

Choice of Efficiency and Economy: The Role of Machine Learning in Advancing Battery Technologies

Fangzhou Xu^{1,a,*}

¹*National University of Singapore (NUS), 21 Lower Kent Ridge Road, Singapore, 119077, Republic of Singapore*

a. XUFA0003@e.ntu.edu.sg

**corresponding author*

Abstract: With the growth in global energy demands and the prominent need for sustainable development, battery technology has attracted significant attention as a crucial means of energy storage. Applying machine learning in battery technology offers new avenues for enhancing battery performance, safety, and lifespan. Through a literature review, this study comprehensively analyses the current status of machine learning in battery capacity prediction, solid electrolyte research and new material screening. It also explores the related research achievements, existing problems and future development directions, emphasising the importance of high-quality data, appropriate model selection, strict validation and interdisciplinary collaboration in promoting the progress of battery technology.

Keywords: Machine learning, Battery technology, Energy storage, Materials science

1. Introduction

In the pursuit of a sustainable and energy-efficient global future, the role of battery technology has become increasingly pivotal. As the demand for clean energy solutions surges, batteries have emerged as the cornerstone of various applications, from revolutionising the automotive industry with electric vehicles to enabling the efficient storage and utilisation of renewable energy sources. Their performance, safety and longevity are not only crucial for technological advancements but also for environmental preservation and economic viability.

Traditional battery technologies, predominantly relying on liquid electrolytes, have long been encumbered by various challenges. Safety issues, such as the risk of leakage and the potential for thermal runaway, have cast a shadow over their widespread adoption. Moreover, limitations in energy density and stability have hindered their ability to meet the ever-escalating demands of modern applications [1,2]. In contrast, solid electrolytes have emerged as a promising alternative, offering enhanced safety profiles due to their non-flammable nature, along with the potential for higher energy densities and extended lifespans. These attributes make them particularly attractive for applications where safety and performance are of utmost importance, such as in consumer electronics and the rapidly expanding electric vehicle market.

However, the path to optimising battery technologies has been fraught with difficulties. The complex and intricate nature of battery chemistries and configurations has posed significant barriers to traditional design and optimisation methods, impeding the pace of innovation. It is within this

context that machine learning (ML) has emerged as a revolutionary force and is poised to transform the landscape of battery research and development [3].

The concept of the State of Health (SOH) stands as a linchpin in assessing the performance reliability of lithium-ion batteries, which have become integral in the quest for efficient and sustainable energy storage. SOH not only reflects the current condition of the battery but also serves as a crucial indicator of its remaining lifespan. Traditional methods for estimating SOH, often reliant on empirical models, have proven inadequate in capturing the full spectrum of the complex, nonlinear ageing processes inherent in battery operation.

This shortcoming has spurred the search for more sophisticated and adaptable approaches, one of which lies in the realm of machine learning. ML, with its remarkable ability to dissect vast datasets and unearth hidden patterns, holds the key to unlocking new frontiers in battery technology. By seamlessly integrating ML into the battery design process, researchers can accelerate the discovery of novel materials, precisely predict battery performance under diverse conditions and optimise manufacturing processes with unprecedented efficiency.

This literature review embarks on a journey to explore the recent strides made in the application of ML for battery design. It delves deep into its profound impact on material discovery, performance prediction and system optimisation. Through a meticulous and comprehensive examination of current studies and methodologies, this review endeavours to spotlight the transformative potential of ML in propelling the next generation of battery technologies forward. In doing so, it aims to address the pressing need for sustainable and reliable energy storage solutions, which are essential for a greener and more prosperous future.

The integration of ML into battery design represents a paradigm shift in materials science and engineering. It not only addresses the long-standing challenges in the development of next-generation energy storage technologies but also holds the promise of revolutionising the way we conceive, design and optimise batteries. Central to this transformation is the identification and development of solid electrolytes that possess high ionic conductivities, robust stability and seamless compatibility with electrodes. This review will not only explore the role of ML in predicting the properties of these solid electrolytes but also investigate how it can guide the selection and optimisation process, thereby paving the way for the creation of more efficient and reliable battery systems.

Moreover, as ML continues to permeate the field of battery technology, it is essential to understand the evolution of the models and techniques employed. From the early days of simple regression models to the current state-of-the-art ensemble and stacking techniques, the field has witnessed a remarkable progression towards more accurate and reliable SOH prediction models. This review will trace this evolution, highlighting the key milestones and breakthroughs that have shaped the current landscape. Additionally, it will scrutinise the challenges and opportunities that arise in the data collection process, the development of robust models and the practical implications of ML predictions in the synthesis and testing of promising materials.

Data quality, feature engineering and the delicate balance between model complexity and interpretability are also critical aspects that will be explored in depth. High-quality data serves as the bedrock for accurate ML predictions and this review will examine the strategies and techniques employed to ensure data integrity and relevance. Feature engineering, the process of transforming raw data into meaningful features, plays a crucial role in enhancing the performance of ML models and its significance will be elucidated. Furthermore, the trade-off between model complexity, which often leads to improved accuracy and interpretability and is essential for understanding and validating the models, will be analysed.

By conducting a detailed and in-depth analysis of recent advancements, this review aims to synthesise the findings of leading researchers in the field. It will not only provide a comprehensive overview of the current state of the art but also outline the challenges that lie ahead. Finally, it will

propose a trajectory for future research, offering insights and recommendations that can guide the continued development and application of ML in battery technology. In doing so, it hopes to contribute to the ongoing efforts to develop more sustainable and efficient energy storage solutions, which are vital for a cleaner and more sustainable future.

2. Application Research of Machine Learning in Battery Technology

2.1. Battery Capacity Prediction

Significant progress has been made in predicting battery capacity for electric vehicles. For example, a research team developed a stacking ensemble machine learning model [4]. They utilised daily operational charging data and incorporated domain-specific features derived from incremental capacity (IC) and differential voltage (DV) analyses to handle the complex degradation behaviours of batteries (Figure 1). Four base machine learning models - linear regression, random forest, Gaussian process and gradient boosting regressors - were combined using a meta-learner to form the stacking ensemble model.

This model achieved a mean absolute percentage error (MAPE) of 0.28% and a root mean squared percentage error (RMSPE) of 0.55% in capacity estimation tasks, demonstrating high predictive power and low estimation errors. However, the model has limitations. It relies on IC and DV analysis, which may not be applicable in all real-world scenarios such as partial charging cycles and fast charging modes. Future research could focus on exploring more advanced machine learning techniques like physics-informed neural networks (PINN) to overcome these limitations and improve the accuracy and applicability of battery capacity prediction models in various practical situations.

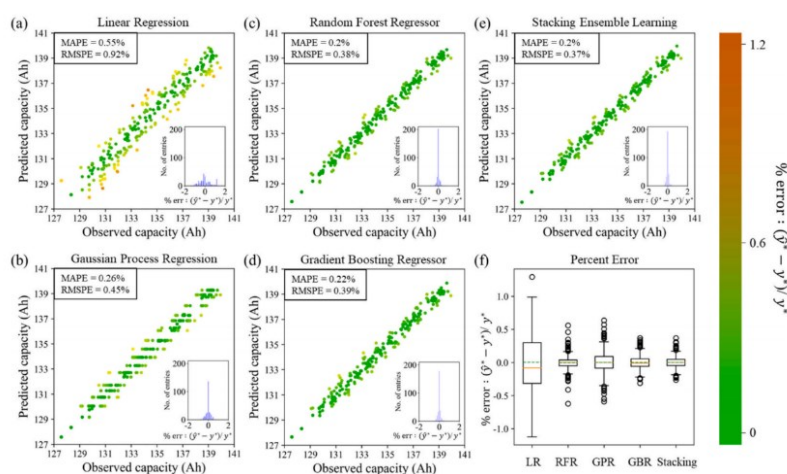


Figure 1: Estimation performance for training set using different machine learning techniques. (a), (b), (c) and (d) are the estimation results with linear regression, Gaussian process regression, random forest regressor and gradient boosting regressor. (e) shows the estimation results using the stacking model. (f) is the average percentage error of each machine learning model.

2.2. Solid Electrolyte Research

Regarding the research on lithium solid electrolytes, efforts have been made to construct a comprehensive database. A database of lithium solid electrolyte conductivities was created, which contains 820 entries from 214 sources, covering a wide range of chemical compositions and temperature ranges [5]. Machine learning techniques were employed to analyse this data. For instance, Compositionally Restricted Attention Based Networks (CrabNets) were used to predict the ionic conductivity of new potential solid electrolyte materials (Figure 2).

This approach has advanced the understanding of electrolyte properties crucial for lithium-ion battery technology. Nevertheless, challenges exist. Data quality and diversity issues were encountered and these factors can impact the effectiveness of machine learning predictions. Future research should address these challenges by improving data collection and preprocessing methods to ensure the reliability and accuracy of machine learning models in predicting the properties of lithium solid electrolytes.

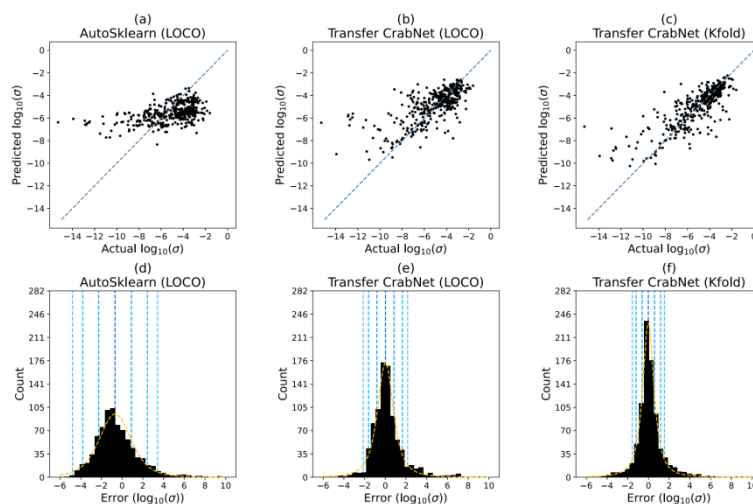


Figure 2: Comparative Analysis of Regression Model Predictions and Performance: This figure presents a visual evaluation of three machine learning regression models - AutoSklearn and Transfer CrabNet - using LOCO-CV and K-fold cross-validation methods. Subplots (a), (b), and (c) depict parity plots showing the relationship between actual and predicted $\log_{10}(\sigma)$ values, while subplots (d), (e) and (f) illustrate the corresponding error distributions. The AutoSklearn model shows a wider spread of predictions, indicating a less precise fit, as seen in (a) and (d). In contrast, the Transfer CrabNet models, particularly when validated with K-fold as shown in (c) and (f), demonstrate a tighter clustering of data points around the line of perfect prediction and a more concentrated error distribution, suggesting superior predictive accuracy and model performance.

2.3. New Material Screening

In the screening of new materials for solid lithium-ion conductors, a large-scale computational screening method was introduced. A study utilised logistic regression models to evaluate over 12,000 lithium-containing crystalline solids based on criteria such as structural and chemical stability, low electronic conductivity and cost-efficiency [6]. This screening reduced the list of candidate materials from 12,831 to 21 promising structures, most of which had not been experimentally examined before. This study demonstrated the power of integrating computational modelling with empirical research.

However, the transition from theoretical or computational models to practical applications still faces hurdles. Future research needs to focus on overcoming challenges such as scalability, manufacturing processes and cost-effectiveness to accelerate the discovery and application of high-performance solid electrolytes for lithium-ion batteries.

3. Challenges and Future Directions

3.1. Data Quality and Representation

High-quality and comprehensive datasets are essential for training accurate machine learning models. However, current datasets may have limitations in representativeness for real-world applications. The

data used in battery research often comes from specific experimental conditions or sources, which may not fully capture the complexity and variability of actual battery usage scenarios. For example, in battery capacity prediction, the data may not cover all possible charging and discharging patterns, temperature variations and ageing conditions.

To address this issue, future research should focus on collecting more diverse and representative data. This could involve conducting experiments under a wider range of conditions, integrating data from multiple sources and using advanced data augmentation techniques. Additionally, improving data preprocessing methods to handle noise, outliers and missing values is crucial to enhance the quality of the data and the performance of machine learning models.

3.2. Methodological Approaches

The variety of machine learning models used in battery technology research indicates the flexibility and adaptability of these techniques. However, selecting the most appropriate model for a specific problem remains a challenge. Different models have their own strengths and weaknesses and their performance may vary depending on the characteristics of the data and the nature of the problem. For instance, while stacking ensemble models have shown good performance in predicting battery capacity, they may be computationally expensive and complex to interpret. In contrast, simpler models like linear regression may be more interpretable but may not capture the complex relationships in the data as effectively.

Future research should focus on developing a better understanding of the relationship between different machine learning models and battery-related problems. This could involve conducting more comparative studies to evaluate the performance of different models under various conditions. Additionally, hybrid models that combine the advantages of multiple models could be explored to improve the accuracy and interpretability of predictions.

3.3. Predictive Performance and Validation

Although the reported results in the literature show high accuracy in some cases, the generalisation ability of machine learning models to unseen data is a crucial factor that needs to be further validated. In many studies, the lack of clear validation data or real-world testing is a common problem. Overfitting, where the model performs well on the training data but fails to generalise to new data, is a potential risk. To enhance the predictive performance and reliability of models, more rigorous validation techniques should be employed. This includes using larger and more diverse validation datasets, implementing cross-validation methods more effectively and conducting long-term and real-world testing. Additionally, developing performance metrics that better reflect the practical requirements of battery applications, such as the ability to predict battery performance under different operating conditions and over the entire lifespan of the battery, is essential.

3.4. Application and Impact

The potential applications of machine learning in battery technology are vast, ranging from improving electric vehicle battery performance to accelerating the discovery of new materials. However, transitioning from theoretical or computational models to practical applications is not straightforward. Scalability is a significant challenge, especially when it comes to implementing machine learning models in large-scale battery manufacturing and management systems. The manufacturing processes need to be adapted to incorporate machine learning-based optimisation, which may require significant changes in production lines and quality control procedures.

Cost-effectiveness is another crucial factor. The implementation of machine learning techniques may involve additional costs in terms of data collection, computing resources and model development.

Future research should focus on developing strategies to overcome these challenges. This could involve collaborating with industry partners to test and optimise machine learning models in real manufacturing and application scenarios, exploring cost-effective data collection and processing methods and developing more efficient machine learning algorithms that can be implemented with limited resources.

3.5. Ethical Considerations and Sustainability

The advancement of battery technologies through machine learning also raises important ethical and sustainability questions. The production of batteries and the extraction of raw materials for battery components can have significant environmental impacts. Additionally, the disposal and recycling of batteries at the end of their lifecycle need to be carefully considered. Machine learning models that predict battery performance and guide material selection may influence the choice of materials and manufacturing processes, which in turn can affect the environmental footprint of batteries.

Future research should incorporate ethical and sustainability considerations into the development of machine learning models. This could involve evaluating the environmental impact of different battery materials and manufacturing processes predicted by machine learning models, developing strategies to minimise the use of scarce or environmentally harmful materials and promoting the recycling and reuse of batteries through machine learning-based optimisation of recycling processes.

4. Conclusion

Machine learning has emerged as a powerful tool in the field of battery technology, offering significant opportunities for improving battery performance, discovering new materials and optimising battery systems. The research reviewed in this paper has demonstrated the potential of machine learning in various aspects of battery technology, from capacity prediction to solid electrolyte research and new material screening. However, several challenges need to be addressed to fully realise the potential of machine learning in practical battery applications.

These challenges include improving data quality and representation, selecting appropriate methodological approaches, enhancing predictive performance and validation, overcoming application and impact hurdles, and considering ethical and sustainability issues. By addressing these challenges through interdisciplinary research and collaboration between academia and industry, the future of battery technology can be significantly enhanced, leading to more sustainable and efficient energy storage solutions.

References

- [1] Tarascon, J.-M., & Armand, M. (2001). Issues and challenges facing rechargeable lithium batteries. *Nature*, 414, 359-367.
- [2] Goodenough, J. B., & Kim, Y. (2011). Challenges for rechargeable batteries. *Journal of Power Sources*, 196(16), 6688-6694.
- [3] Juan, Y., Dai, Y., Yang, Y., & Zhang, J. (2021). Accelerating materials discovery using machine learning. *Journal of Materials Science & Technology*, 79, 178-190.
- [4] Zhao, J., Ling, H., Liu, J. et al. (2023) Machine learning for predicting battery capacity for electric vehicles. *eTransportation*, 15, 100214.
- [5] Hargreaves, C. J., Gaultois, M. W., Daniels, L. M. et al. (2023) A database of experimentally measured lithium solid electrolyte conductivities evaluated with machine learning. *npj Computational Materials*, 9, 9.
- [6] Sendek, A. D., Yang, Q., Cubuk, E. D. et al. (2016) Holistic computational structure screening of more than 12 000 candidates for solid lithium-ion conductor materials. *Energy & Environmental Science* 9(12), 3697-370.