Open Access Research Journal of
Engineering and Technology
Journals home page: https://oarjpublication/journals/oarjet/
ISSN: 2783-0128 (Online)

OARJ OPEN ACCESS RESEARCH JOURNALS

(Research Article)

Check for updates

Prediction of battery charging time and range detection in electric vehicle using machine learning algorithms

K. M. Priya [1, *], Murali Krishnan. S [2] and Mohit. K [2]

[1] *Assistant Professor, Department of Electronics and Communication Engineering, Coimbatore Institute of Technology, Coimbatore 641014, Tamil Nadu, India.*
[2] *Department of ECE, Coimbatore Institute of Technology, Coimbatore 641014, Tamil Nadu, India.*

Open Access Research Journal of Engineering and Technology, 2025, 08(01), 044-056

Publication history: Received on 30 December 2024; revised on 13 February 2025; accepted on 16 February 2025

Article DOI: https://doi.org/10.53022/oarjet.2025.8.1.0026

Abstract

As a key pillar of smart transportation in smart city applications, Electric Vehicles (EVs) are becoming increasingly popular for their contribution in reducing greenhouse gas emissions. The lack of charging infrastructure and range detection is one of the most significant barriers to Electric Vehicle adoption. To address the problem of EV charging and range detection, employing Machine Learning (ML) algorithms to predict charging analysis, which is beneficial to drivers. ML algorithms in Electric Vehicle (EV) refers to the application of computational algorithms and statistical models that enable EVs and their supporting systems. To train the model own data set was created using hardware setup. This hardware setup consists of Battery, ESP Controller, Sensors, and Arduino IDE. Here two type of machine learning algorithms where we used labelled dataset to train the model or algorithms. one is the Support Vector Regression (SVR), other one is Random Forest Algorithm (RFA). It predicts charging patterns, recommend optimal charging times, and estimate charging times based on factors like battery health, current charge level, and environmental conditions. This leads to more efficient charging and better energy management and also going to predict the driving range of EVs based on performance metric like RMSE (Root Mean Squared Error), MAE (Mean Absolute Percentage Error), R Square, SMAPE (Symmetrical Mean Absolute Percentage Error). This information aids drivers in planning routes and making informed decisions. ML in the Electric Vehicle industry transforms data into actionable insights that enhance vehicle performance, optimize energy consumption, and improve user experiences.

Keywords: Electric Vehicles; Machine Learning; Support Vector Regression; Random Forest; Performance Metrics

1. Introduction

As a key pillar of smart transportation in smart city applications, EVs are becoming increasingly popular for their contribution in reducing greenhouse gas emissions. The lack of charging infrastructure and range detection is one of the most significant barriers to Electric Vehicle adoption. Machine Learning (ML) in Electric Vehicles (EVs) refers to the use of computational algorithms and statistical models that enable EVs and their associated systems to learn from data and make informed decisions. Creating own data set using our hardware setup to implement ML models that can predict charging patterns, recommend optimal charging times, and estimate charging times based on factors like battery health, current charge level, and environmental conditions. This leads to more efficient charging and better energy management and predict the driving range of EVs. This information aids drivers in planning routes and making informed decisions. ML in the Electric Vehicle industry transforms data into actionable insights that enhance vehicle performance, optimize energy consumption, and improve user experiences.

It is important for drivers to have a prior knowledge about the amount of fuel remaining in the vehicle and the amount of distance it could travel with the remaining amount of fuel also these data should be accurate and reliable for hassle free travel. In the case of Electric Vehicles (EVs) fuel refers to the charge of the battery. Hence by using Machine Learning

* Corresponding author: K. M. Priya.

algorithms which guarantee accurate predictions is to be implemented on the data set created to predict the SoC of the battery and the range of distance can be covered with the remaining amount of charge left.

The inaccuracy in predicting charging duration, providing timely battery level updates, and offering reliable estimates of remaining travel distance for EVs. Machine Learning algorithms provide high levels of accuracy in prediction which provide actionable in sights.

The main objective of the paper is to implement Machine Learning algorithms to predict the SoC and remaining charging time and the range of distance to cover with the remaining amount of charge in the battery.

Shahriar, S et.al [1]. this paper described about the Machine learning algorithms for the prediction of session duration and energy consumption, they have mainly focused on utilizing historical charging data. In some cases, additional derived features such as vehicle information, charging location information and seasonal information were used the additional input features including weather, traffic and local events and observe its impact on the accuracy of charging behavior predictions. Junzhe Shiet et.al [6] This paper intends to develop a novel algorithm for estimating Remaining Charging Time (RCT) of Electric Vehicles (EVs). Consider charging accuracy in the Constant Current (CC) stage and the charging profile in the Constant Voltage (CV) stage. Mazhar, T et.al [2] this paper illustrates about the emphasize the growing significance in the automotive industry and their impact on the distribution grid and to advocate for the adoption of smart grid solutions and efficient management systems to address the challenges posed by the increasing number of EVs on the road. And explore the role of Information and Communication Technology (ICT) in the development of smart cities and how these technologies can address complex urban issues. NaitMalek, Y et.al [7] introduced the machine learning-based forecasting methods for the State-of- Charge (SoC) of Battery Electric Vehicle (BEV) batteries, aiming to improve the management of energy consumption and charging station utilization in real-world BEV scenarios. Koji Sato et.al [10] this paper intends Battery performance prediction techniques based on Machine Learning (ML) models and lithium-ion battery (LIB) data collected in the real world have received much attention recently. However, poor extrapolation accuracy is a major challenge for ML models using real-world data, as the data frequency distribution can be uneven. Here, investigated the extrapolation accuracy of the ML models by using own data set generated using hardware setup.

The rest of the paper is organized as follows: The proposed system is discussed in Section 2. Section 3 presents Machine Learning Algorithms. Results and Discussion in Section 4. The Conclusion is discussed in Section 5.

2. Proposed system

Machine learning algorithms are computational models that allow computers to understand patterns and forecast or make judgments based on data without the need for explicit programming. These algorithms form the foundation of modern artificial intelligence and are used in a wide range of applications, including image and speech recognition, natural language processing, recommendation systems, fraud detection, autonomous cars etc.

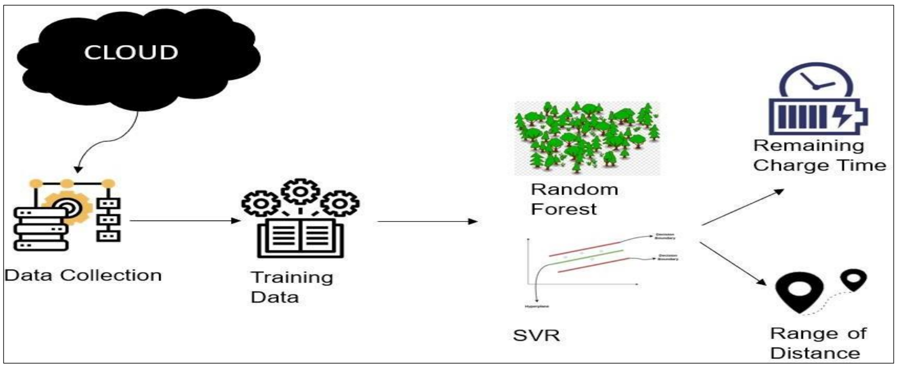

Figure 1 Machine Learning Model

Figure.1 describes Machine learning model is a mathematical representation or computational algorithm that is trained on data to perform a specific task without being explicitly programmed. Machine learning falls under the broader umbrella of artificial intelligence and involves the development of algorithms that can learn patterns and make predictions or decisions based on input data. Supervised learning is a type of machine learning algorithms where we used labelled dataset to train the model or algorithms. The goal of the algorithm is to learn a mapping from the input data to the output labels, allowing it to make predictions or classifications on new, unseen data.

2.1. Hardware Setup and Data Collection

Own datasheet was created using the below hardware setup model. As shown in Figure 2 the voltage and current is measured using sensors and displayed through Arduino Uno and transmitted to cloud via node MCU**.**

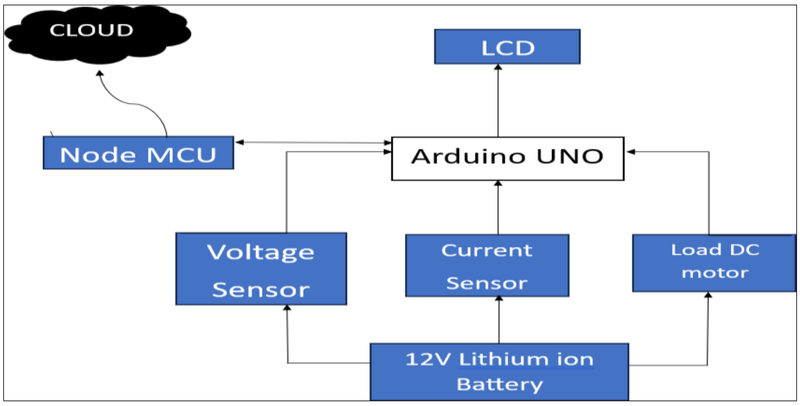

Figure 2 Block diagram of the hardware setup

The basic working principle involves lithium-ion battery is moving from the anode to the cathode during discharge and from the cathode to the anode during charging. The process is reversible, allowing for multiple charge and discharge cycles. Using lithium-ion 12V battery measuring voltage and current with respect to the load given and it is display in LCD using Arduino and data is transfer to cloud via node MCU. The ESP8266 is a low-cost, Wi-Fi-enabled microcontroller manufactured by Espress Systems. It is widely used in Internet of Things (IoT) applications due to its low cost, low power consumption, and Wi-Fi connectivity. Voltage Detection Sensor Module is a simple and very useful module that uses a potential divider to reduce any input voltage by a factor of 5.

ACS712 Module uses the famous ACS712 IC to measure current using the Hall Effect principle. The module gets its name from the IC used in the module, so the final products use the IC directly instead of the module. Blynk is a popular Internet of Things (IoT) platform that allows users to create custom applications for controlling and monitoring their IoT devices. It is designed to be simple and easy to use, even for those without a background in programming.

3. Machine learning algorithms

This section deals about the machine learning algorithms used for prediction of charging time and range detection. Machine learning algorithms are computational models that allow computers to understand patterns and forecast or make judgments based on data without the need for explicit programming. These algorithms form the foundation of modern artificial intelligence and are used in a wide range of applications, including image and speech recognition, natural language processing, recommendation systems, fraud detection, autonomous cars etc.

3.1. Support Vector Regression (SVR)

Support Vector Regression (SVR) is a type of machine learning algorithm used for regression analysis. The goal of SVR is to find a function that approximates the relationship between the input variables and a continuous target variable, while minimizing the prediction error. SVR can handle non-linear relationships between the input variables and the targetvariable by using a kernel function to map the data to a higher-dimensional space.This makes it a powerful tool for regression tasks where there may be complex relationships between the input variables and the target variable.

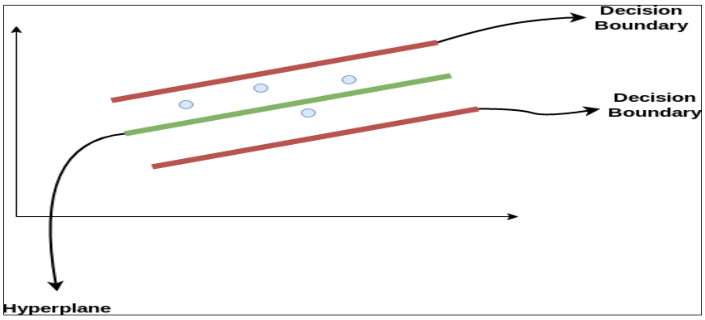

Figure 3 Support vector regression model

Figure 3 describes two red lines as the decision boundary and the green line as the hyperplane. The first thing understand is what is the decision boundary (the danger red line above!). Consider these lines as being at any distance, say 'a', from the hyperplane. So, these are the lines that we draw at distance '+a' and '-a' from the hyperplane. This 'a' in the text is basically referred to as epsilon.

Assuming that the equation of the hyperplane is as follows:

$$Y = wx + b \text{ (equation of hyperplane) –} \quad \ldots\ldots\ldots\ldots \ (1)$$

Then the equations of decision boundary become:

$$wx + b = +a \quad \ldots\ldots\ldots\ldots\ (2)$$

$$wx + b = -a \quad \ldots\ldots\ldots\ldots\ (3)$$

SVR should satisfy:

$$-a < Y - wx + b < +a \quad \ldots\ldots\ldots\ldots\ (4)$$

In SVR, the goal is to minimize the error between the predicted values and the actual values while satisfying a certain margin of error (epsilon). The margin of error allows some data points to be within a certain distance from the predicted values, which can be useful for dealing with noisy data or outliers. The parameter C controls the trade-off between maximizing the margin and minimizing the error. Overall, SVR is a powerful technique for regression analysis that can handle linear and nonlinear relationships between variables and can incorporate different types of kernels to capture complex patterns in the data.

Lee et al. introduced a novel dataset for non-residential EV charging consisting of over 30000 charging sessions. They used Gaussian Mixture Models (GMM) to predict session duration and energy needs by considering the distribution of the known arrival times. The testing dataset included the month of December 2018 and the reported Symmetric Mean Absolute Percentage Errors (SMAPEs) were 14.4% and 15.9% for the session duration and energy consumption, respectively.

In a different approach, ensemble machine learning using SVM, Random Forest (RF), and Diffusion-based Kernel Density Estimator (DKDE) was used for session length and energy consumption predictions. For training, historical charging records from two separate datasets were used, with one of them being public and the other being residential charging. The ensemble model performed better than the individual models in both predictions and the reported SMAPEs were 10.4% for duration and 7.5% for the consumption.

We selected all charging sessions from the ACN dataset that belonged to the 2019 calendar year, ensuring seasonal factors were considered during training. The dataset was split into 80% for model training and 20% for evaluation. During training, K-fold cross-validation was performed with training examples for testing. Model hyperparameters were determined using grid search, evaluating across 5 folds to expedite the process. The models were then evaluated using regression metrics.

In addition to that, Random Forest had the best cross-validation scores for energy consumption, Results from the test set showed that the random Forest performed best, although the improvement using ensemble learning for the prediction. User predictions were also inaccurate.

3.2. Random Forest Algorithm

Random Forest is a famous machine learning algorithm that uses supervised learning methods. It can be applied in both classification and regression problems. It is based on ensemble learning, which integrates multiple classifiers to solve a complex issue and increases the model's performance.Random Forest Algorithm widespread popularity stems from its user-friendly nature and adaptability, enabling it to tackle both classification and regression problems effectively. The algorithm's strength lies in its ability to handle complex datasets and mitigate overfitting, making it a valuable tool for various predictive tasks in machine learning.

Figure.4 demonstrates the Random Forest Tree which is used to combine N decision trees with building the random forest, and the second is to make predictions for each tree created in the first phase.

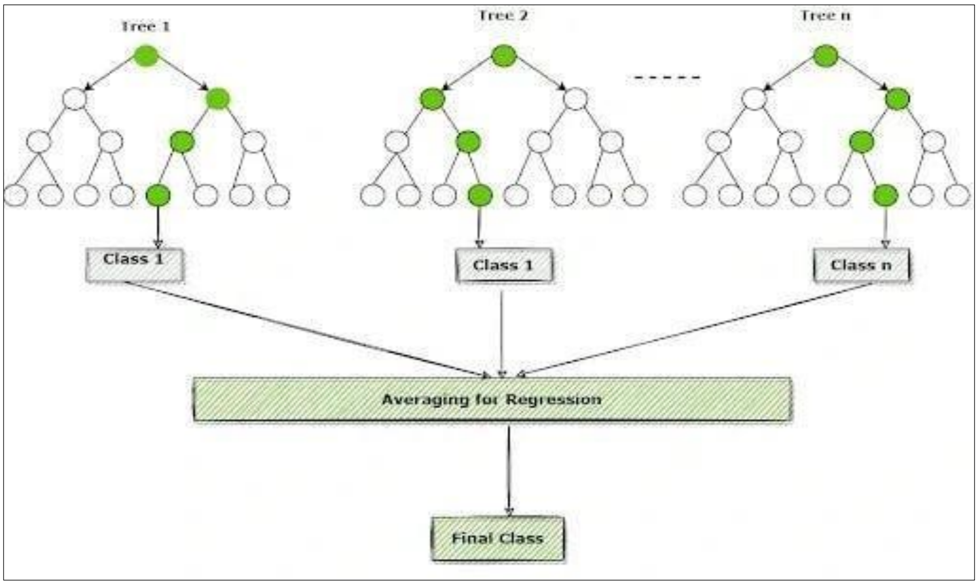

Figure 4 Random Forest Tree

The Random Forest Tree which it demonstrates that to combine N decision trees with building the random forest, and the second is to make predictions for each tree created in the first phase.

- Step 1: In the Random Forest model, a subset of data points and a subset of features is selected for constructing each decision tree. Simply put, n random records and m features are taken from the data set having k number of records.
- Step 2: Individual decision trees are constructed for each sample.
- Step 3: Each decision tree will generate an output.
- Step 4: Final output is considered based on Majority Voting or Averaging for Classification and regression, respectively.

It begins by importing the necessary libraries such as Pandas, Numpy and the Random Forest Classifier from the sklearn library. Then the dataset is read into a Pandas Data Frame and the features and target are separated into two separate variables (X and Y). Next, a Random Forest Classifier model is created and trained using the X and Y variables. Finally, the model is used to predict the target variable using the features.

The following steps form the basis for any machine learning workflow once we have a problem and model in mind:

- State the question and determine required data
- Acquire the data in an accessible format
- Identify and correct missing data points/anomalies as required

- Prepare the data for the machine learning model
- Establish a baseline model that you aim to exceed
- Train the model on the training data
- Make predictions on the test data
- Compare predictions to the known test set targets and calculate performance metric
- If performance is not satisfactory, adjust the model, acquire more data, or try a different modeling technique
- Interpret model and report results visually and numerically

3.3. Mathematical Model

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

- y_i : Actual value for the ith observation
- $\hat{y}_i$: Calculated value for the ith observation
- N : Total number of observations

The Root Mean Squared Error (RMSE) is one of the two main performance indicators for a regression model. It measures the average difference between values predicted by a model and the actual values. It provides an estimation of how well the model is able to predict the target value.

$$\boldsymbol{R^2} = \boldsymbol{1} - \frac{\sum_{i=1}^{n}(y_i - \overline{y_i})^2}{\sum_{i=1}^{n}(y_i - \overline{\mu})^2} \quad \text{………(6)}$$

R-squared is a statistical measure that represents the goodness of fit of a regression model. The value of R-square lies between 0 to 1. Where R-square equals 1 when the model perfectly fits the data and there is no difference between the predicted value and actual value.

$$\boldsymbol{SMAPE} = \left(\frac{1}{N} \sum \frac{|y_i - \hat{y}_i|}{(|y_i| - |\hat{y}_i|)/2} * \boldsymbol{100}\right) \quad \text{………… (7)}$$

SMAPE, or Symmetrical Mean Absolute Percentage Error, is one calculation that you can use to check the accuracy of forecasting methods.

$$\boldsymbol{MAE} = \frac{1}{n} \sum_{i=1}^{n} |y_i - \hat{y}_i| \quad \text{…………(8)}$$

Mean Absolute Error calculates the average difference between the calculated values and actual values.

4. Results and discussion

This an overview of the hardware setup and the results of prediction of battery charging time and range detection in Electric Vehicle using Machine Learning algorithms.

4.1. Hardware setup

Figure 5 shows the hardware setup of how ESP Controller, voltage sensor and current sensors are connected. The connection for the sensors is quite similar, using the pinout of the ESP controller the GPIO pins are defined in the Arduino IDE and are wired accordingly (+5V, GND, 5v5, A0 etc.).

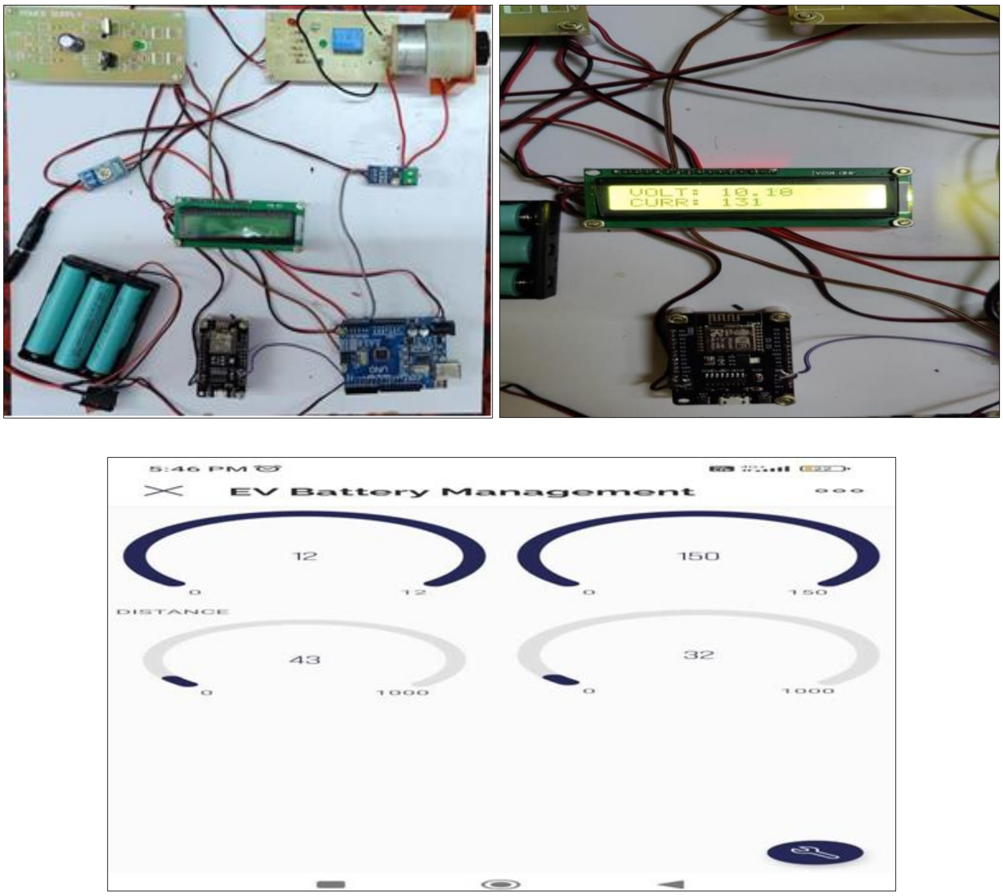

Figure 5 Hardware setup

The software setup for the sensor data and working of the controller was done in Arduino C and BLYNK IoT application displays the sensor outputsfrom the ESP Controllers the same readings is been sent to the LCD Display.

4.2. Range and Charging Time Tabulated values

The below Table 1 refers to estimate the value for various parameters such as RMSE, MAE, R(Square), SMAPE for Random Forest and Table 2 refers to estimate the value for various parameters such as RMSE, MAE, R(Square), SMAPE for SVR.

Table 1 Performance metrics of Range detection for Random Forest

RMSE	MAE	R^2	SMAPE
0.1183	0.083	0.99	0.82 %

Table 2 Performance metrics of Range detection for SVR

RMSE	MAE	R^2	SMAPE
3.14	2.88	0.9968	9.26 %

For rang detection from the results the error value of the random forest getting minimal error value as compared to SVR.The Table 3 refers estimate and compare the charging time RMSE, MAE, R(Square), SMAPE for Random Forest and Table 4 refers estimate and compare the values of RMSE, MAE, R(Square), SMAPE for SVR. RMSE lower value indicate better fit, R(Square) closer to 1 indicates better fit and closer to 0 indicates poorer fit. For SMAPE 0% indicates a perfect fit and 100% indicates the worst fit.

Table 3 Performance metrics of charging Time for Random Forest

RMSE	MAE	R^2	SMAPE
10.884	9.04	0.639	6.95 %

Table 4 Performance metrics of charging Time for SVR

RMSE	MAE	R^2	SMAPE
26.72	22.75	18.89	%

4.3. Comparison between Predictions and True values

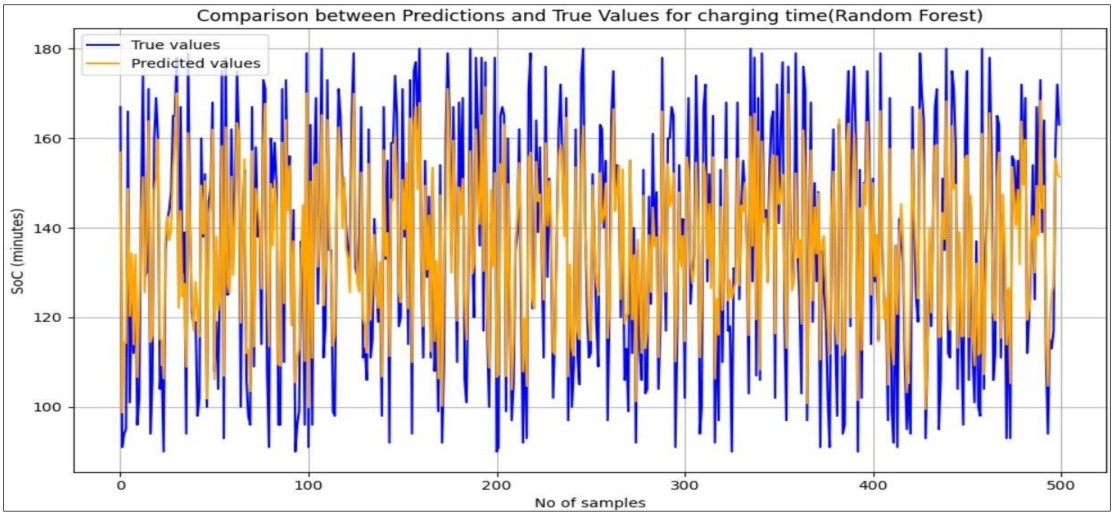

Figure 6 Comparison of Predicated and True value using line graph in Random Forest for Charging time

The above Figure 6 compares the actual value and predicted values using line graph for the Charging time application using Random Forest algorithm.

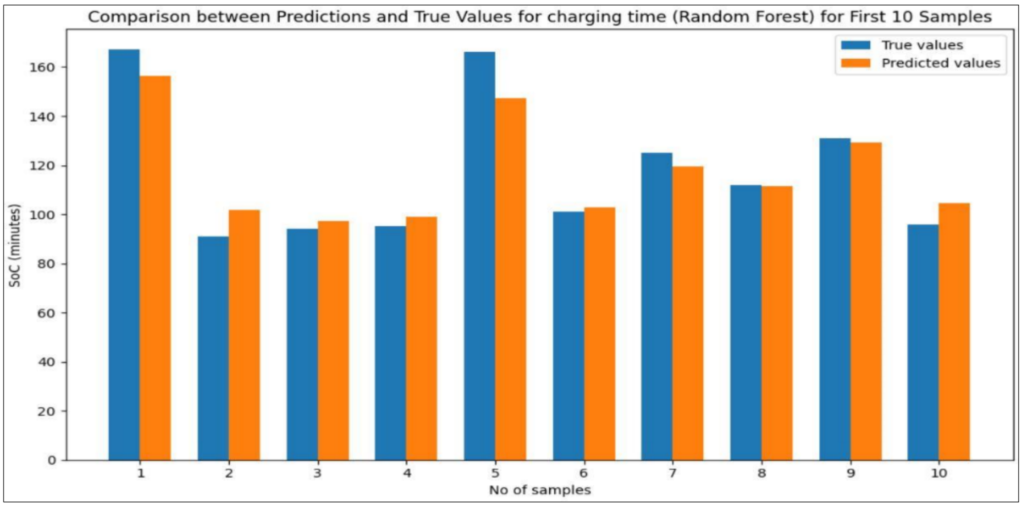

Figure 7 Comparison between predicated and true value of SoC in Random Forest

The above Figure 7 compares the actual value and predicted values of SoC for the first 10 samples for Random Forest algorithm.

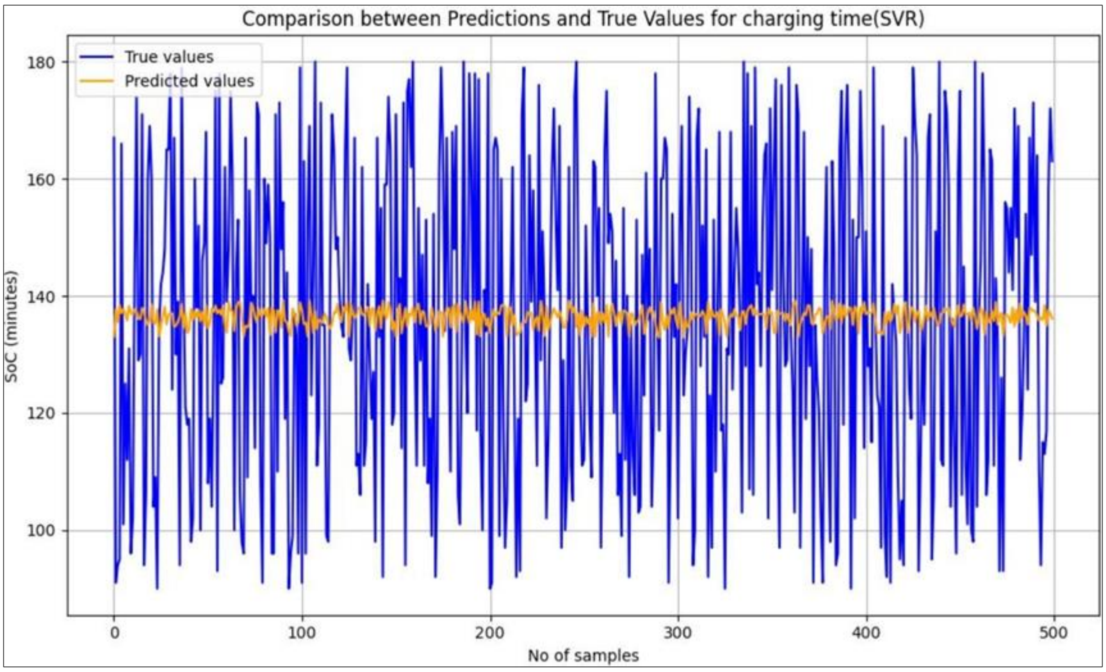

Figure 8 Comparison of Predicated and True value using line graph in SVR for Charging Time

The above Figure 8 compares the actual value and predicted values using line graph for the Charging time using SVR algorithm. Plotted SoC with respect to the number of samples and blue line indicates actual value and orange line indicates predicted value.

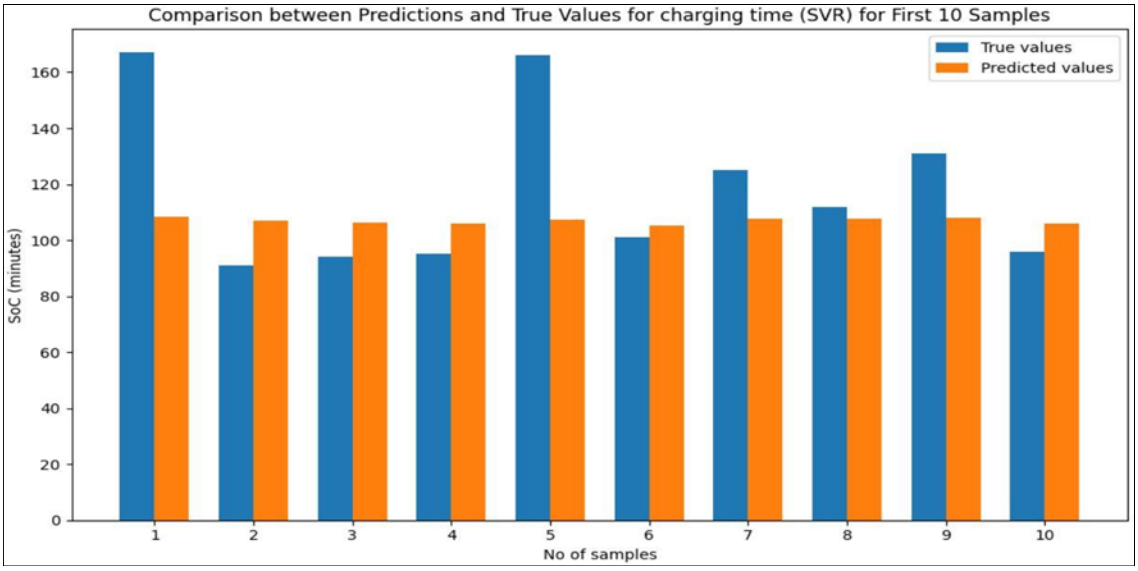

Figure 9 Comparison between predicated and true value of SoC in SVR

The above Figure 9 compares the actual value and predicted values using bar graph for the first 10 samples for the State of Charge using SVR algorithm. Blue bar indicates actual value and orange bar indicates predictedvalue.

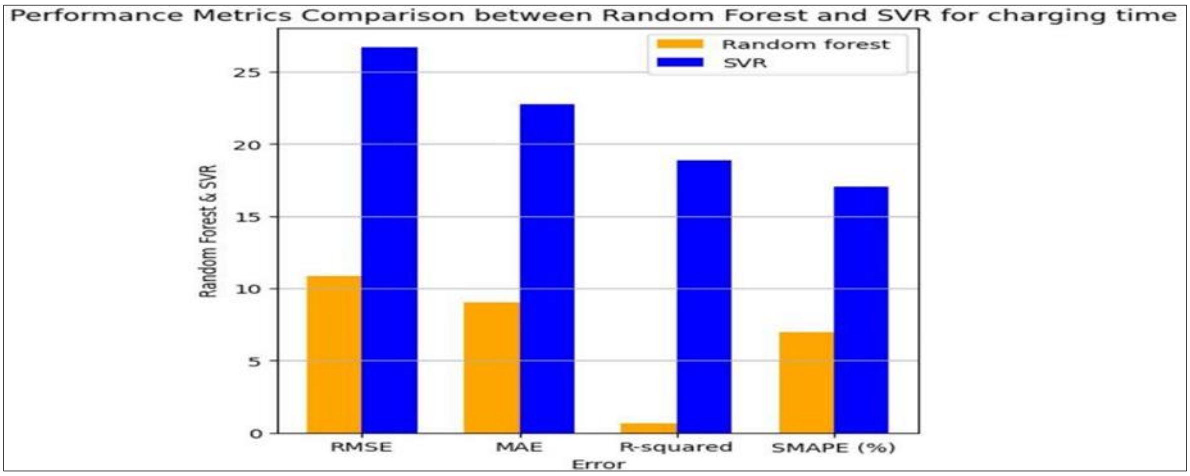

Figure 10 Comparison of Random Forest and SVR with respect to performance Metrics

The Figure 10 compares the Random Forest and SVR using bar graph with respect to performance Metrics for charging Time. In the above results it is proved that the random forest algorithm got minimal error value. The prediction and true value both are nearest to each other.

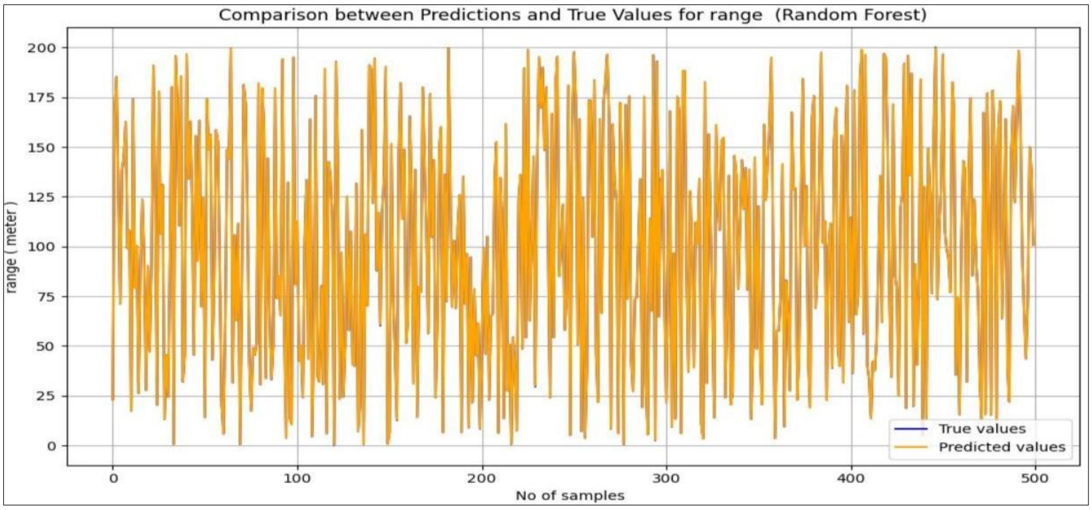

Figure 11 Comparison of Predicated and True value using line graph inRandom Forest for Range Detection

The above Figure11 compares the actual value and predicted values using line graph for the range (in meters) application using Random Forest algorithm.

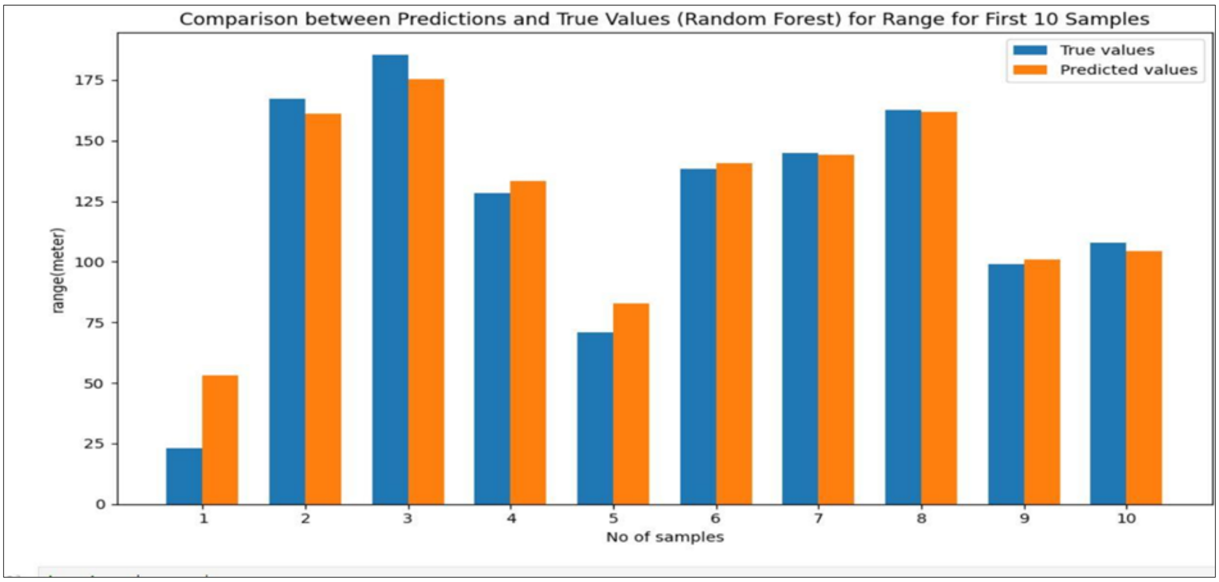

Figure 12 Comparison between predicated and true value of range detection in Random Forest

The above Figure 12 compares the actual value and predicted values usingbar graph for the range detection (in meters) application using Random Forest.

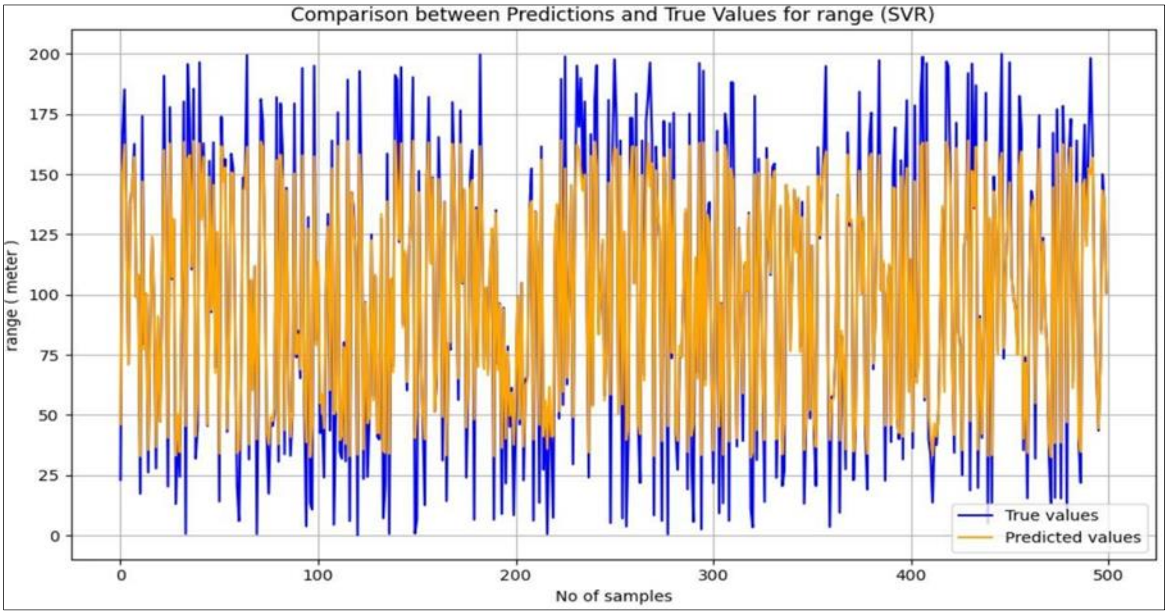

Figure 13 Comparison of Predicated and True value using line graph in SVR for Range Detection

The above Figure 13 compares the actual value and predicted values using line graph for the range (in meters) using SVR algorithm. Blue line indicates actual value and orange line indicates predicted value.

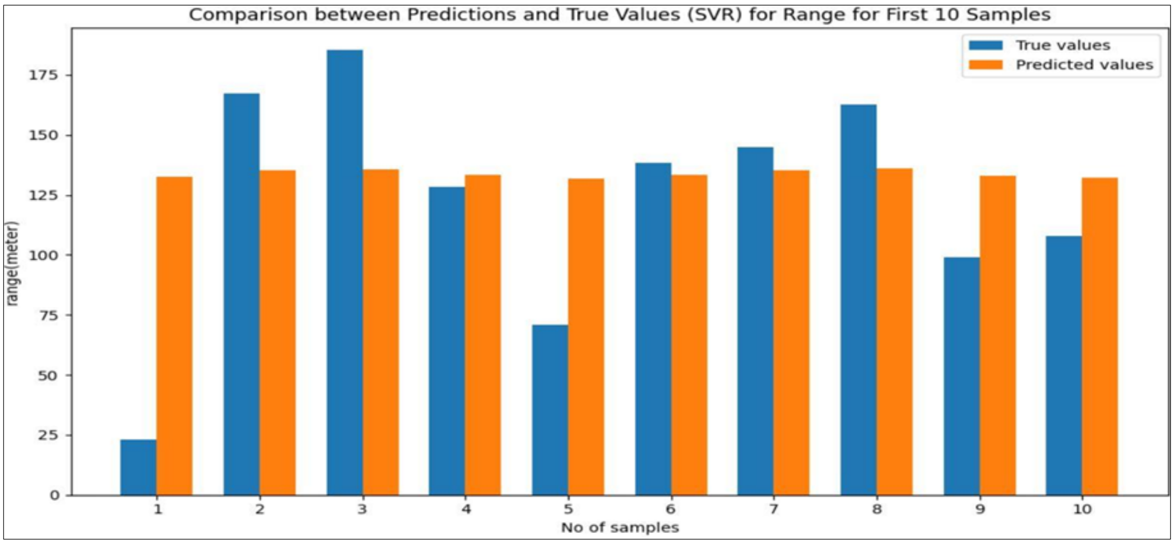

Figure 14 Comparison between predicated and true value of range detection in SVR

The above Figure 14 compares the actual value and predicted values using bar graph for the range detection (in meters) using SVR algorithm.

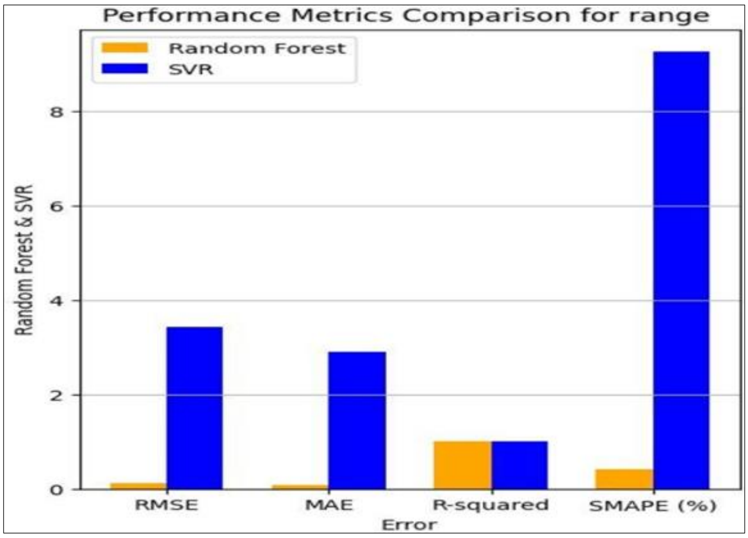

Figure 15 Performance metrics comparison of Random Forest and SVR

The Figure 15 compares the Random Forest and SVR of Performance metrics using bar graph for the with respect to error. Blue bar indicates Random Forest and orange bar indicates SVR.From the results, it is proved that,the Random Forest algorithm having minimal error value so the predicated and true value both are equal in Random Forest.

5. Conclusion

The utilization of Machine learning algorithms, specifically Support VectorRegression (SVR) and Random Forest, for the prediction of Electrical Vehicle (EV) charging time and range detection holds significant promise in enhancing the efficiency and usability of EVs. Through rigorous analysis and comparison, it was found that both SVR and Random Forest models offer valuable insights into charging time prediction and range detection. However, Random Forest demonstrated superior performance compared to SVR in terms of accuracy and reliability.

The successful implementation of Machine Learning algorithms in predicting EVcharging time and range detection signifies a significant step forward in addressing key challenges faced by EV users, such as range anxiety and inefficient charging practices. By providing accurate and reliable predictions, these algorithms empower EV drivers to make informed decisions and effectively plan their journeys, ultimately contributing to the widespread adoption and acceptance of Electric Vehicles.

Compliance with ethical standards

Acknowledgement

All authors have read and agreed to publish the manuscript.

Disclosure of conflict of interest

The authors declare no conflict of interest.

References

[1] Shahriar S. Al-Ali A. R. Ahmed Osman. Nijim M. Prediction of EV Charging Behavior Using Machine Learning. IEEE Access.2022. 9:1576–1590.

[2] Mazhar T. Asif R.N. Malik . Nadeem M.A. Haq I. Iqbal M. Kamran M. Ashraf S.Electric Vehicle Charging System in the Smart Grid Using Different Machine Learning Methods. Sustainability.2023.15: 401-426 .

[3] Huzaifa Rauf. Muhammad khalid. Naveed Arshad. Machine learning in State of health and remaining useful life estimation Theoretical and technological development in battery degradation modelling. IEEE Access. 2022.156: 541-560.

[4] Jun bi Wang. Yongxing wang. Shuai Sun. Wei Guan. Predicting Charging Time of Battery Electric Vehicles Based on Regression and Time-Series Methods. IEEE International Smart Cities Conference (ISC2).2021.19 :893-915.

[5] Khalid, Sundararajan. Acharya . Sarwat A. I. Prediction of Li- Ion Battery State of Charge Using Multilayer Perceptron and Long Short- Term Memory Models. IEEE Transportation Electrification Conference and Expo (ITEC).2022. 15:1108-11033.

[6] Junzhe Shi. Min Tian. Sangwoo Han. Tung-Yan Wu. Yifan Tang. Electric fleet charging management considering battery degradation and nonlinear charging profile. IEEE Access 2020,87, :567-589.

[7] NaitMalek Y. Najib M. Bakhouya M. Essaaidi M.On the Use of Machine Learning for State-of- Charge Forecasting in Electric Vehicle.IEEE International Smart Cities Conference (ISC2), 2019.16:408-428.

[8] Williard. Chen C . Pecht M.State of charge estimation for Li-ion batteries using neural network modeling and unscented Kalman filter- based error cancellation. International Journal of Electrical Power & Energy Systems.2019.62:783–791.

[9] Shimamurs, Watanabe. Ohsawa. Horie H.Research and Development Work on Lithium-ion Batteries for Environmental Vehicles. World Electric Vehicle Journal, 2019.1(1):251–257.

[10] Koji Sato and Yuki Sato.Battery Voltage Prediction Technology Using Machine Learning Model with High Extrapolation Accuracy. Journal of Energy Storage, 2023.30: 401-426.