Article

Not peer-reviewed version

An Analytical Approach for Accurate SoC Estimation by Comparing Different Machine Learning Algorithms for EVs

Mangesh Maurya * , Shashank Gawade , Neha Zope , Manisha Shitole , Prajwal Thorat

Posted Date: 23 August 2024

doi: 10.20944/preprints202408.1725.v1

Keywords: State of Charge (SOC); Machine Learning Algorithm; Lithium-ion (Li-ion) Battery; Accurate Soc Estimation

Article

An Analytical Approach for Accurate SoC Estimation by Comparing Different Machine Learning Algorithms for EVs

Mangesh Maurya [1,*], Shashank Gawade [1], Neha Zope [1], Manisha Shitole [2] and Prajwal Thorat [1]

[1] Department of Mechanical Engineering, School of Engineering and Sciences, MIT-ADT University, Pune, India
[2] Department of Computing Engineering, School of Engineering and Sciences, MIT-ADT University, Pune, India
* Correspondence: mangeshmaurya28@gmail.com

Abstract: The SOC (State of Charge) in Lithium-ion battery packs refers to determining a battery's current energy level or capacity at a given time. SOC is important for many battery-operated systems since it shows how much energy is available. In the earlier stage SOC is calculated by different techniques such as Coulomb counting, open-circuit voltage, Kalman filters, Data-Driven Approaches, and Real-Times Updates so these all methods are used to calculate the SOC Estimation in battery packs. In recent years, these old methods have had some errors and accuracy malfunctions in SOC estimation in battery packs. The SOC estimation errors occur on Two-Wheeler, Three-Wheeler, and four-wheeler of Electric Vehicles (EV). To Overcome this problem in current years the AI&ML Algorithms Methods are applied everywhere related to Accuracy, Solutions & All. To increase the accuracy of SOC Estimations we are going to use Machine Learning Algorithms to improve the SOC in Machine Learning there are various methods such as supervised, unsupervised, and reinforcement. Of these three methods, one of the best methods is to select the Algorithms is apply to the Battery packs to get accurate SOC Estimations. In Machine Learning, we select the Supervised Learning Algorithms. Currently, the project will focus on the use of a Supervised Machine Learning Algorithm which includes an Extra Trees Regressor, Bagging Regressor, K-Neighbors Regressor, Decision Tree Regressor, and Random Forest Regressor for SOC Estimation. In this, we will develop algorithms applied to Lithium-ion Battery Packs to accurately the SOC estimates. In conclusion, the algorithms we will develop by using Machine Learning are compared with the algorithms and analytical Models.

Keywords: state of charge (SOC); machine learning algorithm; lithium-ion (Li-ion) battery; accurate soc estimation

I. Introduction

Electric vehicles (EVs) are a revolutionary and rapidly evolving class of vehicles that are changing the landscape of transportation by providing a more environmentally friendly and sustainable substitute for conventional internal combustion engine (ICE) automobiles. EVs are powered by electricity stored in on-board batteries, which drive electric motors to propel the vehicle. This innovation addresses many of the environmental, economic, and energy challenges associated with conventional gasoline or diesel-powered vehicles [1]. In Electric Vehicles the State of Charge (SOC) Estimations determine the battery's current energy level, which is a vital component of electric vehicles (EVs). SOC is a crucial variable for controlling and enhancing the efficiency, longevity, and range of EV batteries. Drivers and system controllers may make wise choices about charging, driving style, and energy usage with the help of accurate SOC estimations. SOC represents the remainder energy ability of a proportion of the EV battery's overall capacity. For example, a SOC of 50%

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 August 2024 doi:10.20944/preprints202408.1725.v1

indicates that the battery is halfway discharged, while a SOC of 80% suggests that the battery still has 80% of its energy available. The amount of energy stored in a battery at a specific time is predicted using SOC estimate methods using a variety of methodologies, frequently involving mathematical models, measurements of battery voltage and current, temperature adjustment, and occasionally machine learning algorithms. To avoid over-discharging (which might harm the battery) and underutilization (which might leave wasted energy capacity), accurate SOC calculation is crucial [2]. Battery Voltage When there is no current flowing (open-circuit conditions), battery voltage is frequently utilized as a substitute for SOC calculation. It's important to characterize the complete link between SOC and voltage. Electric current entering or leaving the battery has an impact on its state of charge (SOC). Although it needs careful calibration and error handling, accurate measurement and integration of the current over time (coulomb counting) can yield an estimate of (SOC). The temperature of the Battery's performance and behavior is greatly influenced by temperature. To account for differences in capacity and efficiency, temperature compensation is required. Methods for estimating SOC should take into account the battery's temperature and how it may affect voltage and capacity. Initial SOC Knowledge for estimating algorithms, knowing the initial SOC (for example, at the beginning of a trip) is a good place to start. Errors in the estimation can be introduced by using incorrect initial SOC values [3]. Malfunctions or errors in State-of-Charge (SOC) estimations affect Voltage Drift when the voltage-SOC relationship fluctuates over time owing to factors including temperature, battery aging, and cycling, which is one typical source of the mistake. The SOC estimation method may produce unreliable estimates if it only considers the pack.

II. Traditional Method for SOC estimations:

State of Charge (SOC) in an electric vehicle refers to the measurement of the current energy level or capacity of the vehicle's battery pack, typically expressed as a percentage. It represents the remaining available battery energy about its overall capacity. SOC is a crucial parameter in electric vehicles as it provides information about the battery's charging status and helps determine the range or distance the vehicle can travel before requiring a recharge. A higher SOC indicates a more fully charged battery, while a lower SOC indicates that the battery has discharged and requires recharging. Various methods can be used to estimate SOC, including open circuit voltage (OCV) measurement, coulomb counting, and model-based approach

1) Coulomb Counting: coulomb counting is a widely employed method for SOC estimation, based on combining the battery's continuous inflow and outflow of current. The basic equation for Coulomb counting is as follows: SOC(t) = SOC(t-1) + (∫I(t) dt) (Q nominal) In this equation, SOC(t) represents the SOC at a time 't', SOC(t-1) is the SOC before the previous time step, I(t) am current at a time 't', ∫I(t) dt symbolizes the current's time integral., and Q -nominal is the nominal capacity of the battery

$$SOC(t) = SOC(to) + \frac{\int_{to}^{to+\tau} I_{bat}.d\tau}{Q_{rated}} \times 100\%$$

SOC: State Of Charge
I_{bat}: battery current value
Q_{rated}: rated capacity (nominal capacity)

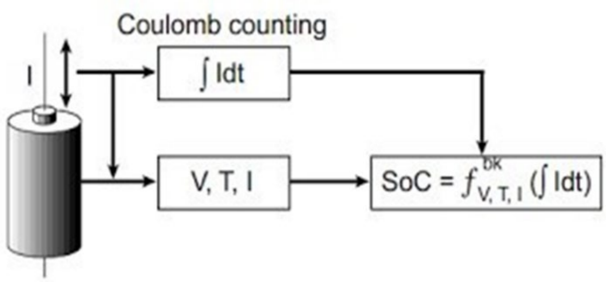

Figure 1. Coulomb counting Method [4].

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 August 2024 doi:10.20944/preprints202408.1725.v1

2) Open-Circuit Voltage (OCV) Method: The Open-Circuit Voltage method estimates SOC by measuring the battery's voltage when it is at rest and relating it to SOC using a per-determined voltage-SOC relationship The Equation for SOC Estimation using the OCV method is typically represented as: SOC = f(OCV)

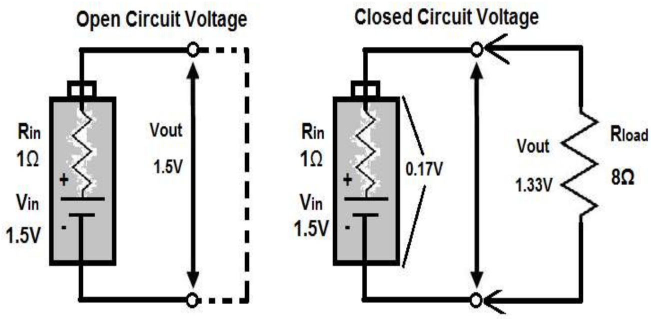

Figure 2. Open-circuits Methods [4].

3) Model-Based Approaches: A Study of Machine Learning Algorithms in State of Charge (SOC) Estimation MIT School of Engineering & Science, M-Tech Mechanical (Electric Vehicles) 4 Model-based approaches utilize mathematical models that represent the battery's behavior and characteristics to estimate SOC.

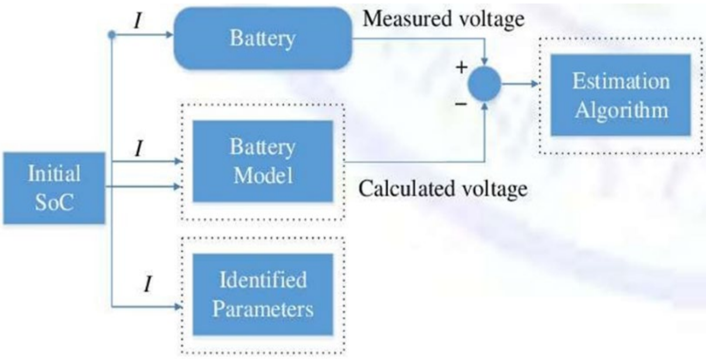

Figure 3. Model-Based Methods [4].

4) Extended Kalman filter (EKF): In indirect approaches, data from other estimated quantities is used to evaluate the SoC. Measurements of the open circuit voltage (VOCV) can be used to calculate SoC [4]. The VOCV of lead-acid and lithium-ion batteries is directly related to their SoC, and this relationship is typically assessed by experimentation.

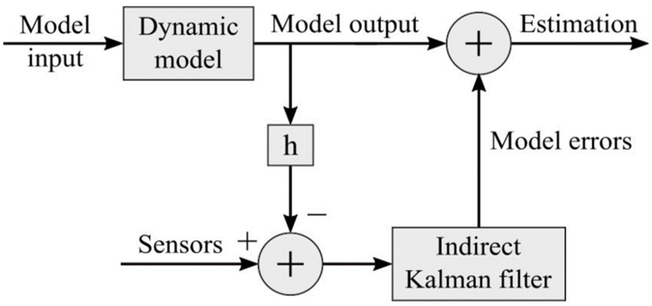

Figure 4. EKF Methods [5].

III. SOC Estimations Errors Manipulated by Machine Learning Methods:

Creating models and techniques that let computers see patterns and make judgments without explicit programming is the goal of the artificial intelligence (AI) discipline of machine learning. Here is a broad overview of machine learning Supervised Learning A labeled dataset, consisting of paired input data and desired output, is used to train the algorithm in supervised learning. Learning a mapping function that can precisely forecast the result for a novel, unforeseen inputs is the aim. Unsupervised Learning involves working with unlabeled data [5]. The algorithm explores the inherent structure in the data to find patterns or relationships without explicit guidance on the output. Semi-Supervised Learning: This type of learning combines supervised and unsupervised learning. The algorithm is trained on a dataset containing labeled and unlabelled data, making it possible for it to use unlabeled data for unsupervised learning and labeled data for supervised learning. Reinforcement Learning This technique entails educating an agent to make choices in a given setting [6,7]. Based on its behavior, the agent receives feedback in the form of rewards or penalties, and the objective is to eventually learn a policy that maximizes the cumulative reward. Algorithm Machine learning algorithms are the mathematical models or sets of rules that enable computers to learn patterns and make predictions. Examples include decision trees, support vector machines, and neural networks. Model: A model is the learned representation of patterns in the data. It can be a mathematical equation, decision tree, neural network, or any other structure that captures the relationships in the training data. The characteristics or attributes of the input data that the algorithm utilizes to generate predictions are known as features. Feature engineering involves selecting, transforming, or creating features to improve the model's performance. Training data is the dataset used to train the machine learning model. It consists of input-output pairs for supervised learning or just input data for unsupervised learning. The loss function measures how well the model's predictions match the actual outcomes. During training, the algorithm adjusts its parameters to minimize the loss, improving the model's accuracy Hyperparameters are configuration settings external to the model that are set before training. Examples include learning rates, regularization parameters, and the architecture of a neural network [7,8]. The earlier Method for calculating SOC Estimations by Traditional Methods Such as an open circuit voltage, Columb counting method, Model-based approaches, and Indirect Kalman filter. While in this method there are errors to find in SOC estimations such as errors in Model Accuracy, Parameter Accuracy, Operating Conditions, Degradation Effects, and Cyclic Variations SOC is an important factor for many battery-operated systems since it shows how much energy is available for usage Soc Estimation is required in a battery it may manipulate the driver's real-time conditions to avoid this problem Soc estimation uses machine learning algorithms to improve Accuracy and estimate it. In Machine Learning Algorithms we select Supervised Learning in that there are regression and Classification we compared five algorithms Decision Tree, ExtraTreesRegressor, Bagging Regressor, Neighbour's Regressor, and Random Forest Regressor. These five algorithms compare and select the best algorithms and implement on to improve the soc estimation and accuracy [9].

IV. Analytical Approach

The methodology in this current project work is divided into two parts, first generating data using an analytical approach and later validating it with an experimental approach. The analytical method starts to generate data from the Test Rig experiment setups such as current, temperature, voltage, and Soc Precisely calculating a battery's state of charge (SOC) is essential in many situations, particularly when dealing with electric cars (EVs) and other battery-operated devices [10,11]. While non-invasive estimation via direct measurement is impractical, machine learning (ML) provides strong tools. BTMS has become a more economical and sustainable solution [12].

1) Gathering of Data: Collect information from a range of sensors and measurements, including temperature, voltage, current, and perhaps other environmental elements that have an impact on battery operation [13].

2) Engineering Features: Extrapolate pertinent characteristics that reflect the status of charge from the gathered data. Voltage characteristics, charge and discharge rates, temperature

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 August 2024 doi:10.20944/preprints202408.1725.v1

dependencies, and other factors pertinent to the particular battery chemistry and application are examples of features [14,15]

3) Selecting a Machine Learning Model: Select the right machine learning methods to estimate SoC. Typical options consist of Regression algorithms including support vector, polynomial, and linear regression [16].

4) Training Models: Utilizing the information gathered, train the chosen machine learning models. Make use of methods like cross-validation to assess and enhance model performance [17,18].

5) SoC Estimation: Use the machine learning models that have been trained to estimate the condition of charge in real time. As new data becomes available, update and enhance the models frequently to increase accuracy and make them more flexible to changing operational conditions [19,20].

A. Datasets

The above data sets Load, Time, Speed, and Crr to test and create data sets for Soc Outputs.

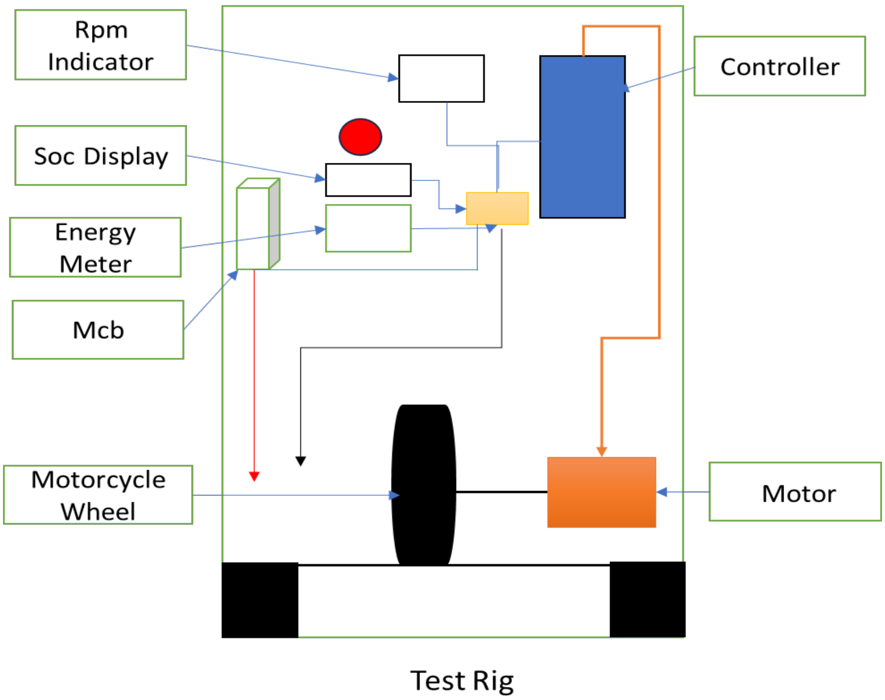

Figure 5. Experiments Setups.

The Test Rig System is used to generate the Data of Different loads, speeds, Time, and Table.1, The test Rig tests the motor efficiency, VI characteristics, heat, and motor Performance. In this setup, we record and generate data such as current, temperature, and voltage. Show in Table.2

Table 1. Data Sets.

Load	Time	Speed	Crr
60.0	43	30	0.1
70.1	18	45	0.1
70.1	67	15	0.1
70.1	67	45	0.1
70.1	18	15	0.1
102.5	43	10	0.1
102.5	43	30	0.1

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 August 2024 doi:10.20944/preprints202408.1725.v1

Table 2. Soc Output.

Sr No	Current	Voltage	Temperature	SOC Output %
1	21.88	50.08	29	84
2	21.71	50.11	29	84
3	21.51	50.08	29	84
4	21.22	50.01	29	84
5	21.27	50.02	29	83.91
6	21.88	50.08	29	83.91

Above the Test Rig setup is designed and developed for analysis of the load and speed effects on the motor. The Rig indicates the Current, Temperature, Voltage, and Soc Display.

V. Results and Discussions

We are using Python software for analytical purposes and testing your data sets to run the models and tests and train the models using machine learning algorithms. Using Python software is compatible with users for ML coding and libraries. etc.

Table 3. Comparison of Models.

Model	Adjusted-Squared	R-Squared	RMSE	Time Taken
DecisionTreeRegressor	0.99	0.99	1.37	0.03
ExtraTreesRegressor	0.99	0.99	1.46	0.48
Bagging Regressor	0.99	0.99	1.49	0.08
KNeighborsRegressor	0.99	0.99	1.52	0.04
RandomForestRegressor	0.99	0.99	1.56	0.64

Training RMSE: 1.24
Testing RMSE: 1.37
Training Accuracy: 0.99
Testing Accuracy: 0.99
Therefore, after the combination of all the available data, we get thefollowing results....
Training Accuracy: 99%
Testing Accuracy: 99%

The above Machine Learning Code compares the five algorithms as DecisionTreeRegressor, ExtraTreesRegressor, Bagging Regressor, KNeighborsRegressor, and Random Forest Regressor

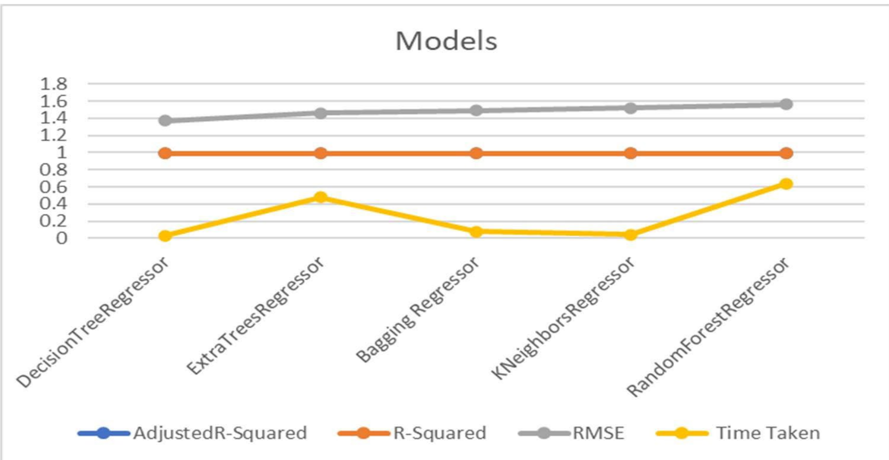

Figure 6. Represents the Graph comparing Five Models.

The comparison of five machine learning algorithms as per given data and finalization of the decision tree regressor.

1) In a regression model, the statistical expression for the portion of the variation of a dependent variable can be accounted for by one or more independent variables as R-squared (R^2) [21]. Its range is 0 to 1, where 0 indicates that there is no explanatory power and 1 indicates that the independent variable(s) fully explains the variance in the dependent variable. [22].

2Adjusted R-squared is a modified version of R-squared that takes the number of predictors in the model into consideration. (also known as adjusted R^2) [23]. It discourages overfitting by penalizing the inclusion of pointless predictors in the model. When a more informative predictor is introduced to the model, the adjusted R-squared usually increases or stays the same. However, if a less informative predictor is included, it usually drops [23,24].

3) The difference between values predicted by a model or estimator and the observed values is measured by the Root Mean Squared Error or RMSE [25]. It is the square root of the mean squared discrepancies between the actual observation and the prediction. RMSE gives a general idea of the size of the error in the model and has the same unit of measurement as the dependent variable [26].

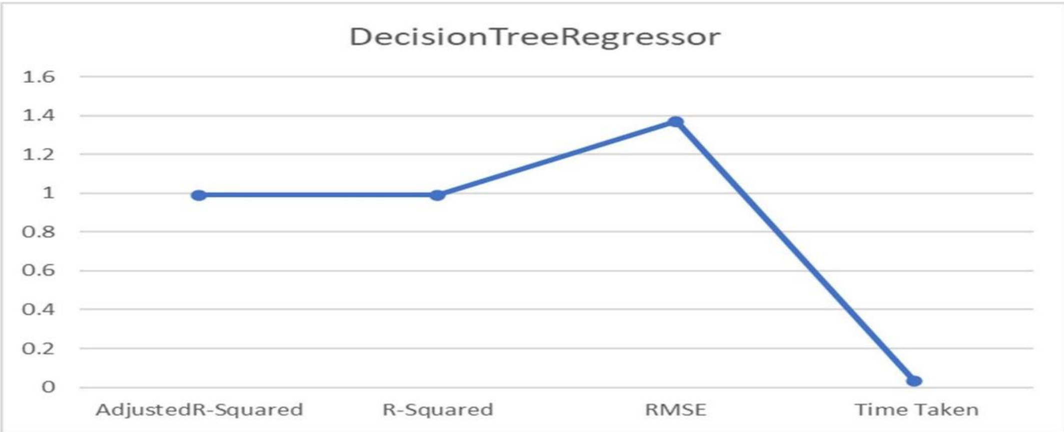

Figure 7. Decision Tree Regressors.

The Above Fig shows the suitability of your data sets and provides the best accuracy for Soc estimations.

VI. Conclusion and Future Scope

The results obtained from the Machine learning model are shown in the above figure. These results are verified against the previous literature. The study will also aid in understanding how the Improve the Soc Estimation of Battery Packs. Some key points first generate data major factors for soc data are current, voltage, and temperature. SOC is linearly proportional and voltage is in continuous form. The system will get more accurate as well as increase the battery life cycle, over and under voltage protections.

1. The best-suited ML from the analytical calculations of Datasets was found to be Decision Tree Regressors
2. The battery Soc Accuracy improves by using ML Algorithms to select a Decision Tree (0.99,1.37,0.03) an analytical approach that motivates its use in experimental analysis.

The future scope of the work also lies in validating the analytical results obtained for Big Data Sets at different Battery Loads and comparing them using the results obtained through experimental analysis.

Acknowledgment: A heartfelt thank you to the authors of the School of Engineering and Sciences at MIT-ADT University for complete guidance throughout the study.

Abbreviations

SoC-state of charge
Li-ion- lithium-ion
ETR- Extra Trees Regressor
K-NR - K-Neighbors Regressor
DTR- Decision Tree Regressor
RFR- Random Forest Regressor
CC- Coulomb Counting
EKF-Extended Kalman filter
OCV- Open Circuits Voltage
MBM-Model-Based Methods
EV-Electric Vehicles
BR-Bagging Regressor
RMSE-Root mean squared error

References

1. D. Abdulqader, A. Mohsin Abdulazeez, and D. Zeebaree, "Machine Learning Supervised Algorithms of Gene Selection: A Review," Apr. 2020.
2. M. W. Libbrecht and W. S. Noble, "Machine learning applications in genetics and genomics," Nature Reviews Genetics, vol. 16, no. 6, pp. 321–332, 2015.
3. J. Wang, P. Neskovic, and L. N. Cooper, "Training Data Selection for Support Vector Machines," in Advances in Natural Computation, vol. 3610, L. Wang, K. Chen, and Y. S. Ong, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2005, pp. 554–564.
4. D. Maulud and A. M. Abdulazeez, "A Review on Linear Regression Comprehensive in Machine Learning," Journal of Applied Science and Technology Trends, vol. 1, no. 4, pp. 140–147, 2020.
5. G. Carleo et al., "Machine learning and the physical sciences," Reviews of Modern Physics, vol. 91, no. 4, p. 045002, 2019.
6. R. Xiong, J. Cao, Q. Yu, H. He, and F. Sun, "Critical Review on the Battery State of Charge Estimation Methods for Electric Vehicles," IEEE Access, vol. 6, pp. 1832–1843, 2017, doi: 10.1109/ACCESS.2017.2780258
7. Hendrycks, Dan, Kimin Lee, and Mantas Mazeika. "Using pre-training can improve model robustness and uncertainty." In *International conference on machine learning*, pp. 2712-2721. PMLR, 2019.
8. Li, Yuan, et al. " Comparative study of the influence of open circuit voltage tests on state of charge online estimation for lithium-ion batteries. Quot; Ieee Access 8 (2020): 17535-17547.
9. Gong, Dongliang, Ying Gao, and Yalin Kou. " Parameter and State of Charge Estimation Simultaneously for Lithium-Ion Battery Based on Improved Open Circuit Voltage Estimation Method. " Energy Technology 9.9 (2021): 2100235.
10. Irshada, Mariyam, and V. Kumar. "SMOTE and ExtraTreesRegressor based random forest technique for predicting Australian rainfall." *International Journal of Information Technology* 15, no. 3 (2023): 1679-1687.
11. Mohammadi, Fazel. " Lithium-ion battery State-of-Charge estimation based on an improved Coulomb-Counting algorithm and uncertainty evaluation. Quot; Journal of Energy Storage 48 (2022): 104061.
12. Ng, Kong Soon, et al. " Enhanced coulomb counting method for estimating state-of-charge and state-of-health of lithium-ion batteries. Quot; Applied Energy 86.9 (2009): 1506-1511.
13. Wei, Zhongbao, Jiyun Zhao, Hongwen He, Guanglin Ding, Haoyong Cui, and Longcheng Liu. "Future smart battery and management: Advanced sensing from external to embedded multi-dimensional measurement." *Journal of Power Sources* 489 (2021): 229462.
14. J. H. Ahn and B. K. Lee, "High-Efficiency Adaptive-Current Charging Strategy for Electric Vehicles Considering Variation of Internal Resistance of Lithium-Ion Battery," IEEE Trans. Power Electron., vol. 34, no. 4, pp. 3041–3052, Apr. 2019.
15. Z. Xia and J. A. Abu Qahouq, "State-of-Charge Balancing of LithiumIon Batteries with State-of-Health Awareness Capability," IEEE Trans. Ind. Appl., vol. 57, no. 1, pp. 673–684, Jan. 2021.
16. Chandran, Venkatesan, Chandrashekhar K. Patil, Alagar Karthick, Dharmaraj Ganeshaperumal, Robbi Rahim, and Aritra Ghosh. "State of charge estimation of lithium-ion battery for electric vehicles using machine learning algorithms." *World Electric Vehicle Journal* 12, no. 1 (2021): 38.
17. D. N. T. How et al., "State-of-Charge Estimation of Li-Ion Battery in Electric Vehicles: A Deep Neural Network Approach," IEEE Trans. Ind. Appl., vol. 56, no. 5, pp. 5565–5574, Sep. 2020
18. Zhang, J.; Wang, Q.; Meng, F.; Shi, H.; Xi, Y. New Energy Vehicle Battery SOC Evaluation Method based on Robust Extended Kalman Filterd. J. Phys. Conf. Ser. 2022, 2196, 012037. [CrossRef]

 doi:10.20944/preprints202408.1725.v1

19. Bates, Stephen, Trevor Hastie, and Robert Tibshirani. "Cross-validation: what does it estimate and how well does it do it?" *Journal of the American Statistical Association* 119, no. 546 (2024): 1434-1445.
20. Hossain, M.; Haque, M.; Arif, M. Kalman filtering techniques for the online model parameters and state of charge estimation of the Li-ion batteries: A comparative analysis. J. Energy Storage 2022, 51, 104174. [CrossRef]
21. Murphy, Sierra. "Regression." In *Translational Orthopedics*, pp. 137-140. Academic Press, 2024.
22. Yuan, H.; Han, Y.; Zhou, Y.; Chen, Z.; Du, J.; Pei, H. State of Charge Dual Estimation of a Li-ion Battery Based on Variable Forgetting Factor Recursive Least Square and Multi-Innovation Unscented Kalman Filter Algorithm. Energies 2022, 15, 1529. [CrossRef]
23. Cui, Z.; Wang, L.; Li, Q.; Wang, K. A comprehensive review on the state of charge estimation for lithium-ion battery based on neural network. Int. J. Energy Res. 2021, 46, 5423–5440. [CrossRef]
24. Liu, X.; Dai, Y. Energy storage battery SOC estimate based on improved BP neural network. J. Phys. Conf. Ser. 2022, 2187, 012042. [CrossRef] Energies 2023, 16, 2155 16 of 16
25. Hu, C.; Cheng, F.; Ma, L.; Li, B. State of Charge Estimation for Lithium-Ion Batteries Based on TCN-LSTM Neural Networks. J. Electrochem. Soc. 2022, 169, 030544. [CrossRef]
26. Meng, J.; Luo, G.; Ricco, M.; Swierczynski, M.; Stroe, D.-I.; Teodorescu, R. Overview of Lithium-Ion Battery Modeling Methods for State-of-Charge Estimation in Electrical Vehicles. Appl. Sci. 2018, 8, 659.