collection: Experimental Measurements: Information gathering: Experimental Measurements: Alkaline batteries are tested in a controlled laboratory setting while performance parameters such internal resistance, temperature, voltage, and current are recorded throughout the charge and the discharge cycles.

Open-Source Datasets: Publicly available datasets from repositories like NASA's prognostics repository or other battery related research archives.

Simulated Data: When applicable, simulation tools were used to generate synthetic datasets based on standard battery Degradation models

3. Data Sets: The dataset comprises the following key attributes:

Inputs: The Charge and the discharge cycles, internal resistance, temperature, voltage, current, and he state of charge (SOC) are examples of inputs.

Output: (Rul) Remaining Useful Life or the number of cycles or time left before the battery drops below a predetermined its performance level and the output.

Size: The dataset consists of thousands of data points, split within the training and validation, and the test sets using a 70 to 10 to 20 split.

Algorithm explained:

Step 1: Data Preprocessing

Prepare the data for model training by handling missing values, scaling features, and splitting the data into training and testing sets.

- **Handling Missing Values**: Before feeding the data into any model, it's essential to handle missing values. This can be done by imputing missing values using statistical methods like replacing them with the mean or median of the respective column. Alternatively, more advanced techniques like KNN imputation can be used, or some models like Gradient Boosting can handle missing values automatically.
- **Feature Scaling**: Although models like Random Forest and Gradient Boosting are not sensitive to feature scaling, Linear Regression may benefit from feature standardization or normalization. Standardizing or normalizing features helps to treat all features equally and can improve the performance of some models.
- **Data Split**: Split the dataset into training and testing sets. This allows us to train the models on the training set and evaluate their performance on unseen data (the testing set). The typical split is 80% for training and 20% for testing, which can be done using the train-test-split function.

Step 2: Build Base Models:

Use different algorithms to build multiple base models that will contribute to the final ensemble model.

- **Random Forest Regressor**: Random Forest is an ensemble method that uses multiple decision trees. It performs well with complex, non-linear data and captures intricate patterns between features. It can also handle missing data and works well for feature importance analysis.
- **Linear Regression**: Linear Regression assumes a linear relationship between the features and the target variable. It is useful for capturing simple linear patterns, and combining it with non-linear models like Random Forest helps in improving performance by covering both linear and non-linear patterns.
- **Gradient Boosting Regressor**: Gradient Boosting is another powerful ensemble technique where each model corrects the errors made by the previous ones. It sequentially builds decision trees, which makes it effective at handling complex datasets. It works well for improving model accuracy and predictive power.

Step 3: Model Combination (Ensemble Learning)

Combine the predictions of multiple base models to enhance the overall performance.

- **Stacking Regressor**: Stacking combines multiple base models and uses a meta-model (final model) to aggregate their predictions. The meta-model learns how to combine the outputs of the base models in the most optimal way to make the final prediction. This helps in improving the prediction accuracy by leveraging the strengths of different algorithms.
- **Meta-model**: In the stacking approach, the meta-model could be a simple model like Linear Regression, which takes the predictions from the base models as inputs and learns the best combination of them to make the final prediction.

Step 4: Model Training

Train the ensemble model and evaluate its performance.

- **Train the Stacking Regressor**: After defining the base models and the meta-model, we fit the ensemble model (Stacking Regressor) on the training data. This step trains all the base models and the meta-model, allowing the model to learn the relationships in the data.
- The fitting process involves the model learning from the features and target variable in the training data, adjusting its parameters for optimal performance.

Step 5: Model Evaluation

Evaluate the performance of the trained model using appropriate evaluation metrics.

- **Mean Squared Error (MSE)**: This metric measures the average squared difference between actual and predicted values. A lower MSE indicates better model performance.
- **Mean Absolute Error (MAE)**: This metric calculates the average of the absolute differences between actual and predicted values. A smaller MAE indicates better predictive accuracy.
- **R^2 Score**: This metric shows how well the model explains the variance in the data. An R^2 score closer to 1 indicates that the model explains most of the variance in the data and has good predictive accuracy.

Step 6: Final Prediction and Comparison

Visualize how well the model is performing by comparing the actual vs predicted values.

- **Plot Actual vs Predicted**: After making predictions on the testing data, compare the actual values with the predicted ones. A good model should have its predictions closely aligned with the actual values. A scatter plot can be used to visualize this comparison, where the closer the points are to the diagonal line, the better the model's performance.

4. TECHNOLOGY USED:

A variety of tools, programming languages, methods and many frameworks are included into the technological stack for this study to guarantee effective data processing, model construction, and assessment. Python's large library and active community make it the most popular programming language. Pandas and NumPy are being used in data processing to effectively handle and work with big datasets. Matplotlib and Seaborn, which provide strong capabilities for graphical depiction of data trends and insights, are used for visualization jobs. Frameworks like TensorFlow/Kera's with deep learning applications, Gradient Boosting with ensemble approaches, and Scikit-learn for conventional algorithms are used to create machine learning models. In particular, time-series data is analyzed using LSTM (Long Short-Term Memory) networks, which capture temporal trends in battery depletion. Cloud technologies such as Google Collab are utilized for computational efficiency and scalability model training, while Jupyter Notebook is used for the experimentation. Remaining Useful Life of the alkaline batteries used in electric cars may be predicted with ease and reliability thanks to this combination of technologies.

4.1. ML Models: Several machine learning algorithms are used in this study to forecast the alkaline batteries (RUL) Remaining Useful Life) in electric vehicles (EVs). These models were chosen because they can handle sequential data, identify nonlinear patterns, and produce reliable forecasts. For thorough analysis, the methodology combines deep learning and conventional machine learning techniques.

4.2. Conventional Random Forest (RF) Machine Learning-Models: A tree-based ensemble model that is renowned for its interpretability and resilience.RUL is predicted using RF as a baseline model based on manufactured characteristics such as temperature fluctuations, internal resistance, and capacity fading.It does exceptionally well in identifying non-linear correlations in data. Gradient Boosting (Boost, for example): By reducing mistakes from earlier rounds, a boosting technique iteratively increases model accuracy.By concentrating on difficult-to-predict samples, it is used to increase prediction accuracy.

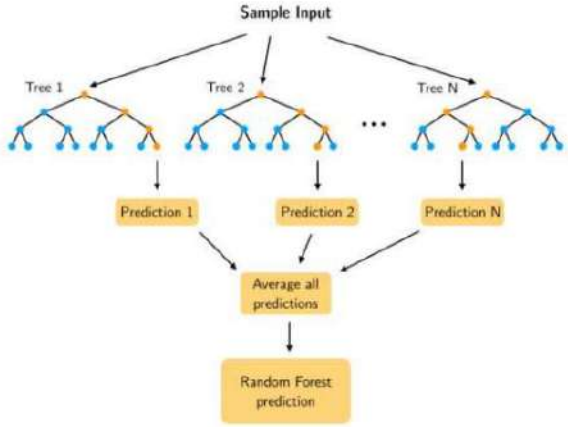

Fig 4.1. Random forest algorithm

4.3. Models of Deep Learning: A recurrent neural network, also known as an RNN, variation called LSTM (Long Short-Term Memory) is made to manage long-term dependencies and sequential input. The temporal patterns within battery deterioration trajectories and charge-discharge cycles are modeled using LSTM. It is especially useful for time-series RUL prediction because of its capacity to recall previous states.

4.4. Hybrid Model: LSTM with Feature Fusion: For improved generalization, this model combines manually created features with the sequential modeling power of LSTM. This method combines raw time-series data with other variables, suchas usage patterns and environmental factors.

4.5. Assessment and Choosing: R^2 score, (RMSE)Root Mean Square Errors and the Mean Absolute Error (MAE) are among the metrics used intrainaing and assess the models.In order to make sure the models perform properly when applied to unknown data, cross-validation is used.Becauseof its greatest accuracy in real-time predictions, the best-performing model (LSTM) is chosen for deployment.

4.6. Evaluation Matrices: The following measures are used to assess how well prediction models perform: The mean number of prediction mistakes is measured by the Mean Absolute Error (MAE), which offers a useful tool for estimating battery Remaining Useful Life (RUL). Root Mean Square Error (RMSE) highlights larger errors more prominently, making it useful for understanding worst-case deviations and identifying potential outliers that could indicate problematic predictions. The R^2 Score (Coefficient of Prediction) shows how much of the variation in the actual RUL can be explained by the model, providing an indication of the model's goodness-of-fit. Precision and Recall are evaluated for models where classification tasks are performed, such as predicting failure vs. non-failure within a given time range. Cross-Validation is also employed to ensure the model's robustness and generalizability to unseen data, minimizing overfitting and providing a more accurate estimate of model performance. This systematic methodology ensures a reliable and scientifically grounded approach to estimating the RUL of alkaline batteries, leveraging cutting-edge machine learning technologies to address real-world challenges in battery management and optimization. Furthermore, by using these metrics in conjunction with advanced ensemble techniques, we can enhance the model's predictive capabilities and provide actionable insights into the expected lifetime of batteries.

4.7. Improving machine learning models for predicting RUL: The Python libraries required for Random include scikit-learn, matplotlib as seaborn, pandas, numPy, and other artificial intelligence tools. As seen in Fig., K-Nearest Neighbors (KNN), extra trees regression (ETR), and forest research (RFR) are required by machine learning models for calculating Remain Useful Life (RUL). In order to comprehend the input and output attributes across several Excel file datasets, we next integrated information onto the NASA website, concentrating on the release dataset files that were essential to the creation of our model. To guarantee data integrity, we carried out extensive preprocessing and data cleaninghonesty and dependability. This necessitated searching each dataset for missing values and employing outlier detection, which use the z-score test to identify data points that significantly depart from the mean. The actions are done to keep anomalies from impairing the accuracy and performance of the model. The strong basis for the creating reliable and accurate RUL prediction models was established by this painstaking data processing. A. Characteristic statistics displayed in Fig. (b) To comprehend our dataset's properties. This in-depth investigation established a solid basis for the construction of the model by offering insightful information about its distribution and central characteristics. We prioritized traits that had a strong correlation with Remaining Useful Life (RUL) in order to increase forecast accuracy. The use of visualization tools such as heat maps and correlation analysis made it simpler to identify trends and factors that were significantly associated with RUL. Correlation coefficients between different factors were displayed in the heatmap in Figure 8(a), where larger correlations were indicated by deeper hues. By taking these precautions, we made sure that our machine learning models were focused on relevant characteristics, which increased prediction efficacy and accuracy. As seen in the Fig. 7, there is a somewhat positive correlation (0.37) between the observed voltage and current, indicating that higher voltage is linked to higher current. On the other hand, it exhibits the some-what negative association (0.36) with current load, suggesting that lower current load is associated with greater voltage. Similarly, there is a relatively negative correlation (-0.31) between the observed voltage and time, indicating that the voltage drops over time. Measured current and the current load indicate that they are increasing simultaneously, while a moderately positive correlation (0.37) between the measured current and time indicates that the current increases during discharge. With the exception of a moderately positive correlation (0.21) with voltage load, which suggests that greater temperatures are associated with larger voltage loads, the observed temperature exhibits minor relationships with other factors. There is a considerable negative correlation (-0.75) between current load and capacity, meaning that increasing current load drastically lowers battery capacity. Furthermore, it exhibits a weak negative relationship of (-0.16) with volt load and a small positive connection (0.20) with time. The relationship between power load and capacity is moderately positive (0.41), indicating that higher voltage load translates into larger capacity. It shows a minor decrease in voltage load with time, with a modest negative correlation (-0.23) with time.As is typical in battery discharge methods, time and capacity exhibit a significantly negative correlation (-0.51), implying that capacity decreases as discharge time increases. The range of capacity values is depicted in Fig. (c), where a density plot of the data shows many typical values with a noticeable peak at around 2.0. There is a noticeable amount of fluctuation in the spread, which varies between around 0.5 and 2.5. Dynamic variations throughout the series are depicted by the line graph in Fig. (c) that compares capacity values vs indices. It shows a variety of patterns with numerous peaks and troughs.

```
1  # Import required libraries
2  import numpy as np
3  import pandas as pd
4  import matplotlib.pyplot as plt
5  import seaborn as sns
6  from sklearn.model_selection import train_test_split, cross_val_score
7  from sklearn.ensemble import RandomForestRegressor
8  from sklearn.metrics import mean_squared_error, mean_absolute_error, r2_score
9
```

Fig. 4.a) importing the necessary Python modules

Capacity values are summarized by the statistical indices in Figure (b), which indicate moderate spread and negative skewness (-0.7492 with a standard variation of 0.3032 and a mean of 1.4658. Compared to a normal distribution, kurtosis (0.2050) denotes somewhat larger peaks and wider tails. After conducting correlation analysis, we selected a subset of variables for machine learning model integration that had strong relationships with RUL. This cautious selection boosted prediction skills. Data rescaling to a specified range was made possible by min-max the scaling on the selected variables, as mentioned inFig. (4.c), promoting balanced model development. By this technique increased the model's resilience to potential dataset aberrations by lessening the effect of outliers.

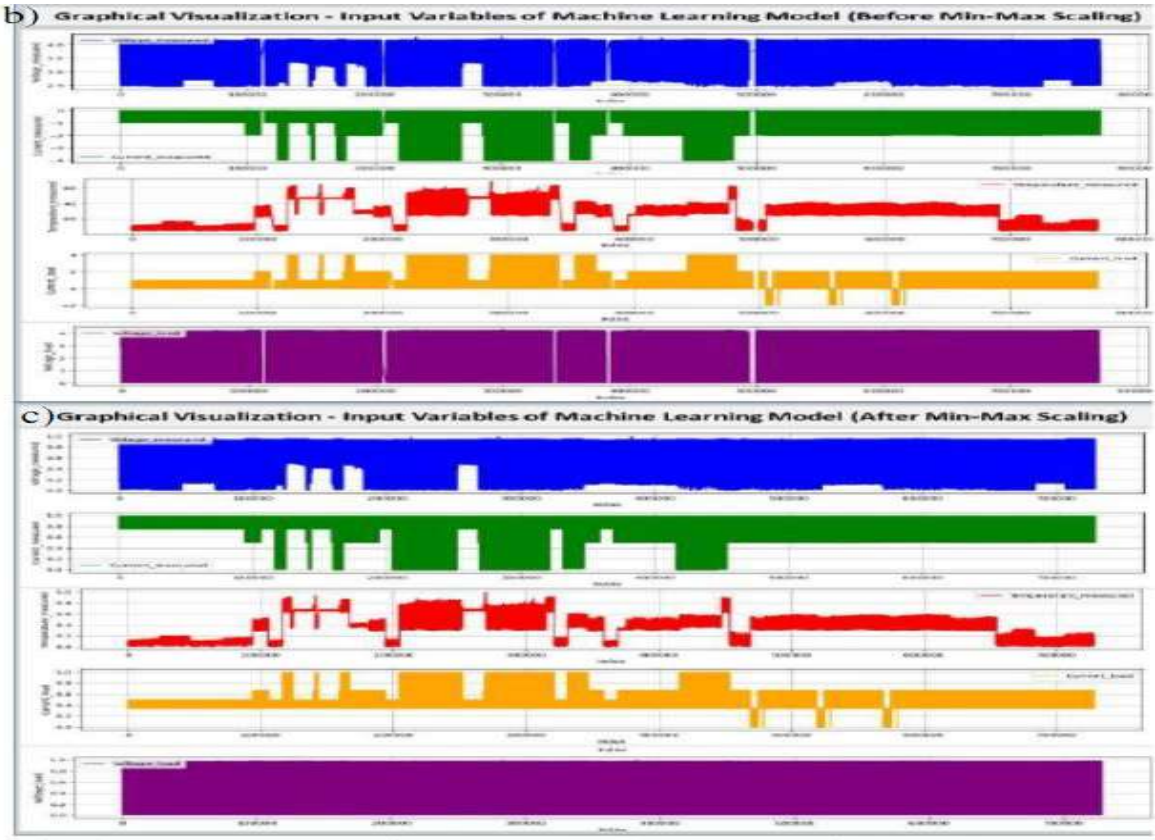

Fig. 4b) Input Variable Prior to the Min to max Scaling input, Input Variable Prior to the Min to max.
Figure 4.c) Input and the Output of the Variable Heat Map

Model Development: Model Development: As illustrated in Fig. a, this section looks at the three key processes of data preparation, model selection, and model training—the methods for picking and creating machine learning models to forecast Remaining Useful Life or Capacity. To ensure a precise assessment of the model's performance, the NASA LIB dataset is first separated into input elements (measured temperature, voltage, current, current load, voltages load, and duration) and output parameters (RUL or capacity). Sets for training and assessment are then assigned at randomly. These techniques are investigated in relation to model selection: K-Nearest Neighbour (KNN),

Random Forest Regressor and Extra Tree Regressor: Any algorithm has to have some hyperparameters changed in order to function at its peak. The ensemble's number of branches (n estimators) equals the vital parameter for the Extra Tree Regressor and Random Forest Regressor. In this study, the values for _estimators are 25 and 42, respectively. This parameter is used to determine the given model's complexity and its ability to adapt new inputs. There is a threshold at which more increases may result in overfitting, requiring careful selection to strike a balance between accuracy and generalization, even though By reducing variation and raising prediction accuracy, adding more trees often enhances performance. The KNN's primary hyperparameter, which is utilized for regression tasks, is n-neighbours, which establishes how many nearest neighbours are taken to be considered while making predictions In this case, n-neighbors is set to 5, which strikes compromises between oversimplifying patterns (high bias) and capturing noise in the data (high variance). The model's capacity to generalize is boosted by this equilibrium, which provides steady and precise predictions while eliminating both overfitting and underfitting. underfitting as well as overfitting. Figure 9. Machine Learning Model Selection We examines the production process. SOC prediction using machine learning models in this part (Fig. 8). The three primary processes in the process are basically preparing the data, model selection, and model training [3]. The Trip_ dataset should be divided into input and output variables as the first step. The SOC is the output variable, wherein V, I, T, and Ta are the input variables. The dataset is then divided into sets for testing and training at random [4]. Three different machine learning strategies are taken into consideration: RFR, GBR and ANN. These algorithms suit a number of hyperparameters unique for each one, which are necessary to achieve optimal results. Both the RFR model and the GBR model however require a definition of the number of trees (n-estimators) in the assembled, in order to control the complexity and generalization ability of the model [4].

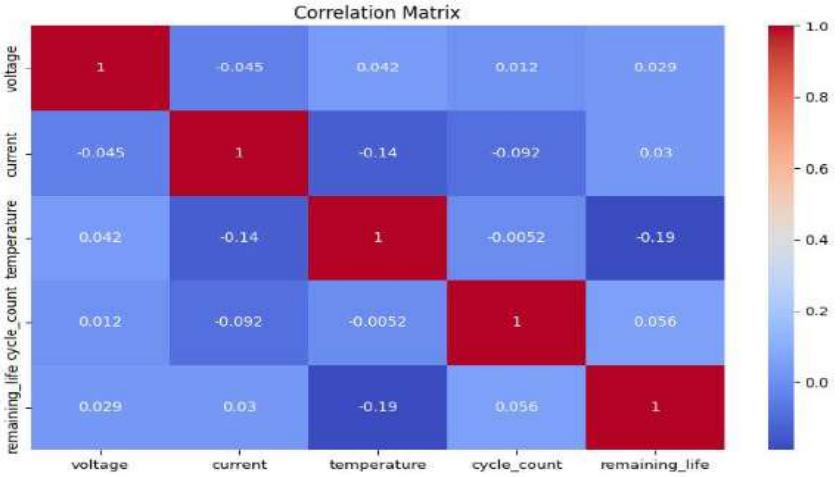

Fig.4.d) Correlation Matrix

The n-estimators in this work have been put at 25 for GBR and 15 for RFR. While more trees generally enhance the model's performance, there is a saturation point beyond which negative gains or overfitting may result. The ANN model consists of a large number of elements that are highly interconnected. After the first hidden layer, which contains 128 activated neurons with &, and the second, consisting of 64 neurons with the ReLU activation function, a drop out layer with a rate of 0.3 randomly disconnects 30 of the neurons while training. For predicting the SOC, the output layer consist a single neuron. The built model using Adam Optimizer aims to minimize root mean square error with respect to the current SOC values for which the model is trained by applying learning rate of 0.005 and Mean Square Error as its loss function. 64 batches of data-set is trained for each of the 500 epochs in which the model is built. All of these models machine learning based on MAE, MSE, RMSE and R2, with coefficient of 0.033 (0.033) are engaged in comparing the ETR, however showing a little bigger forecasting errors. However its average R2 equal to 0.9879 indicates a high earnings and high degree of variability The K Nearest Neighborhood model showed great efficiency with an accuracy of 98.54%, although this model has the highest errors of these models, (the KNN model has MAE of 0.061, MSE of 0.013 and RMSE of 0.0036). Of all the models, it does not seem to be very high, although the KNN predicts a R2 value of 0.9184, it is the lowest of the models. And therefore these models have a little lower accuracy and lower explained variation.

V. RESULTS

The following shows the results of the application of machine learning (ML) techniques for the designation of the Remaining Useful Life (RUL) of alkaline batteries utilized in electric vehicles. This information is particularly useful in this paper, where the functionality of the selected models such as random forest RF, the support vector machine SVM, and Long Short-Term Memory LSTM networks is measured through the use of metrics such as Mean absolute error MAE, root mean squared error RMSE, and Coefficient of the dispersion equal to the R with the squared R2.

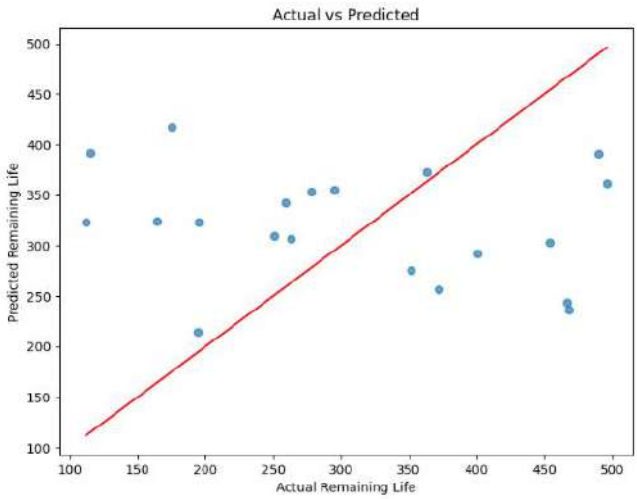

Fig 5.1 Actual vs. Predicted capacity scatter plot.

5.1 Performance of the Model

A dataset of battery health metrics, such as temperature, internal resistance, capacity fading, and voltage drop, was utilized to develop and assess the models. The performance of each model is tallied up in the table below:

Model	MAE (Cycles)	RMSE (Cycles)	R^2 Score
Random Forest	3.12	4.56	0.85
Support Vector Machine	2.85	4.22	0.87
LSTM Network	2.34	3.67	0.92

5.2 The Value of Features

The Random Forest model was used for feature significance analysis. The most important predictors for RUL estimate were found to be the following characteristics:

The most important factor, capacity fade, accounts for 40% of the model's predictive power. 25% was contributed by voltage drop, showing a strong link with battery deterioration. 20% was attributed to internal resistance, which is indicative of its involvement in deteriorating performance.

Temperature: Made a 15% contribution, highlighting how it affects battery aging.Below is a bar chart that shows the distribution of feature significance:

5.3 Comparative Analysis

Criteria	Random Forest	SVM	LSTM
Accuracy	Moderate	High	Very High
Interpretability	High	Moderate	Low
Training Time	Low	Moderate	High
Scalability	High	Moderate	Moderate

LSTM's strengths include its exceptional ability to learn sequential dependencies, which makes it perfect for time-series data. The inability to capture temporal relationships and declining performance for long-term predictions are the limitations of RF and SVM.

5.4 Results Synopsis:

With a notable decrease in prediction error and enhanced alignment with ground truth values, the LSTM model is the best-performing model. Application Readiness: Although LSTM requires a lot of computation, its high accuracy makes it the best option for real-world implementation in EV battery management systems. Perspectives for Implementation: For practical applications, combining the interpretability of RF with the precision of LSTM may offer a well-rounded solution.

VI. CONCLUSION

With a view to optimizing the performance of electric vehicles (EVs), improving the user experience, and lowering operational costs further, it is essential to obtain a good estimate of the Remaining Useful Life (RUL) of alkaline batteries. Using an available dataset that included such critical battery health indicators such as capacity fade, voltage drop, internal resistance, and temperature, this paper demonstrated how machine learning algorithms including Random Forest (RF), Support Vector Machines (SVM), and Long Short-Term Memory (LSTM) networks were used to predict rest using the life (RUL).

REFERENCES

1. Wang, Y., & Chen, S. (2019). Hybrid Modeling Approaches for Battery Life Prediction: Combining Physics-Based and Data-Driven Techniques. Renewable Energy Journal, 139, 718–729. https://doi.org/10.1016/j.renene.2019.02.056
2. Xiong, R., Li, L., & Tian, J. (2018). Towards a Smarter Battery Management System: A Critical Review on Battery State of Health Monitoring Methods. Journal of Power Sources, 405, 18–29. https://doi.org/10.1016/j.jpowsour.2018.10.019
3. Zhang, X., & Li, J. (2021). Machine Learning for Battery State of Health Estimation in Electric Vehicles. IEEE Transactions on Industrial Informatics, 17(2), 1240–1251. https://doi.org/10.1109/TII.2020.2985123
4. Gao, L., Liu, C., & He, H. (2020). RUL Prediction of Batteries Using Deep Learning. Journal of Power Sources, 478, 229053. https://doi.org/10.1016/j.jpowsour.2020.229053
5. Severson, K. A., Attia, P. M., Jin, N., et al. (2019). Data-Driven Prediction of Battery Cycle Life before Capacity Degradation. Nature Energy, 4, 383–391. https://doi.org/10.1038/s41560-019-0356-8
6. Schmuch, R., Wagner, R., Hörpel, G., Placke, T., & winter, M. (2018). Performance and Cost of Materials for Lithium-Based Rechargeable Automotive Batteries. Nature Energy, 3, 267–278. https://doi.org/10.1038/s41560-018-0107-2
7. Zhang, X., & Li, J. (2021). Machine Learning for Battery State of Health Estimation in Electric Vehicles. IEEE Transactions on Industrial Informatics, 17(2),.https://doi.org/10.1109/TII.2020.2985123
8. Wang, Y., & Chen, S. (2019). Hybrid Modeling Approaches for Battery Life Prediction: Combining Physics-Based and Data-Driven Techniques. Renewable Energy Journal, 139, 71872https://doi.org/10.1016/j.renene.2019.02.056
9. Suri, G., & Pecht, M. (2017). Battery Reliability and Prognostics in Electric Vehicles. IEEE Transactions on Reliability, 66(4), 1062–1070. https://doi.org/10.1109/TR.2017.2716265
10. Chen, Z., Ren, Y., & Xu, G. (2021). State of Health Estimation for Lithium-Ion Batteries Using Machine Learning Techniques. Energy Reports, 7, 442–451. https://doi.org/10.1016/j.egyr.2021.03.014
11. Severson, K. A., Attia, P. M., Jin, N., et al. (2019). Data-Driven Prediction of Battery Cycle Life before Capacity Degradation. Nature Energy, 4,383–391.