A Data-Driven Machine Learning Method for Battery Capacity Estimation

Bachelorarbeit

zur Erlangung des akademischen Grades „Bachelor of Science (B.Sc.)" im Studiengang Wirtschaftsingenieur der Fakultät für Elektrotechnik und Informatik, Fakultät für Maschinenbau und der Wirtschaftswissenschaftlichen Fakultät der Leibniz Universität Hannover

vorgelegt von

Name: Hassani

Vorname: Maria

[redacted]

in: [redacted]

Prüfer: Prof. Dr. Michael H. Breitner
Betreuer: M. Sc. Yifan Wang

Hannover, den 01. Juli 2024

Table of Contents

1. Introduction

This chapter provides an overview of the key elements of this bachelor thesis, including the background of the topic, the problem statement, the motivation and contribution of the work and the research questions. By examining the underlying research questions, the research goals are defined. Additionally, the structure of this thesis is presented.

1.1 Background of the Topic

The depletion of fossil fuels and the rise in environmental pollution have made the development of clean and sustainable energy sources, along with new technologies for energy conversion and storage, increasingly urgent (Kumar *et al.*, 2018). In recent years, some of the most effective conversion and power-saving technologies such as lithium-ion batteries (LIBs), supercapacitors, and fuel cells have been rapidly advanced and widely used in electric vehicles, mobile phones, and other consumer electronic devices (Kumar et al., 2019). Owing to their high power density, consistent performance, long lifespan, and low self-discharge rate (Kumar et al., 2018b), LIBs have demonstrated significant potential for numerous applications in electric vehicles and they also contribute to the pollution challenges associated with the lifecycle of these vehicles (Tan et al., 2020) . However, the storage capacity of LIBs diminishes with prolonged use due to their inherent electrochemical properties. This degradation results in increased safety risks and higher operational costs in electric propulsion systems (Hu et al., 2020). To address these challenges, the implementation of a battery management system (BMS) is essential. The BMS provides online supervision and verification, enhancing the performance, health, and safety of LIBs-based electric drive systems (Wassiliadis et al., 2023). A critical function of BMS is the accurate estimation of the state of health (SOH), which has become increasingly essential for ensuring the safe operation of electric vehicles (Voronov et al., 2018). The SOH is a crucial index for assessing battery service life, reflecting both its aging and reliability. However, due to the complex internal degradation mechanisms of batteries, direct measurement of SOH is not possible, making accurate SOH estimation a significant challenge (Eddahech et al., 2012).

1.2 Problem Statement

Accurate SOH estimation is essential for the effective management and performance of the LIBs, especially in applications such as electric vehicles and portable electronic devices. Methods such as equivalent circuit models-based and electrochemical models-based (Li et al., 2022) are traditional approaches for capacity estimation and have limitations. In contrast, data-driven capacity estimation methods are highly universal and flexible, so they have gained significant attention (Sui et al., 2021).

1.3 Motivation and Contribution of the work

The motivation for this research stems from the urgent need to address the persistent challenges associated with the efficiency and longevity of LIBs in electric vehicles, mobile phones, and other consumer electronic devices Despite significant advancements in energy storage technologies, there is still a notable gap in developing comprehensive solutions for accurately estimating the SOH of these batteries. This gap has profound implications for safety, operational costs, and performance reliability across various applications. In recent years, the rapid adoption of electric vehicles, advancements in consumer electronics, and the global push for sustainable energy solutions have spurred increased interest in developing more advanced BMS. However, current approaches often fall short due to their limitations in accurately monitoring and predicting the complex degradation mechanisms within LIBs. These shortcomings hinder the overall effectiveness of BMS, affecting the efficiency, safety, and reliability of electric propulsion systems and consumer electronics. Addressing these challenges is crucial. Enhancing the accuracy of SOH estimation can lead to significant improvements in battery management, contributing to the broader goal of sustainable and clean energy usage. By focusing on these issues, this research aims to contribute to the development of more reliable and efficient technologies for electric vehicles, mobile phones, and other consumer electronics. Ultimately, this work supports the transition to greener transportation solutions and more sustainable electronic devices.

This thesis makes a number of important contributions to the field of battery SOH estimation. The primary contributions are as follows:

A thorough analysis of different hyperparameters for the Long Short-Term Memory (LSTM) model is conducted, along with the calculation of the Importance Ip index. This provides new insights into model optimization and identifies the most significant features in SOH prediction. The proposed solution can be applied to developing more reliable and efficient technologies for electric vehicles, mobile phones, and other consumer electronics. By demonstrating its effectiveness in improving SOH prediction, this practical implementation highlights the potential for real-world impact in these industries. By addressing the gaps in current research and providing robust solutions, this work lays the foundation for future studies and practical implementations in battery SOH estimation. The contributions of this thesis are expected to pave the way for more advanced research and development, ultimately enhancing the accuracy of battery health prediction.

1.4 Research Questions

To address the challenges and opportunities identified in the background of the topic and problem statement sections; the main goal of this bachelor thesis is to develop an effective data-driven machine learning (ML) method for battery capacity estimation using the LSTM model on the NASA battery dataset. The research will address the following questions:

What is the suitable parameter setting for LSTM to predict SOH?

Which feature plays the most important role in battery SOH prediction?

1.5 Thesis Structure

This bachelor's thesis is divided into six chapters. The first chapter provides a brief overview of the background of the topic, the problem statement, the motivation and contribution of the work, the research questions, and the thesis structure. The second chapter presents the literature review and is divided into two sections: the basics of battery technology and capacity estimation (SOH) methods, as well as the definition and importance of ML in battery capacity estimation. It also explains how ML is used in battery analysis and gives an overview of relevant concepts and algorithms. The methodology is presented in the third chapter. It begins with an explanation of data collection and preprocessing, followed by feature engineering. The model selection and architecture are then explained. The methodology section continues with the training procedure and evaluation metrics, culminating with the presentation of the proposed framework. Chapter four explains the experimental results, including the performance evaluation of the proposed method and the analysis of results. Chapter five discusses the interpretation of the findings, the implications of the study, and its practical applications. The thesis concludes with a summary of the most important highlights and discusses recommendations for future research.

6. Conclusion and Future Research

The conclusion of this thesis presents a comprehensive summary of its key contributions, findings, limitations, and recommendations for future research. It highlights the successful development and validation of a data-driven ML method for estimating the SOH of LIBs, demonstrating the effectiveness of the LSTM model. The conclusion also outlines potential areas for further research aimed at enhancing the model's general applicability, accuracy, and practical implementation in real-world BMSs.

6.1 Summary of Contributions

This thesis has introduced a data-driven ML method for estimating the SOH of LIBs using the LSTM model. By making use of the NASA battery dataset, this research has been able to successfully demonstrate the potential of the LSTM model to provide accurate and reliable predictions of SOH. The key contributions will reflect a well-summarized framework whereby the data preprocessing, feature engineering, model training, and evaluation processes have been done. Results showed the need to select an appropriate ML technique and optimize the hyperparameters for better prediction accuracy and model performance.

6.2 Key Findings

The research yielded several key insights. Firstly, the LSTM model significantly outperformed traditional methods for estimating the SOH of batteries, demonstrating its ability to capture complex temporal patterns in data. The importance of feature engineering and data normalization was highlighted, as these processes greatly enhanced the model's prediction capabilities. The findings indicate that optimal performance is achievable through careful hyperparameter optimization, particularly regarding the number of epochs, batch sizes, and neuron configurations. Additionally, the calculation of the importance index was crucial for identifying the most significant features for SOH estimation. Finally, the experimental results confirmed that the proposed approach provides a robust and scalable solution for real-time SOH estimation in practical applications.

6.3 Limitations of the work

Despite the promising results and potential applications of this thesis, there are several limitations that should be acknowledged. Due to constraints such as limited time and resources, certain aspects of the research were not as comprehensive as desired.

Limited Dataset:

The thesis relies heavily on the NASA dataset, which, while comprehensive, does not encompass the diversity of all possible battery conditions and types. Only four specific batteries (B0005, B0006, B0007, and B0018) were used in this research, which may limit the generalizability of the findings. A broader dataset with varied battery types and conditions could provide more robust and widely applicable results. The lack of multiple datasets restricts the ability to validate the model across different battery chemistries and operational conditions. Incorporating additional datasets from various sources would enhance the reliability and applicability of the model.

Model Selection:

Due to time constraints, the thesis employed a relatively simple LSTM model. While effective, the model does not explore the full range of all ML and deep learning techniques available. More sophisticated models, such as hybrid models combining LSTM with other neural network architectures or incorporating attention mechanisms, could potentially yield more accurate predictions. Limited resources and time prevented a comprehensive comparison with other advanced models like Convolutional Neural Networks (CNNs) or Transformer-based models. Future research could benefit from a comparative analysis of these models to identify the most effective approach for SOH estimation.

Scalability and Computational Resources:

The computational resources required for training and deploying more complex models were not available. Future studies could explore the use of high-performance computing environments or cloud-based solutions to overcome these limitations and enable the use of more complex models.

In conclusion, addressing these limitations in future research could lead to more accurate and generalizable models for battery SOH estimation, ultimately enhancing the practical applications and reliability of BMS.

6.4 Recommendations for Future Work

Future research will expand the dataset to include many different types of batteries and their operational conditions, thus making the model more general. Moreover, further investigations on other advanced ML techniques and hybrid models may assist in increasing the accuracy and robustness of the prediction. Implementing the proposed model into real-life BMS is also essential to evaluate its practical utility and performance. Lastly, continuous refinement and optimization of the LSTM model using different architectures and training strategies would be performed for even higher accuracy and efficiency in estimating battery capacity.