Tiako P.F. (ed) *Competitive Tools, Techniques, and Methods*. Chronicle of Computing. OkIP. CAPE24#11

https://doi.org/10.55432/978-1-6692-0007-9_12

Machine Learning-Based SOC Prediction for Lithium-ion Batteries in Electric Vehicles

Abdulrahman Mahmoud Eid
Department of Computer Engineering
University of Sharjah
Sharjah, UAE
U20105546@sharjah.ac.ae

Ali Bou Nassif
Department of Computer Engineering
University of Sharjah
Sharjah, UAE
anassif@sharjah.ac.ae

Chaouki Ghenai
Department of Sustainable and Renewable Energy Engineering
College of Engineering
Sharjah, UAE
cghenai@sharjah.ac.ae

Heba Yahia Youssef
Department of Electrical Engineering
University of Sharjah
Sharjah, UAE
U18105651@sharjah.ac.ae

Latifa A. Alkhaja
Department of Computer Engineering
University of Sharjah
Sharjah, UAE
U18100910@sharjah.ac.ae

Hajar Humaid Almazrouei
Department of Computer Engineering
University of Sharjah
Sharjah, UAE
U18102546@sharjah.ac.ae

***Abstract*— Developing an effective and accurate battery management system that can predict the state of charge of electric vehicles is essential to enhance the safety and efficiency of electric vehicles. The assessment of the state of charge of the battery is important not only for determining the amount of energy that is available from the battery but also for determining how long the battery will last. This paper provides a brief understanding of how the state of charge estimation was predicted before and after the era of Machine Learning (ML). In addition, it proposes an accurate and fast state of charge estimation for lithium-ion batteries in electric vehicle applications using machine learning. The proposed model is designed to be generalizable across various inputs, applicable to both new and old batteries, and robust under different charging and discharging scenarios. Model performance and accuracy were evaluated using predefined metrics such as Root Mean Square Error and Mean Absolute Error. Among the four machine learning algorithms that achieved an approximate error of 0.75, while Multiple Linear Regression (MLR) model was chosen for its lightness and speed in training and testing. The findings of this research contribute to the advancement of Battery Management System (BMS) design and implementation for enhancing the efficiency and safety of Lithium-Ion Batteries (LIBs) in real-world driving scenarios. In the future, further research might be conducted to study the implementation of a variety of deep learning algorithms, as well as the estimation of battery health and the remaining useful life.**

Keywords*— *battery management system, electric vehicles, state of charge, machine learning, and lithium-ion batteries

I. Introduction

Lithium-Ion Batteries (LIBs) have proven to be the most efficient and popular energy storage option in Electric Vehicles (EVs). Extensive research and development have been conducted to improve their safety, dependability, and durability [1], [2]. This indicates a bright future for LIB usage in EVs and other vehicles and makes it an ideal research subject. LIBs possess distinct characteristics such as high efficiency, extended cycle life, low discharge rate, and high output voltage. Because LIBs are efficient at achieving increased performance over time, special consideration must be given to their working conditions to reduce physical damage, aging, and thermal runaways. As a result, there is a pressing need to develop a Battery Management System (BMS) capable of precisely monitoring, estimating, and managing battery SOC, among other statuses [3]. An energy management system and control are essential for any electric vehicle that uses batteries to improve efficiency on the road. This can only happen if power is continuously transferred from the energy storage system to the wheels of the car in response to actual demand [4].

BMSs are widely used in a variety of portable electrical and electronic devices; however, BMS implementation in EVs is still in its early stages due to the significant difference in cell count and power ratings between EVs and small-scale devices. The BMS continuously monitors the current, voltage, and temperature to obtain the battery parameters (e.g., operational voltage, current, power consumption, State of Charge (SOC), aging, internal impedance, ambient temperature, etc.) to keep the battery within its safe limits during charging and discharging [5]. Thus, one of the primary considerations is safeguarding the batteries and EVs from thermal runaways and explosions. Accurately predicting the battery's remaining range is also crucial, particularly since long-distance travel can result in discharges of up to 80% or higher [6]. As a result, proper battery protection is critical during charging and discharging. To maximize the battery's efficiency and safety, a highly effective BMS capable of predicting SOC and other critical battery functions is essential.

Recently, there has been a surge in academic and research interest in developing SOC estimation models using machine learning (ML) techniques [7]. The primary objective of this study is to develop a highly efficient SOC estimation model that can be integrated along with the BMS to enhance EV performance. To achieve this goal, a comprehensive analysis of the Panasonic18650PF Li-ion battery dataset [8] was conducted to evaluate the proposed SOC estimation model's effectiveness. The efficacy of four distinct ML models was analyzed, and their performance was benchmarked against each other. The results demonstrate that proper data preparation, even with simpler ML algorithms, can outperform complex models in prediction accuracy, highlighting its potential to significantly advance the field of EV BMS.

Presented at the 2024 OkIP International Conference on Advances in Power and Energy (CAPE) in Oklahoma City, OK, USA, and Online, on October 3, 2024

A. Literature Review

Researchers in [9] introduced SOC Estimation for LIBs using Deep Forward Networks (DFNN) with the flexibility of choosing between two or four hidden layers. They evaluated their model's performance on the Panasonic 18650PF dataset. Their proposed models were only evaluated by Mean Absolute Error (MAE). The proposed two hidden layers model demonstrated an MAE of 1.60%, while the one with four hidden layers achieved 2.0%. In [10], researchers also proposed SOC Estimation for LIBs using DFNN but with two hidden layers only. Panasonic 18650PF dataset has also been utilized to evaluate the model's performance. Their proposed model was evaluated using MAE only and achieved 1.1%. Academics in [11], proposed SOC Estimation for LIB using DNN with five hidden layers. They evaluated the model's performance on the earlier-mentioned dataset. Their model was assessed by two metrics and demonstrated an MAE of 0.24% and a Mean Square Error (MSE) of 0.11%. The scholars in [7], proposed four distinct models to estimate SOC for Li-ion batteries using Multi-Layer Regression (MLR), Multi-Layer Perceptron (MLP), Support Vector Regression (SVR), and Random Forest (RF). Their proposed models were also tested on the Panasonic 18650PF dataset. Their suggested models showed corresponding MAEs for MLR, MLP, SVR, and RF of 3.16%, 3.18%, 3.10%, and 0.82%. Furthermore, these models' Root Mean Square Error (RMSE) values were 4.32%, 4.21%, 4.30%, and 1.44%, respectively.

Following on the previously mentioned works done to build an efficient SOC estimation model, and evaluated using the Panasonic 18650PF, it can be seen that they suffer from the following:

1) Utilizing Deep Learning (DL) models inefficiently where complex models were applied in an insufficient way which results in a medium to low prediction model performance, or a time-consuming model.

2) Extra hidden layers in DL models do not implicitly indicate a better model performance. On the other hand, it might only increase complexity and time consumption.

3) Employing classical ML algorithms without proper data analysis and preparation. This results in a poor performance model.

The contribution of this paper is as follows:

1) Provide a brief understanding of how the SOC estimation was predicted before and after the era of ML.

2) Build and propose four SOC estimation models employing MLR, SVR, RF, and MLP, and verifying them using the Panasonic 18650PF dataset.

3) Evaluate the performance of the proposed models considering the several prediction errors, learning, and prediction speed.

The remaining part of this paper is organized as follows: Section 2 provides a brief understanding of how SOC estimation was predicted before and after the era of ML. Section 3 illustrates the proposed methodology implemented to build an efficient SOC estimation model. Results and discussion will be presented in Section 4. Lastly, Section 5 concludes with the conclusion and final remarks.

II. State-of-Charge Definition and Estimation before and after the era of ML

SOC is the proportion of a battery's remaining charge capacity in relation to its maximum capacity when fully charged [12]. The amount of capacity left in the battery, which is equivalent to the fuel gauge in ordinary vehicles, is a crucial factor in the driving experience as represented in Eq.1.

$$SOC = 1 - \frac{\int i\, dt}{C_n} \tag{1}$$

Where 'I' is the current either charging or discharging. 'dt' is the infinitesimal time interval. 'C_n' is the nominal capacity of the battery.

The C_n steadily declines over time due to variations in external load and the battery's internal chemical reactions, resulting in non-stationary and non-linear battery degradation characteristics. Furthermore, significant SOC errors may occur because of terminal reading build-up, necessitating periodic value recalibration [13].

A battery's SOC reflects its usable capacity, and it is one of the most important states to monitor in a battery to optimize performance and extend battery life. Accurate SOC estimation can help drivers make informed decisions about their vehicle's charging schedule. It also supports the BMS to prevent overcharging and over-discharging, both of which can be dangerous [14].

A. SOC Estimation Methods

Many researchers have been interested in SOC estimation, and numerous alternative approaches have been proposed [15]. Method categorization is a difficult process because most approaches involve the integration of two or more approaches as well as the use of various heuristic or deterministic mathematical tools [12]. However, before the era of ML, SOC estimation methods were classified into two main ways: conventional, and model-based methods. After the evolution of ML and it is revolutionary effectiveness on real-world applications [16], data-driven methods came into the picture to support building ML models.

1) Conventional Estimation Methods: Coulomb-counting, Open-circuit Voltage, and Impedance Spectroscopy are examples of conventional estimation methods. The theory behind these models is based on experimental testing to determine the electric components (impedances, resistances, capacitors, etc.) that make up the presumed equivalent circuit models that represent the battery [17].

2) Model-based Estimation Methods: Model-based SOC estimation strategies, also known as white-box models, are created with a knowledge of the underlying processes in mind. The traditional method is the model-based method, which can solve a wide range of problems, particularly in the engineering field. This approach typically requires the practitioner to have an extensive understanding of the system or the process to build robust rules that accurately mimic the system's behavior. Because of that, model-based SOC estimation techniques may be extremely powerful and precise[18].

3) Data-driven Estimation Methods: In contrast, with the evolution of ML, a recent methodology based on data made it possible to verify and witness the effectiveness of ML in the field of BMS. This method becomes more popular with the availability of robust computers and vast volumes of data. Data-driven techniques, also known as black-box models, are built on real-world data with little or no understanding of the underlying processes. Because the data-driven technique is heavily reliant on analyzing process data, practitioners are not required to have a comprehensive, domain-specific knowledge of the underlying procedure. This method can be used to build a SOC estimation model with little prior knowledge of the battery's internal characteristics and chemical interactions. However, because data-driven techniques rely heavily on data, the accuracy and usefulness of the model are heavily influenced by data quality. Unbalanced data, for example, could cause a model's decision-making process to be biased [10].

III. Research Methodology

Fig. 2 illustrates the applied measures to build the most effective SOC estimation model. The methodology begins with dataset analysis and ends with evaluation metrics. This section details the steps taken to develop the SOC estimation system, starting with a description of the selected dataset, followed by the procedures used for preparing and processing the dataset to make it ready for the chosen ML algorithms. Finally, the section concludes with an overview of the ML models and evaluation criteria used to assess the proposed model.

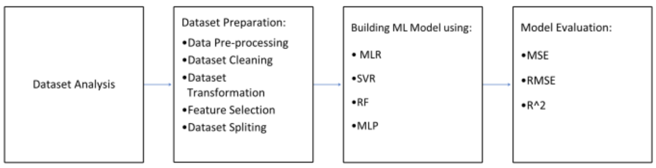

Fig. 2. The workflow of the SOC estimation model utilizing ML.

A. Dataset Description and Analysis

The Panasonic18650PF Li-ion battery dataset used in this study was generated by researchers in [8]. This dataset is very useful for evaluating and testing the effectiveness of any ML model, as it contains a variety of attributes that can be used as input features for the proposed model based on the battery test. Voltage, current, battery temperature, power in Wh, and capacity in Ah are all included in the dataset. The drive cycles in this dataset include the US06, HWFET, UDDS, and LA92, as well as mixed cycles and charging and discharging tests. It has six temperatures for each type: -10°C, -20°C, 0°C, 10°C, 25°C, and 40°C, for a total of 24 driving cycles, six mixed cycles, and six charge tests. This dataset was chosen for its reliability and prior use in several studies, as cited in the literature. This facilitates evaluating the model's effectiveness and making valid comparisons. Before starting the preparation and processing, the feature distribution was analyzed. Fig. *3*. illustrates the histogram of feature distributions where ‘Voltage’ and ‘WhAccu’ have many lower values, while ‘Temperature’ has multiple peaks, indicating diverse operating conditions. On the other hand, ‘Current’ and ‘Capacity’ sample distributions are close to normal, with some outliers visible.

1) Dataset Preparation: Data preparation is crucial for building any effective ML model. It involves several steps:

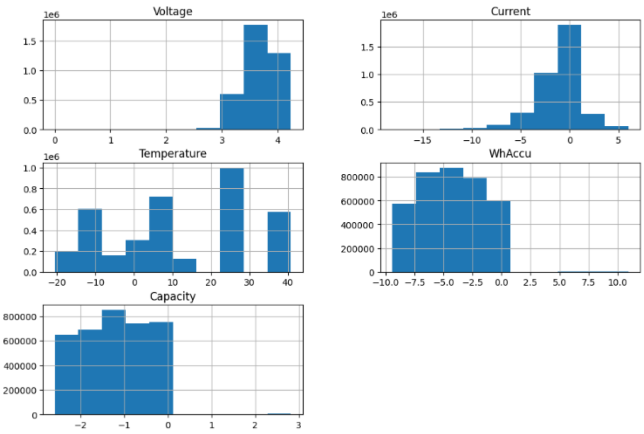

Fig 1. Panasonic LG Li-ion battery dataset feature distributions.

data preprocessing, data cleaning, data transformation, feature selection, and data splitting. This section will go through the steps taken to build the final model [19].

2) Data Preprocessing: Measurement noise in collected data is inevitable when using sensors, making data preprocessing a critical first step in ML modeling. To reduce the needed time to extract useful information from each data file. An automated code has been created to extract the useful features from dataset sets such as voltage, current, temperature, power, and finally capacity, and then reformat everything and add the header in a CSV file.

3) Data Cleaning: Data cleaning is vital for removing outliers, corrupted data, or missing points. It began by cleaning the chosen dataset, removing duplicates, and missing points, and then removing outliers using the IQR test. Finally, outliers were checked using the box plot as shown in Fig. 3, which shows the effectiveness of the cleaning steps taken.

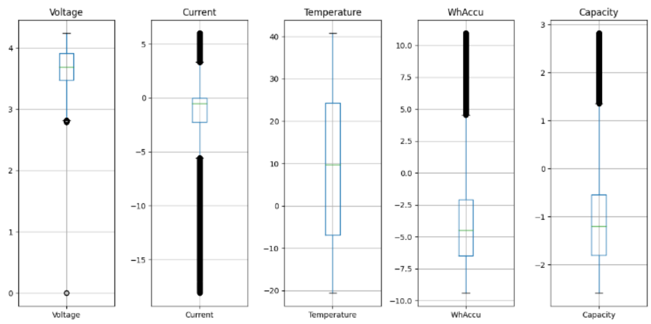

Fig. 3. Features box plot after cleaning the outlier samples.

4) Data transformation: Data transformation is typically necessary when the dataset contains data in various formats or when multiple datasets are combined. In this scenario, a MinMax scalar was implemented, which resulted in rescaling the data to a range between 0 and 1. This assisted in reducing the model's sensitivity to the feature scale [20].

5) Feature Selection: Feature selection is critical in the model-building process. The presence of correlated and insignificant characteristics reduces the model's performance. The process began by removing the time stamp and power properties. The power was removed because of its direct relationship to both current and voltage, as the correlation between the inputs could disrupt model performance. ‘Voltage’, ‘Current’, ‘Temperature’, ‘Power in Wh’, and

'State of Charge' in Ah are all possible characteristics. The correlation between features was then analyzed as Fig. 4, which provided the significance of all attributes, and all features as input were significantly correlated with the SOC. As a result, all the remaining parameters were used as inputs to the model ('Voltage,' 'Current,' and 'Temperature'), with the SOC as the output.

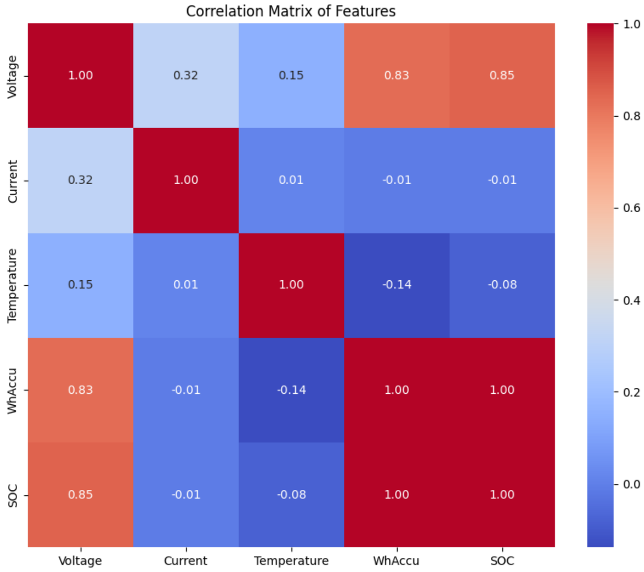

Fig. 4.Features box plot after cleaning the outlier samples.

6) Data Splitting: The dataset was then divided into X and Y columns, where X represents the independent variable, model inputs, and Y represents the dependent variable, model output. The independent and dependent variables were then divided into training and testing sets. The training set was used to train the model, while the test set was used to evaluate the model's performance on new data. Due to the large size of the chosen datasets, the dataset was divided into 80% training and 20% testing for both X and Y using the hold-out method. This strategy is appropriate if the goal is to compare models on the test dataset based on model accuracy and select the best model.

B. Machine Learning Training

Four different ML algorithms were utilized, and their performances were tested and compared. The choice of algorithms was based on their actual performance post-implementation. Supervised learning algorithms were employed because the available data is already labeled and does not require additional labeling efforts.

In addition, the model's output is the SOC, which is a continuous value. Thus, using regression algorithms is appropriate in this case. After a review of the literature, the following four algorithms were employed: Multilayer Perceptron (MLP), Support Vector Machine (SVR), Linear Regression (LR), and Random Forest (RF).

1) Multiple Linear Regression (MLR): MLR is a simple linear regression extension in which the algorithm seeks the best straight-line fit between the multiple inputs and the output. The following relation maps the inputs Xn to the predicted value Ŷ as shown in Eq. 2:

$$\hat{Y} = b0 + b1X1 + \cdots + bnXn \quad (2)$$

The coefficients bn are determined during the model's training to minimize prediction errors relative to the actual output. The default MLR was utilized for the proposed model.

2) Support Vector Regression (SVR): SVR is conceptually similar to Support Vector Machine in that it establishes a decision boundary between different classes. Instead of classifying the observations, it finds a linear function for regression. It is defined as an optimization problem that begins with the creation of a convex-insensitive loss function to be minimized and ends with determining the flattest tube containing most of the training cases. The default MLR was utilized for the proposed model.

3) Random Forest (RF): The RF algorithm is a supervised ML algorithm that can be used for classification as well as regression. It makes use of ensemble learning, a technique that combines multiple classifiers to provide accurate predictions in complex situations. The prediction is made by RF algorithms using bagging or bootstrap aggregation on the results of multiple decision trees. The RF used consists of 100 trees to make predictions.

4) Multilayer Perceptron (MLP): The MLP is a type of neural network, which is a universal approximator that can map any relationship between a system's inputs and outputs, regardless of its complexity. Its working principle is based on the human brain, assigning weights (w) to each input to indicate its significance to the output during training. Each node is referred to as a neuron and has its activation function. The number of neurons, network layers, and activation functions all vary depending on the application and impact model prediction performance. The MLP utilized has one hidden layer with 100 neurons, where it attempts to learn up to 500 times to find the best solution.

C. Model Evaluation

Performance metrics were employed to compare and evaluate model prediction accuracy. The SOC estimation output in the intended application is a continuous value. Thus, all performance metrics are for regression and not classification. There are several methods for evaluating:

1) The Mean Absolute Percentage Error (MAPE) measures the accuracy of a prediction system. It expresses this accuracy as a percentage and is calculated as the average absolute percent error for each period minus actual values divided by actual values. The lower the MAPE number, the better a model can forecast values, as shown in Eq. 3:

$$MAPE = \frac{1}{n}\sum_{t=1}^{n} \left| \frac{A_t - F_t}{A_t} \right| \times 100\% \quad (3)$$

Where At represents the actual value, and Ft represents the forecast value. Their difference is divided by At. The absolute value of this ratio is summed for every forecasted point in time and divided by the number of fitted points n1.

2) Root Mean Square Error (RMSE) is the square root of MSE. It calculates the standard deviation of the residuals from the model's fit line. The RMSE statistic provides information on a model's short-term performance by allowing a term-by-term comparison of the actual difference between estimated and measured values. The lower the value, as with

MSE, the better the model's performance. Where n is the number of data points in both equations, as shown in Eq. 4:

$$RMSE = \sqrt{\frac{\sum_{i}^{n}=1\,(P_i - O_i)^2}{n}} \quad (4)$$

Where Pi is the predicted value, and Oi is the observed value for the ith observation in the dataset. The sample size is defined by n.

3) The mean absolute error (MAE), like the previous two metrics, measures the error for continuous predictions, as shown in Eq.5:

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n} \mid y_i - \hat{y}_i \mid \quad (5)$$

Where yi is the actual value, y^i is the predicted value, and n is the number of observations.

IV. Result and Discussion

Table 1 demonstrates the exceptional performance of the four proposed models. Even though they all share similar low errors, their computing time and resources required differ. A high performance of MLR indicates that the dataset is well prepared and processed. Where powerful algorithms such as MLP, SVR, and RF can manage certain levels of unprocessed and complex datasets. However, this will come at the cost of the utilized resources and time consumed.

TABLE 1. PERFORMANCE EVALUATION OF THE FOUR PROPOSED MODELS

Model/Error	MAE	MAPE	RMSE	R2	Training Time (sec)	Testing Time (sec)
MLR	0.75	0.7%	1.2%	0.87	1	0.01
MLP	0.73	0.75%	1.05%	0.89	517	4
SVR	0.7	0.8%	0.975%	0.90	990	12
RF	0.7	0.72%	1.05%	0.88	490	8

Table 1 presents the remarkable efficacy of the four suggested models in estimating SOC. An exceptionally high degree of performance was noted for the four models (MLR, MLP, SVR, and RF). With MAE in the range between 0.7 to 0.75, every model showed remarkable precision. These outcomes show how carefully the dataset was prepared and processed to guarantee the highest possible quality.

The computing efficiency of the implanted algorithms varied, though, even though their accuracy levels were comparable. One second was all that the MLR model needed to train and 0.01 seconds to evaluate, which made it stand out for its exceptional speed. On the other hand, more sophisticated models like MLP, SVR, and RF took a lot longer to train and test. MLP and SVR took more than 500 and 990 seconds, respectively.

The MLR model's low resource requirements and computational efficiency make it particularly suitable for real-time SOC estimation in battery management systems (BMS) for electric vehicles (EVs). This efficiency allows the MLR model to be integrated into existing BMS without significant hardware upgrades. The successful data pretreatment efforts are attested to by the good MLR performance, indicating that the dataset is suitable for linear modeling.

In cases where the data has been prepared thoroughly, the simplicity and speed of MLR provide a significant advantage. It attains high performance without the computational burden posed by more complex models. This choice underscores the value of thorough data preparation, which can often reduce the need for more elaborate modeling techniques.

However, it is important to note that while the MLR model offers computational efficiency, it may not capture the full complexity of relationships between variables in more intricate datasets. This limitation could affect accuracy in scenarios involving extreme conditions or irregular usage patterns. Furthermore, the dataset used is specific to the Panasonic18650PF Li-ion battery, which may limit the generalizability of the models to other battery types or configurations.

Further investigation, the Violin plot of residuals for the MLR was plotted to demonstrate the distribution of the errors (residuals) as shown in Fig. 5. It can be seen clearly that residuals are concentrated around zero, which indicates the high accuracy of the model predictions. In addition, Fig. 6 represents the MLR model prediction error plot, which shows the difference between the actual and prediction values. The residuals are fairly distributed around the horizontal line which indicates a well-performed model in most of the cases.

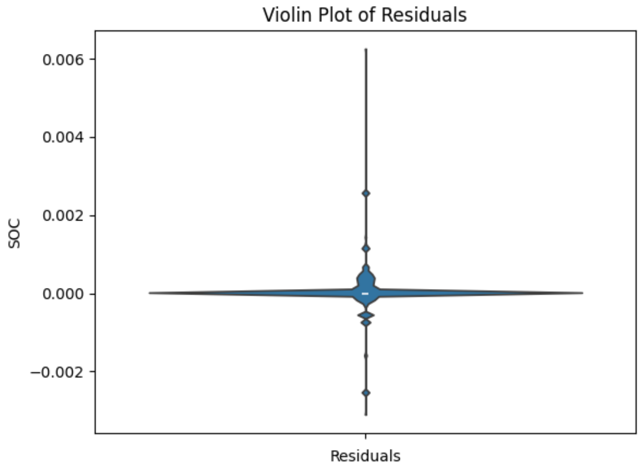

Fig. 5. MLR Violin plot of residuals.

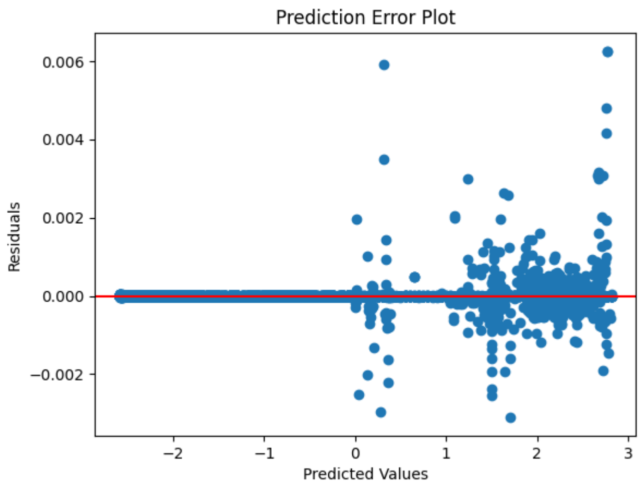

Fig. 6. MLR prediction error plot.

V. CONCLUSION

In conclusion, providing an efficient SOC estimation for electric vehicles is of utmost importance for several reasons such as preventing disastrous situations, and enhancing diver experience. This work contributes to the field by providing a brief understanding of how SOC estimation was predicted before and after the era of ML. The findings underscore the critical role that accurate SOC estimation plays not only in vehicle performance but also in the safety and reliability of electric vehicles, which is increasingly important as the adoption of EVs continues to grow globally.

Adding to that, Four ML-based SOC models using the Panasonic 18650PF dataset were tested and evaluated. The performance of these models was assessed and compared, highlighting their superiority over existing literature. Notably, MLR, MLP, SVR, and RF achieved comparable results in MAE with approximately 0.7, where MLR was the fastest in both training and testing, thus chosen. These results indicate that while traditional methods may still have merit, machine learning approaches offer significant advantages in terms of speed and accuracy, paving the way for their adoption in real-world applications.

Moving forward, it is essential to explore the integration of various deep learning algorithms that can potentially enhance prediction accuracy, particularly in more complex scenarios. Additionally, investigating battery health assessment and remaining useful life estimation could provide more comprehensive insights into battery performance and longevity. Such research could significantly impact the development of smarter, more efficient battery management systems, ultimately contributing to the sustainability and reliability of electric vehicles. Future studies should also consider diverse datasets, including various battery chemistries and configurations, to enhance the generalizability of the models.

REFERENCES

[1] Y. Ding, Z. P. Cano, A. Yu, J. Lu, and Z. Chen, "Automotive Li-Ion Batteries: Current Status and Future Perspectives," Electrochemical Energy Reviews, vol. 2, no. 1, pp. 1–28, 2019, doi: 10.1007/s41918-018-0022-z.

[2] J. Duan et al., "Building Safe Lithium-Ion Batteries for Electric Vehicles: A Review," Electrochemical Energy Reviews, vol. 3, no. 1, pp. 1–42, 2020, doi: 10.1007/s41918-019-00060-4.

[3] K. W. E. Cheng, B. P. Divakar, H. Wu, K. Ding, and H. F. Ho, "Battery-Management System (BMS) and SOC Development for Electrical Vehicles," IEEE Trans Veh Technol, vol. 60, no. 1, pp. 76–88, 2011, doi: 10.1109/TVT.2010.2089647.

[4] H. Musbah, H. H. Aly, and T. A. Little, "Energy management of hybrid energy system sources based on machine learning classification algorithms," Electric Power Systems Research, vol. 199, p. 107436, 2021, doi: https://doi.org/10.1016/j.epsr.2021.107436.

[5] S. Park et al., "Review of state-of-the-art battery state estimation technologies for battery management systems of stationary energy storage systems," Journal of Power Electronics, vol. 20, no. 6, pp. 1526–1540, 2020, doi: 10.1007/s43236-020-00122-7.

[6] M. A. Hannan, Md. M. Hoque, A. Hussain, Y. Yusof, and P. J. Ker, "State-of-the-Art and Energy Management System of Lithium-Ion Batteries in Electric Vehicle Applications: Issues and Recommendations," IEEE Access, vol. 6, pp. 19362–19378, 2018, doi: 10.1109/ACCESS.2018.2817655.

[7] H. Y. Youssef, L. A. Alkhaja, H. H. Almazrouei, A. B. Nassif, C. Ghenai, and M. A. AlShabi, "A machine learning approach for state-of-charge estimation of Li-ion batteries," in Artificial Intelligence and Machine Learning for Multi-Domain Operations Applications IV, SPIE, 2022, pp. 674–682.

[8] P. J. Kollmeyer, "Panasonic 18650PF Li-ion Battery Data," Mendeley Data.

[9] A. B. de Lima, M. B. C. Salles, and J. R. Cardoso, "State-of-Charge Estimation of a Li-Ion Battery using Deep Forward Neural Networks," Sep. 2020, [Online]. Available: http://arxiv.org/abs/2009.09543

[10] A. B. de Lima, M. B. C. Salles, and J. R. Cardoso, "Data-driven state-of-charge estimation of the Panasonic 18650PF Li-ion cell using deep forward neural networks," in 2021 14th IEEE International Conference on Industry Applications, INDUSCON 2021 - Proceedings, Institute of Electrical and Electronics Engineers Inc., Aug. 2021, pp. 1546–1550. doi: 10.1109/INDUSCON51756.2021.9529774.

[11] E. Chemali, P. J. Kollmeyer, M. Preindl, and A. Emadi, "State-of-charge estimation of Li-ion batteries using deep neural networks: A machine learning approach," J Power Sources, vol. 400, pp. 242–255, Oct. 2018, doi: 10.1016/j.jpowsour.2018.06.104.

[12] M. Darwish, S. Ioannou, A. Janbey, H. Amreiz, and C. C. Marouchos, "Review of Battery Management Systems," in 2021 International Conference on Electrical, Computer, Communications and Mechatronics Engineering (ICECCME), 2021, pp. 1–6. doi: 10.1109/ICECCME52200.2021.9590884.

[13] M. Zhang and X. Fan, "Review on the state of charge estimation methods for electric vehicle battery," World Electric Vehicle Journal, vol. 11, no. 1. MDPI AG, 2020. doi: 10.3390/WEVJ11010023.

[14] C. Chinese Association of Automation. Youth Academic Annual Conference (33rd : 2018 : Nanjing Shi, M. IEEE Systems, Chinese Association of Automation, and Institute of Electrical and Electronics Engineers, Lithium-ion Battery SoC Estimation Based on Online Support Vector Regression.

[15] M. A. Hannan, M. S. H. Lipu, A. Hussain, and A. Mohamed, "A review of lithium-ion battery state of charge estimation and management system in electric vehicle applications: Challenges and recommendations," Renewable and Sustainable Energy Reviews, vol. 78. Elsevier Ltd, pp. 834–854, 2017. doi: 10.1016/j.rser.2017.05.001.

[16] A. M. Eid, A. B. Nassif, B. Soudan, and M. N. Injadat, "IIoT Network Intrusion Detection Using Machine Learning," in 6th International Conference on Intelligent Robotics and Control Engineering, IRCE 2023, 2023. doi: 10.1109/IRCE59430.2023.10255088.

[17] D. Saji, P. S. Babu, and K. Ilango, "SoC Estimation of Lithium Ion Battery Using Combined Coulomb Counting and Fuzzy Logic Method," in 2019 4th International Conference on Recent Trends on Electronics, Information, Communication Technology (RTEICT), 2019, pp. 948–952. doi: 10.1109/RTEICT46194.2019.9016956.

[18] S. Neupert and J. Kowal, "Model-Based State-of-Charge and State-of-Health Estimation Algorithms Utilizing a New Free Lithium-Ion Battery Cell Dataset for Benchmarking Purposes," Batteries, vol. 9, no. 7, p. 364, Jul. 2023, doi: 10.3390/batteries9070364.

[19] A. M. Eid, B. Soudan, A. B. Nassif, and M. Injadat, "Correction: Comparative study of ML models for IIoT intrusion detection: impact of data preprocessing and balancing," Neural Comput Appl, vol. 36, no. 19, p. 11661, 2024, doi: 10.1007/s00521-024-09841-5.

[20] A. M. Eid, B. Soudan, A. B. Nassif, and M. Injadat, "Enhancing intrusion detection in IIoT: optimized CNN model with multi-class SMOTE balancing," Neural Comput Appl, vol. 36, no. 24, pp. 14643–14659, 2024, doi: 10.1007/s00521-024-09857-x.