



Technische Hochschule
Ingolstadt

MASTER THESIS

“Enhancing Lithium-Ion Battery State
of Health Estimation through Data
Collection”

M. Eng. International Automotive Engineering

Rupen Sarvaiya

2024

Technische Hochschule Ingolstadt

FACULTY OF ELECTRICAL ENGINEERING AND INFORMATION TECHNOLOGY

International Automotive Engineering

Master Thesis

Enhancing Lithium-Ion Battery State of Health Estimation through Data Collection

in collaboration with
CARISSMA Forschungszentrum

Author
Rupen Ashokbhai Sarvaiya
Matr. No. : 00101121

Supervisor
Carlos antônio rufino júnior

First Examiner
Prof. Dr. Hans-Georg Schweiger

Second Examiner
Prof. Dr. Armin Arnold

Issued on: 12.10.2023

Submitted on: 05.02.2024

DECLARATION

I declare that I have authored this thesis independently, that I have not used other than the declared sources/resources, that I have not presented it elsewhere for examination purposes, and that I have explicitly indicated all material that has been quoted either literally or by consent from the sources used. I have marked verbatim and indirect quotations as such.

Ingolstadt, 05.02.2024

A handwritten signature in black ink, appearing to read 'Rupen Ashokbhai Sarvaiya', written over a horizontal line.

Ingolstadt, _____

Rupen Ashokbhai Sarvaiya.

ACKNOWLEDGMENTS

I would like to express my profound appreciation to the CARISSMA Research Centre for giving me the chance to pursue my master's thesis at their research centre. The tremendous knowledge and resources they offered were crucial to the effective completion of my research.

I would like to express my gratitude to my supervisor, Carlos Antônio Rufino Júnior, for his invaluable assistance and direction during the entire duration of my thesis. The valuable input and deep insights he provided had a vital role in improving the quality of my work.

I express equal gratitude to my academic supervisors, Prof. Dr. Hans-Georg Schweiger and Prof. Dr. Armin Arnold, for their proficient guidance and commendations. Their inputs played a crucial role in enhancing the quality of my research. I would like to sincerely thank my freinds rahil and ronak for their unwavering encouragement, inspiration, and support throughout my academic journey. They have consistently served as a source of inspiration, lending a sympathetic ear and wise counsel when I needed it. Their commitment to ensuring my success has been genuinely priceless, and I appreciate everything they have done for me.

A particular thank you also goes to my parents, bhavna and ashok, who have always supported and encouraged me in my career. I doubt that without them and their assistance, I would be where I am now.

ABSTRACT

The increasing demand for electric vehicles EV in recent years has led to a growing need for advanced BMS that can accurately estimate the state of health SOH of batteries. The SOH is a critical parameter that determines the performance and lifespan of batteries, and accurate estimation of these parameters is essential for optimizing battery utilization and improving the overall efficiency and reliability of EV. Accurately estimating the SOH of batteries in real driving conditions is a challenging task due to the dynamic nature of driving cycles, which can cause significant variations in battery behavior. Moreover, the accuracy of existing estimation techniques is often affected by factors such as battery degradation, temperature variations, and non-linearities in battery behavior.

To address these challenges, researchers and engineers have developed a wide range of techniques and algorithms for estimating the SOH of batteries in real driving conditions. These techniques include model-based approaches, data-driven methods, and hybrid techniques that combine both model-based and data-driven approaches. The objective of this master thesis is to critically review the existing literature on estimation techniques for SOH in real driving cycles, identify the strengths and limitations of different approaches, and propose a novel estimation technique that can overcome the limitations of existing approaches. The proposed technique will be evaluated using real-world data obtained from a test vehicle.

It is vital to do a precise assessment of the condition of these batteries in order to guarantee that they can be used safely and to prevent explosions that may possibly be catastrophic. The challenges that were discussed before could be solved with the assistance of prediction models. The purpose of this research is to evaluate the accuracy of predictions made by a variety of machine learning algorithms on the state of the battery. In order to achieve this result, time series forecasting techniques are used to data metrics. It was shown that Long Short-Term Memory LSTM models perform very well when it comes to the creation of forecasts that can be relied upon. An accurate forecast made with the aid of machine learning models may assist in increasing sales of electric vehicles and ensuring that these batteries are used in a secure manner.

ACRONYMS

AI	Artificial Intelligence
ANN	Artificial Neural Network
BEV	Battery Electric Vehicle
BMS	Battery Management System
BPNN	Back Propagation Neural Network
BTS	Battery Testing System
BYD	Build Your Dreams
C-ECOS	Carrisma Institute Of Electric connected And Secure Mobility
CC-CV	Constant Current-Constant Voltage
CIA	Capacity Incremental Analysis
DDM	Data Driven Models
DL	Deep Learning
DOD	Depth Of Discharge
DVA	Differential Voltage Analysis
ECM	Equipment Circuit Models
EIS	Electrochemical Impedance Spectroscopy
ELM	Extreme Learning Machine
EM	Electrochemical Models
EOL	End Of Life
EV	Begining Of Life
EV	Electric Vehicle
GNN	Gated Neural Network
GP	Gaussian Process
GRU	Gated Recurrent Uint
HEV	Electric Vehicle
HM	Hybrid Models
HPPC	Hybrid Pulse Power Characterisation
IC	Integrated Circuit
LCO	Lithium Cobalt Oxide
LDV	Light Duty Vehicle
LFP	Lithium Ion Phosphate
LIB	Lithium Ion Battery
LMO	Lithium Manganese Oxide
LSTM	Long Short Term Memory

MAE	Mean Absolutr Error
ML	Machine Learning
MSE	Mean Squared Error
NCA	Lithium Nickel Cobalt Aluminium Oxide
NCM	Nickel Cobalt Manganese
NMC	Lithium Nickel Manganese Cobalt Oxide
NN	Neural Network
OCV	Open Circuit Voltage
OEM	Original Equipment Manufacturer
P2D	Pseudo 2D Model
PBM	Physics Based Models
PHEV	Plugin Hybrid Electric Vehicle
RBF	Radial Basis Functions
RC	Resistor Capacitor
RMSE	Root Mean Squared Error
RNN	Recurrent Neural Network
RUL	Remaining Useful Life
RVM	Relevant Vector Machine
SEI	Solid Electrolyte Interphase
SOC	State Of Charge
SOH	State of Health
SOP	State Of Power
SPM	Single Particle Model
SUV	Sport Utility Vehicle
SVM	Support Vector Machines

List of Figures

1	Global lithium-ion battery demand and sales	2
2	Lithium-ion battery Basic Structure and Internal Detailed Structure	7
3	Prismatic Cells- Cylindrical Cells- Pouch Cells	14
4	RUL Estimator Models	19
5	Battery demand by mode, 2016-2022	20
6	BMS block diagram	21
7	BMS Functionalities	22
8	An unrolled recurrent neural network.	23
9	LSTM Gates.	25
10	LSTM forget gate.	26
11	LSTM input gate.	27
12	LSTM output gates.	27
13	SOH ESTIMATION METHOD	34
14	Sorting battery SOH prediction techniques	35
15	Overview of machine learning approaches	36
16	Classification of ML-based SOH prediction algorithms	37
17	LSTM	40
18	GRU	41
19	Generic implementation flowchart for SOH prediction using ML algorithms	42
20	Diagram of the Individual Cell Arrangement and Experimental Test Setup.	46
21	Diagram illustrating the Configuration of a Controlled Temperature Cham- ber for Cell Testing	47
22	LFP CELL	49
23	Detailed description of cyclic aging data	52
24	Dataset	53
25	LSTM structure	56
26	The schematic diagram of dropout neural network model. (a) standard neu- ral network with 2 hidden layer; (b) neural network after applying Dropout technology	57
27	Cycling aging of cells for different DOD and c-rates	61
28	Cell 01	62
29	Cell 25	62

List of Tables

1	Vehicle Data sheet	39
2	Data sheet of the LFP cell used	48
3	LSTM Model Result	64
4	GRU Model Result	65
5	Linear Regression Model Result	66
6	Polynomial Regression Model Result	67
7	Decision Tree Regression Model Result.	68
8	Random Forest Regression Model Result.	69

Table of Contents

Declaration	I
Acknowledgments	II
Abstract	III
Acronyms	IV
1 Introduction	1
1.1 Motivation	2
1.2 Thesis Objective	3
1.3 Thesis Contribution	4
1.4 Thesis Outline	5
2 Theoretical Background	6
2.1 Li-Ion Battery	6
2.1.1 Lithium-ion Battery Composition	7
2.1.2 Characteristics of Lithium-ion Battery	7
2.2 Most Common Lithium-Ion Batteries found in EVs	9
2.2.1 Li-ion batteries that do not contain cobalt	9
2.2.2 Li-ion batteries that contain Cobalt	11
2.3 Different Types of Li-ion Cells	13
2.3.1 Cylindrical Cells	13
2.3.2 Prismatic Cells	13
2.3.3 Pouch cells	13
2.3.4 Characterization of a LiFePO ₄ battery	14
2.4 Battery Cell Parameters and Specifications	15
2.4.1 Methods for Forecasting Remaining Useful Life	19
2.5 Background Research	20
2.5.1 Battery demand for Electric Vehicles continues to rise	20
2.6 Battery Management Systems (BMS)	21
2.6.1 BMS Functionalities	22
2.7 An Overview of Recurrent Neural Network (RNN) and Long Short-Term Memory (LSTM)	23
2.7.1 What is Recurrent Neural Network (RNN)?	23
2.7.2 What is (LSTM) and How it Works	25
2.7.3 Implementation of a recurrent neural network using Long Short-Term Memory (LSTM)	28
3 Literature Review	29
3.1 Battery Aging Modelling	29
3.1.1 Physics-Based Models (PBMs)	30
3.1.2 Empirical/Semi-Empirical Models	31
3.1.3 Equivalent Circuit Models (ECMs)	31
3.1.4 Electrochemical Models (EMs)	32

3.1.5	Data-Driven Models (DDMs)	32
3.1.6	Hybrid Models (HMs)	33
3.2	Approaches to Predicting the State of Health	33
3.3	Machine Learning for Predicting Battery State	35
3.3.1	Overview of Machine Learning Approaches	36
3.3.2	Machine Learning Algorithms for Battery State Prediction	36
3.3.3	Comparison of ML algorithms in SOH estimation (LSTM),(GRU) and Regression	38
4	Methodology	42
4.1	Test Methodology	42
4.1.1	State of Health (SOH) Prediction Using Machine Learning Algorithms	42
4.2	Test Bench Description	44
4.3	An explanation of the test configuration	44
4.4	Description of the LFP single cells	48
4.4.1	Selecting the charge/discharge current for capacity and energy measurements	49
4.4.2	Loading Procedure	50
4.4.3	Test Plan and Cyclic Aging Data Collection	50
4.5	Feature Engineering	51
4.5.1	Data Description	51
4.5.2	Data processing and Analysis	53
4.5.3	Data Aggregation and Data Cleaning	54
4.5.4	Removing Outliers and Handling Missing Values	55
4.6	Machine Learning Models	55
4.6.1	LSTM Model:	56
4.6.2	Dropout Technique:	57
4.6.3	Model performance metrics:	58
4.7	Model Running Time	59
5	Results and Discussion	60
5.1	Analysis and Visualisation of Data:	60
5.1.1	Cycle Aging with Voltage vs Discharge Capacity	60
5.1.2	Discharge Voltage Analysis	61
5.2	Model Implementation	63
5.2.1	LSTM Model	63
5.2.2	GRU Model	64
5.2.3	Regression Model	66
6	Conclusion and Future Work	71
6.1	Conclusion	71
6.2	Future Work	72
	Literature references	X
A	Appendix	XIX
A.1.1	Hardware Specifications:	XIX

A.1.2 Software utilized:XIX

1. INTRODUCTION

The demand for power lithium batteries will continue to develop rapidly on a global scale in the coming years as traditional fuel vehicle manufacturers expand their presence in the new energy vehicle market. Concurrently, the cost pressure on power lithium battery firms has steadily grown due to the new energy subsidy policy. In response to the rapidly evolving new energy industry, mainstream power lithium battery companies both domestically and internationally have announced a major expansion of their production in an effort to increase their production and sales scale, increase the scale effect, and lower their production costs per unit of product. Lithium-ion batteries appealing properties—such as a large energy storage capacity, rapid charging, and little loss of charge over time, make them ideal for usage in electric and hybrid cars. Manufacturers of parts for new energy cars are beginning to view lithium batteries as essential products, and many major automakers have boosted their research expenditure in this field as the government continues to support the development of these vehicles. A new energy electric vehicle's battery life range and cost are two factors that many buyers will carefully evaluate before making a purchase[39].

There are primarily two aspects to battery aging: calendar aging and cycle aging. The term "calendar aging" describes the long-term impacts of battery storage. Cycle aging, in contrast, is linked to the effects of charging and discharging cycles on battery utilization. Problems with the chemical composition of batteries, such as high temperatures, might arise from time to time. Events of a hazardous nature may result from these conditions. Consequently, in order to reduce hazards, it is crucial to track the age of batteries and evaluate the state of charging and discharging procedures. We can keep potentially disastrous events from happening by examining the battery's health and age on a regular basis[7].

The State of Health (SOH) of these batteries determines their performance and longevity, and it is a crucial parameter for battery management systems. The SOH is a measure of the battery's capacity, internal resistance, and other physical and chemical characteristics, which degrade over time due to use and aging. The estimation of the SOH is challenging, especially in real driving cycles, where the battery experiences different current and temperature conditions, and its performance may vary significantly. This thesis aims to develop a methodology for estimating the SOH of lithium-ion cells in real driving cycles using a Battery Testing System (BTS)[65]. The BTS is a versatile tool that can simulate different driving scenarios and measure the battery's response in terms of voltage, current, and temperature. The data obtained from the BTS can be used to model the battery's behavior and estimate its SOH. However, this requires a thorough understanding of the battery's electrochemical and thermal properties and the development of accurate models to predict its performance. This thesis will review the existing literature on the SOH estimation of lithium-ion batteries and identify the gaps and challenges in this field. Then, it will propose a methodology based on the BTS and the electrochemical-thermal model to estimate the SOH in real driving cycles. The proposed methodology will be validated using experimental data obtained from a real electric vehicle. The results of this thesis will contribute to the development of reliable and accurate methods for the

SOH estimation of lithium-ion batteries, which can help to optimize their performance and reduce their environmental impact. The proposed methodology can also be extended to other types of batteries and applications, such as stationary energy storage systems and portable devices[65].

1.1. Motivation

Electric vehicle sales have increased significantly during the past years, setting a new high of 6.6 million in 2021. The mainstream marketing of electric vehicles is heavily dependent on the development of lithium-ion battery energy storage technology. According to estimates, between 2022 and 2030, the world’s demand for lithium-ion batteries will increase by more than five times, reaching 2000 gigawatt-hours (GWh), depicted in Figure 1[90]. To make the electric market expand, the government’s continuing assistance, as well as ongoing research into Electric Vehicle (EV) batteries, charging networks, and design to decrease the initial cost of electric cars, will be required. Increasing vehicle range, lowering charging time, and adding technology features to EV are other important reasons for quicker EV industry growth. A thorough understanding of the battery’s health progression derives from analyzing SOH characteristics such as capacity fading, impedance increase, and internal resistance fluctuations over time and number of cycles. This empirical understanding serves as the foundation for modeling, forecasting, and extrapolating degradation trajectories, allowing for the prediction of future performance deterioration, degradation rates, and eventual end-of-life scenarios for battery systems in a variety of operating environments. Accurate state-of-health SOH prediction of lithium-ion batteries is required to increase performance and ensure the safe operation of the battery energy storage system[29].

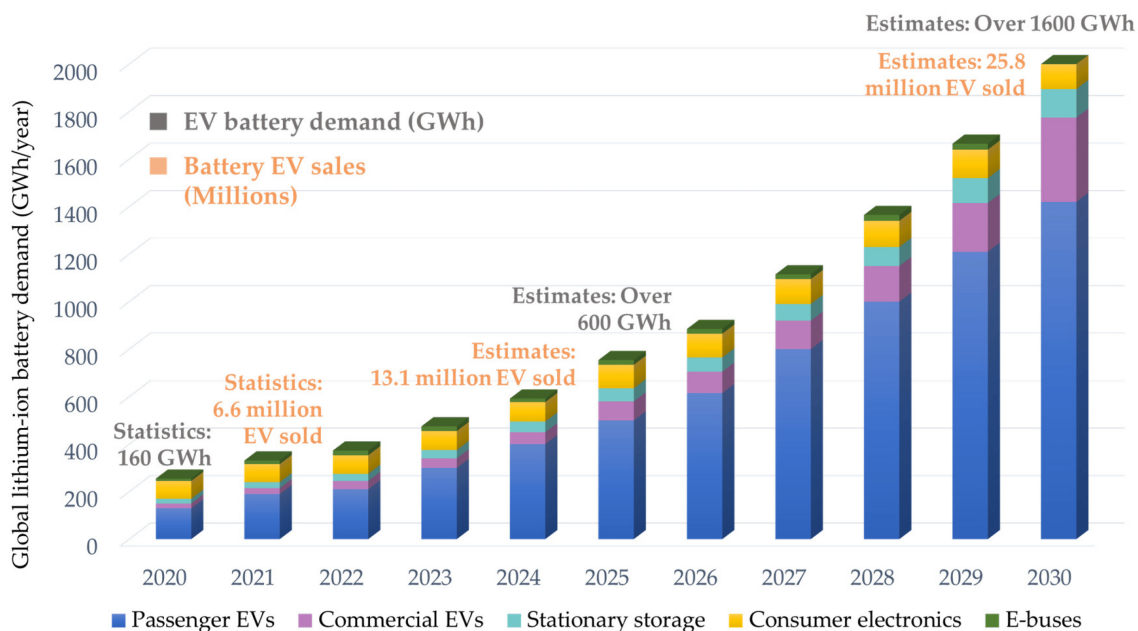


Figure 1: Global lithium-ion battery demand and sales [90]

Over time, minor flaws and problems in battery cells can develop into major failures, therefore precise prediction is essential for dependable and long-lasting performance. It is still difficult to anticipate battery system development based on time-sensitive sensor data, despite advances in our knowledge of failure processes. Due to different aging and failure processes, dynamic operating circumstances, and data-related problems including noise, incompleteness, and variability among cells and batches, this task is made even more difficult in large-scale EV applications[91].

There are still significant concerns about the safety of commercial lithium-ion batteries due to flaws and failures in real-world applications. After a lengthy period of incubation, even a slight increase in risk throughout the battery’s operating lifespan might develop into a safety hazard—a fire and explosion known as thermal runaway. Nonlinear multi-scale electro-chemical systems are difficult to model and forecast because of uncertainties in the materials and manufacturing processes, changing operating and ambient conditions, and lack of reliable dataset. Solving real-world physical problems with noisy and incomplete data and ambiguous boundary conditions complicates this problem[18].

The SOH of the Lithium Ion Battery (LIB) is often determined using the direct measurement method, the Open Circuit Voltage (OCV), coulomb counting, and Electrochemical Impedance Spectroscopy (EIS) approaches. For instance, the correlation between the OCV curves and the battery capacity is used to evaluate the SOH of LIB. However, due to the restrictions associated with the aforementioned methodologies, the direct measurement method is difficult to use for online SOH calculation. For example, in order to measure an accurate OCV curve, the LIB needs to be in equilibrium, which means that the OCV technique requires a considerable relaxation time. On the other hand, the EIS approach requires installing additional gear. A potential strategy to improve the forecast accuracy of SOH of LIB has been the use of data-driven strategies, which get beyond the drawbacks of both model-based and direct measurement approaches. Numerous kinds of Artificial Neural Network (ANN), Deep Learning (DL), Gaussian Process (GP), Support Vector Machines (SVM), and linear regression[29].

1.2. Thesis Objective

1. Gather and compile a superior dataset from battery testing that provides insight into the aging processes that occur in Lithium Ion Phosphate (LFP) cells in practical application scenarios.
2. Perform a thorough analysis of the gathered data to find trends, correlations, and important variables affecting the State Of Charge (SOC), temperature, and cycle conditions of LFP cells.
3. Investigate and contrast various machine learning methods and algorithms to create precise prediction models that estimate the SOH of LFP cells.
4. To learn more about the degradation processes and how they affect the SOH of LFP cells, look at the connections between battery performance, aging mechanisms, and important components.

5. Compare the anticipated and actual measured SOH levels from LFP cell testing to assess the precision and dependability of the created prediction models. The best prediction model architecture for LFP cell SOH prediction should be identified by analyzing the effectiveness of various machine learning methods and algorithms.
6. Discuss the viability and efficiency of data-driven methods for forecasting the SOH of LFP cells, taking into account possible real-world applications.
7. To guide the study, enhance the data analysis, and address the particular goals and difficulties related to the prediction of the LFP cell SOH, review the relevant scientific literature and research findings.

1.3. Thesis Contribution

Assessment of the best model architecture's accuracy for SOH prediction in LIB, yielding important insights into the efficacy of data-driven methodologies. The development of battery health estimation methods and their possible real-world applications across a range of sectors would benefit from this conclusion. Evaluation of the effects of various charging and discharging profiles on LIB SOH, allowing for the development of ideal charging plans that reduce battery deterioration and increase usable life. This result will assist optimize LIB operating conditions and have consequences for battery management. Examining the connection between LIB SOH and environmental variables including humidity and temperature. The results of this research will make it easier to create efficient environmental control and thermal management systems that will reduce battery aging and enhance overall battery performance. Proof of the feasibility of using data-driven models to forecast LIB SOH, hence minimizing reliance on expensive and time-consuming physical testing methods. This result will aid in the creation of practical and affordable battery health monitoring methods.

Creating prediction models especially for LFP cells using the data set and modeling strategies established. This will enhance the prediction models' precision and suitability for this particular battery chemistry and yield important insights into the aging processes of LFP cells. Using actual LFP cell data, prediction models are validated to determine their accuracy and dependability in forecasting the SOH of LFP cells. This result will showcase the prediction models' potential for real-world use in battery management systems and help to improve LFP cell characterisation and maintenance techniques. Comparing and assessing various machine learning models or algorithms for LIB SOH prediction. This result will shed light on the advantages and disadvantages of different modeling strategies and assist in determining the best methods for precise and trustworthy SOH estimate.

Working together with industry partners to confirm that the prediction models built in real-world settings are feasible and scalable. This result will demonstrate the research's business applicability and make it easier for data-driven methods for SOH estimate in commercial battery systems to be adopted. Addition of derived knowledge in battery research and formulation of suggestions for further study in the area of battery health assessment. This result will aid in directing and stimulating more study and developments in the comprehension of and management of LIB aging.

1.4. Thesis Outline

- **Chapter 1** This chapter sets the stage for the thesis by outlining the motivations behind the research, the objectives of the thesis, the contributions it seeks to make to the field, and an overview of the thesis structure.
- **Chapter 2** This chapter delves into the technical aspects of lithium-ion batteries, including their composition, characteristics, the types found in electric vehicles EV, different cell types, and battery management systems. It also introduces the concepts of Recurrent Neural Networks RNN and Long Short-Term Memory LSTM networks.
- **Chapter 3** A comprehensive review of existing models for battery aging, methods for predicting battery health, and machine learning approaches for battery state prediction, highlighting the various algorithms and their comparative performance.
- **Chapter 4** This chapter describes the experimental approach, including the test methodology, description of the test bench, test configuration, and detailed information on feature engineering and machine learning models used for State of Health SOH prediction.
- **Chapter 5** Analysis and visualization of the data collected from cycle aging tests are presented, followed by the implementation details and performance evaluations of LSTM, GRU, and Regression models.
- **Chapter 6** The final chapter summarizes the findings of the thesis, discusses the implications of the research, and proposes directions for future studies to expand on the work presented.

2. THEORETICAL BACKGROUND

The theoretical framework, outlined in Section 2, is developed through a sequence of specific subjects, each contributing to a thorough comprehension of lithium-ion batteries (Li-ion) and their practical uses. Section 2.1 commences the investigation by conducting a thorough analysis of Li-ion batteries, including their composition and fundamental attributes. The following part, 2.2, focuses on the widespread use of Li-ion batteries in Electric Vehicles EV, specifically differentiating between cobalt-containing and cobalt-free variations, and exploring the various cell types—cylindrical, prismatic, and pouch cells.

In Section 2.4, the focus of the discussion turns to Battery Cell Parameters and Specifications. This section explains the measurements that determine battery performance and introduces techniques for predicting Remaining Useful Life (RUL). Section 2.5 of the document introduces Background Research, which provides context by emphasizing the increasing need for batteries in Electric Vehicles. This helps to place the theoretical debate within the wider context of battery consumption.

The investigation concludes in Section 2.6, where the pivotal significance of Battery Management System (BMS) is examined. Section 2.6.1 elaborates on the diverse capabilities of BMS, which encompass the supervision of charging and discharging operations, ongoing monitoring of cell conditions, and the assurance of overall safety. This theoretical framework provides a thorough understanding of Li-ion batteries, covering its fundamental composition, several applications in Electric Vehicles, and the crucial role of Battery Management Systems.

Finally, Section 2.7 provides a comprehensive introduction to advanced neural network models, namely Recurrent Neural Network (RNN) and Long Short Term Memory (LSTM). The section concludes by implementing LSTM into recurrent neural network structures after thoroughly examining the structures and functionality of these models. This theoretical framework provides a thorough comprehension of lithium-ion batteries and serves as a foundation for future research. The objective of this research is to utilize these insights to investigate battery degradation and optimize performance.

2.1. Li-Ion Battery

Batteries are an essential means of electricity storage. The Li-ion battery, a rechargeable battery, was invented in 1912 by American physicist Gilbert Lewis. These batteries have a high power density and energy density, are ecologically benign, have a long life, and are widely utilised in electronic devices. However, for high-power applications such as energy storage systems and electric vehicles, a large number of batteries are linked in parallel to form a battery pack. This causes problems with stability, coherence, cost, and safety. Because of these issues, the use of Li-ion batteries is restricted. Because of the impacts on charge rate, voltage range, and temperature, the operating system of Li-ion batteries should be kept within the boundaries. If these limits are exceeded, battery performance may suffer and security issues will arise. To ensure the safe functioning of the Li-ion battery, the capacity of the battery and its overall service life must be evaluated[75].

2.1.1. Lithium-ion Battery Composition

Lithium-ion cells consist of two electrodes. The positive (cathode) is cast on the aluminium current collector and the negative (anode) on a copper current collector. The two electrodes are separated by a thin electron-insulating layer (called separator) soaked with the electrolyte (liquid or solid ionically conducting media). During the charge (i.e., upon application of an anodic current) lithium ions are extracted from the lithium-containing cathode, shuttled through the separator/electrolyte, and intercalated into the anode host structure. Simultaneously, the electrons are flown through the external electric circuit in the same direction. This process spontaneously progresses in the reversed direction on discharge, thus providing the electrical power as shown in Figure 2[75].

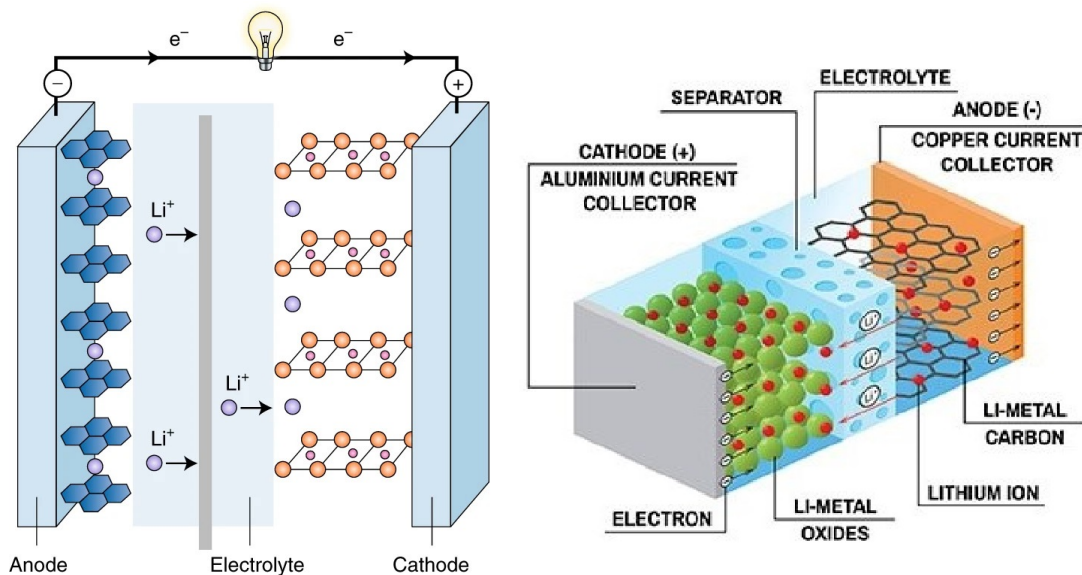


Figure 2: Lithium-ion battery Basic Structure and Internal Detailed Structure [75]

2.1.2. Characteristics of Lithium-ion Battery

A battery is composed of five crucial elements: the anode, cathode, current collectors, electrolyte, and separator. The anode and cathode, acting as the positive and negative terminals, are made from materials specifically designed for electrochemical processes. The current collectors, which round them, serve to give structural reinforcement and enable the effective transmission of electrical charges. The electrolyte, usually in a liquid or gel state, functions as a conductive medium, allowing the migration of ions between the anode and cathode, thereby facilitating the passage of electric current. The separator serves as a physical barrier that prevents direct contact between the anode and cathode. Its purpose is to maintain the battery's structural integrity and prevent the occurrence of short circuits.

A lithium-ion battery consists of several essential components that form its fundamental structure:

Anode: The anode, usually composed of graphite, serves as the electrode responsible for collecting lithium ions during the charging process. Discharging the battery results in the release of lithium ions from the anode.

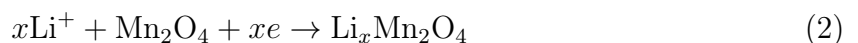
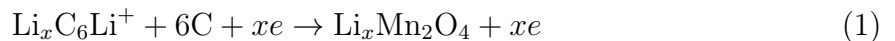
Cathode: The cathode is an electrode that accepts lithium ions during the discharge process. It is typically made of lithium transition metal oxides, such as lithium cobalt oxide, lithium manganese oxide, or lithium iron phosphate.

Separator: The separator is a porous substance that is placed between the anode and cathode. Its purpose is to facilitate the movement of lithium ions while also preventing direct contact between the electrodes, which could lead to a short circuit.

Electrolyte: The electrolyte is a conductive solution that consists of lithium salts, such as lithium hexafluorophosphate (LiPF_6), dissolved in an organic solvent. It facilitates the transfer of lithium ions between the anode and the cathode during the process of charging and discharging.

During the charging process of a lithium-ion battery, lithium ions move from the cathode to the anode via the electrolyte, where they are then stored in the graphite anode material. The process is followed by the movement of electrons across an external circuit, enabling the battery to store electrical energy.

Electrons, which are liberated during the oxidation process, traverse the external circuit in accordance with Equation (1), where as Li^+ ions navigate within the cell. Upon the electrons becoming available, these ions undergo a transformation into lithium (Li), which is then retained within the electrodes, as illustrated in Equation (2)[3].



During the process of discharge, lithium ions migrate from the anode to the cathode through the electrolyte, while simultaneously releasing electrons into the external circuit. These electrons can then be harnessed to provide power to various devices. The flow of ions and electrons in the battery is responsible for completing the electrochemical processes and generating the electric current.

Lithium-ion batteries possess numerous benefits compared to alternative rechargeable battery options. These advantages include a notably high energy density, a lightweight construction, and a minimal self-discharge rate. Nevertheless, they also possess certain drawbacks, including susceptibility to elevated temperatures, the potential for thermal runaway if mishandled, and the likelihood of capacity deterioration with time.

Proper charging and discharging processes, temperature management, and safety measures are essential for maximising the performance and lifespan of LIB. Continual research and development endeavours concentrate on enhancing its energy density, safety characteristics, and cost efficiency to satisfy the growing need for energy storage in diverse industries.

2.2. Most Common Lithium-Ion Batteries found in EVs

Lithium Nickel Manganese Cobalt Oxide (NMC), possess the second highest energy density after Lithium Nickel Cobalt Aluminium Oxide (NCA). They are being used in a range of electric powertrains, electric tools, and recreational vehicles, including scooters and ebikes. Following the advent of widespread electric vehicle production, the majority of research and development efforts have been concentrated on NMC [2].

NCA, or lithium nickel cobalt aluminium oxide, is currently one of the most widely used chemistries in lithium-ion batteries. Tesla was the first to adopt NCA in the cylindrical 18650 cell format, while other companies primarily used prismatic or pouch shaped cells. Subsequently, Tesla discovered a method to enhance the cells capacity by incorporating a narrower separator and enlarging the batteries container from 18650 to a 21700 cell in their US Model 3 [37].

Nevertheless, Tesla successfully launched its Tesla Model 3 equipped with a LFP battery in China in 2020, and they expressed contentment with the results. Recently, it was announced that all of their Model 3 and Model Y vehicles will be universally equipped with this particular type of batteries[78].

2.2.1. Li-ion batteries that do not contain cobalt

Lithium iron phosphate (LFP)

In 1996, JB Goodenough developed Lithium Iron Phosphate as a replacement for lithium cobalt oxide, which had an unstable structure when subjected to overcharging. Lithium Iron Phosphate serves as a cathode substitute. LFP has gained significant popularity since then, mostly because of its exceptional thermal stability, extended cycle life, and remarkable tolerance. When the LFP is maintained at high voltage for extended periods, it can withstand full charge conditions and experiences less strain compared to other lithium-ion cells, making LFP unique. However, the conductivity of this material is relatively lower when compared to lithium metal oxides[23].

Two effective options for enhancing structural conductivity and surface conductivity are metal doping and conducting material coating on LFP electrodes. These methods significantly increase the achievable capacity while maintaining a reasonable charge and discharge current density. Furthermore, nanoscale components have the ability to significantly reduce transportation distances. So far, numerous battery manufacturers have effectively promoted carbon-coated nano-LFP materials. Until recently, a significant proportion of electric vehicles, including models such as BYD, Chevrolet Spark, and BMW Active Hybrid 3 and 5 series, have been utilising the LFP technology. The primary application of LFP technology by the Build Your Dreams (BYD) Group is in their electric vehicles, including electric fork-lifts, buses, and cars, to provide power for their drive-trains. LFP reduces its nominal voltage by 3.3 volts per cell compared to cobalt-based cells. Lower temperatures adversely affect the performance of most batteries, while storing them in hot temperatures reduces their lifespan[16].

This applies to LFP cells as well. LFP batteries have a higher self-discharge rate compared to other types of lithium-ion batteries, which might result in balance issues. Utilising superior cells and advanced electronic control systems can mitigate this issue, while concurrently leads to increased manufacturing expenses. Despite the implementation of these solutions, the LFP technology still faces a constraint in terms of energy density, which is roughly 120 Wh kg⁻¹, in comparison to the maximum value of around 170 Wh kg⁻¹. The energy density of this figure falls well below the minimum requirement of (>250 Wh kg⁻¹) for the upcoming generation of electric vehicles by the year 2025. Furthermore, LFP technology is particularly suitable for demanding tasks like trucks and buses, as opposed to mobile electronic devices with lesser volumetric density. Due to this defect, the OEMs are likely to favor NMC, NCA, and the combinations of NMC/ NCA and Lithium Manganese Oxide (LMO)/ NMC because of their higher specific energy. Due to its affordability and extended lifespan, LFP offers significant potential for market penetration in the power supply industry, in addition to its existing applications in the automotive sector[16].

Li-Manganese Oxide (LMO)

In 1983, M. Thackeray and his colleagues initially introduced the spinel LMO. Moli Energy introduced the lithium-ion battery cathodes to the market in 1996. LMO exhibits a three-dimensional diffusion of lithium ions Li⁺, in addition to a more robust spinel structure. Moreover, LMO is more cost-effective and non-toxic compared to Lithium Cobalt Oxide (LCO). Li⁺ ions have the ability to occupy [Mn204] polyhedral frameworks, which enhances their rating capabilities compared to materials with two-dimensional frameworks for Li⁺ diffusion[16].

The low capacity of LMO, which is 148 mah g⁻¹, is a significant drawback. In addition, the electrolyte's instability results in poor performance, causing degradation of Mn (uneven conversion of Mn³⁺ ions to Mn²⁺ and Mn⁴⁺) and reduced capacity. LMO exhibits moderate safety characteristics and comparatively low specific energy when compared to the other primary lithium-ion compounds[11].

Despite its drawbacks, LMO possesses the advantage of being the most cost-effective among numerous electrode materials. This attribute holds significant importance in facilitating the introduction of electric vehicles into the market. Most LMO batteries are combined with NMC to enhance the energy density and extend the lifespan of electric car usage. This mix provides optimal performance for each system, and the majority of electric vehicles, including the Nissan Leaf, Chevy Volt, BMW I3, and others, utilize LMO/NMC blends. The LMO component, typically accounting for thirty percent of the LMO/NMC combo, significantly enhances speed at a high current level, while the remaining NMC portion contributes to a longer range. In the foreseeable future, the demand for LMO will mostly be influenced by the combination of LMO / NMC in electric vehicles. This particular combination of cathodes has the potential to replace LFP (Lithium Iron Phosphate) in electric automobiles and electric buses in China[16].

Lithium Titanate (LTO)

Batteries using lithium titanate negative electrodes have been in existence since the 1800s. The negative electrode of a Li-titanate cell contains graphite instead of graphite, and this material adopts a spinel structure. An alternative electrode option could be lithium manganese cobalt oxide. Spinel lithium titanate is highly regarded as an advantageous electrode material because it maintains its size during lithiation, resulting in an impressive electrode lifespan. Additionally, it offers enhanced safety due to a very stable discharge plateau at around 1.55 V vs Li / Li⁺. The inadequate performance at high power levels can be attributed to the low conductivity of this material and the weak Li⁺ diffusion coefficient.

However, this issue can be resolved by reducing the lengths of the li-ion transport paths by the use of appropriate nanostructures. Additionally, enhancing the electronic conductivity can be achieved by implementing techniques like as doping, surface coating, and including composites with superior electronic conductors, such as carbon materials. Due to their exceptional safety standards, these batteries are utilised in portable medical applications[48].

2.2.2. Li-ion batteries that contain Cobalt

Li Cobalt Oxide (LCO)

LCO is a widely utilised lithium-ion compound. It is represented by the following chemical symbols: Lithium cobalt oxide (LiCoO₂) Li-cobalt is frequently chosen for applications such as personal computers, mobile phones, and digital cameras because of its elevated specific energy. The battery consists of a cathode made of cobalt oxide and an anode made of graphite carbon. During discharge, the movement of lithium ions from the anode to the cathode is facilitated by the layered structure of the cathode. During the process of charging, the direction of the flow of electric current is reversed.

Like other lithium-ion batteries with cobalt, these batteries also include a graphite anode, which results in a limited cycle life due to changes in the electrolyte interface. To enhance durability, load capacity, and affordability, modern applications use nickel, manganese, and/or aluminium. Due to the exorbitant price of cobalt and the enhanced efficiency achieved by combining it with other active cathode elements, Li-cobalt is experiencing a decline in popularity, primarily in favour of Li-manganese, but also towards Lithium Nickel Manganese Cobalt and Li-Nickel Cobalt Aluminium Oxide. The drawbacks of Li-cobalt are its limited lifespan and inadequate thermal stability[8].

Li Nickel Cobalt Aluminium Oxide (NCA)

Cobalt aluminium oxide has been available for specialised uses since 1999. The battery possesses both high energy density (the amount of energy stored per unit volume) and high power density (the rate at which energy is supplied from the battery). Additionally, it has a long lifespan. These attributes are also present in the NMC battery, so rendering the two batteries highly comparable. Although NCA batteries, like other lithium batteries, are not highly safe, certain safety management procedures must be used in electric vehicles[48].

Additionally, the higher production cost of this product limits its applicability for various uses. Tesla is now recognised as the sole electric vehicle manufacturer that utilises NCA chemicals in its battery production. Tesla asserts that its NCA batteries contain a lower amount of cobalt compared to NMC811 batteries, which consist of 80% nickel, 10% manganese, and 10% cobalt[48].

Li Nickel Manganese Cobalt (NMC)

A highly successful lithium-ion system is achieved by combining a cathode composed of nickel, manganese, and cobalt NMC. Like Li-manganese, these systems can be customised to function as Energy Cells or Power Cells. As an illustration, an 18650 cell including NMC chemistry has a capacity of approximately 2,800 mAh and can provide a current of 4A to 5A under moderate load conditions. However, if the NMC chemistry in the same cell is optimised for specific power, the capacity decreases to around 2,000 mAh, but it can produce a continuous discharge current of 20 A. A silicon-based anode can achieve a capacity of 4,000 mAh or more, but this comes at the cost of reduced loading capability and a shorter cycle life. The addition of silicon to graphite presents the disadvantage of causing the anode to expand and contract during the process of charging and discharging, hence resulting in mechanical instability of the cell[8].

The key to NMC success rests in the amalgamation of nickel and manganese. Analogously, the combination of sodium and chloride, which are individually harmful, transforms into table salt, a substance used as a flavour and food preservative. Nickel is renowned for its high specific energy but lacks stability. On the other hand, manganese possesses the advantage of creating a spinel structure, which results in reduced internal resistance, but it provides a lower specific energy. By combining the metals, their individual strengths are mutually enhanced[8].

NMC is the preferred battery technology for power tools, e-bikes, and other electric power-trains. The cathode composition consists of equal parts of nickel, manganese, and cobalt, sometimes referred to as 1-1-1. Cobalt is both costly and scarce. Battery makers are decreasing the amount of cobalt used but with slight differences in performance. An effective composition is Nickel Cobalt Manganese (NCM)532, consisting of 5 parts nickel, 3 parts cobalt, and 2 parts manganese. Additional combinations include NMC622 and NMC811. Cobalt acts as a stabiliser for nickel, which is a highly energetic and reactive substance[8].

2.3. Different Types of Li-ion Cells

Apart from the chemistry of the cell, its shape and size also have significant effects on its capacity and performance. There are mainly 3 types of cells available on the market: cylindrical, prismatic, and pouch cell. Here, a brief introduction to these 3 types of cells is given.

2.3.1. Cylindrical Cells

To fabricate lithium-ion cylindrical cells, anode foil, separator, and cathode foil are rolled into long segments and inserted into a rigid cell housing or "can" made of aluminium or stainless steel. Following the positioning of safety discs into the can containing the liquid electrolyte, the electrodes are attached to the outer battery[64].

terminals, which are the top and bottom of the cell in this instance. To hermetically enclose the cell, the upper disc assembly is crimped shut. The cylindrical cell design is economical, long-lasting, and capable of cycling, but it is cumbersome and has a low packaging density as a result of the presence of space cavities[55].

2.3.2. Prismatic Cells

Prismatic cells share the same structure as cylindrical cells, however, they employ a flat rectangular casing to reduce the overall thickness of the cell. The electrode/separator assembly can be either cylindrical, where it is rolled, or rectangular, where it is stacked like a deck of cards. The battery terminals can be positioned as contact pads either on the top or on the side of the housing. The prismatic cell's thin and compact design is highly suitable for integration into consumer devices, especially in cases where convenient battery replacement is a need[64].

The use of aluminum or steel housing ensures structural stability, durability, and resistance to dampness. The cells' slender rectangular form offers excellent spatial efficiency, adaptability, and thermal resilience. Nevertheless, these cells exhibit several drawbacks, including inadequate temperature control in comparison to pouch cells, reduced lifespan, and elevated manufacturing expenses[55].

2.3.3. Pouch cells

Pouch cells are similar to prismatic cells, with a slender rectangular shape. Their structure consists of rectangular arrangements of separate layers of electrodes and separators. However, instead of a stiff metal casing, they employ a flexible polymer/aluminum "bag" that is laminated. The electrodes are equipped with tabs on one side, which are fused with battery terminal tabs protruding from the top of the bag. The assembly is fully coated with a liquid electrolyte, and the bag is hermetically sealed using heat[64].

Pouch cells offer cost, weight, and thickness savings by removing the inflexible housing. However, the flexible pouch is susceptible to swelling, which can lead to issues with durability, reduced capacity, and compromised safety. This particular cell has a greater energy density in comparison to other designs due to its reduced weight[55].

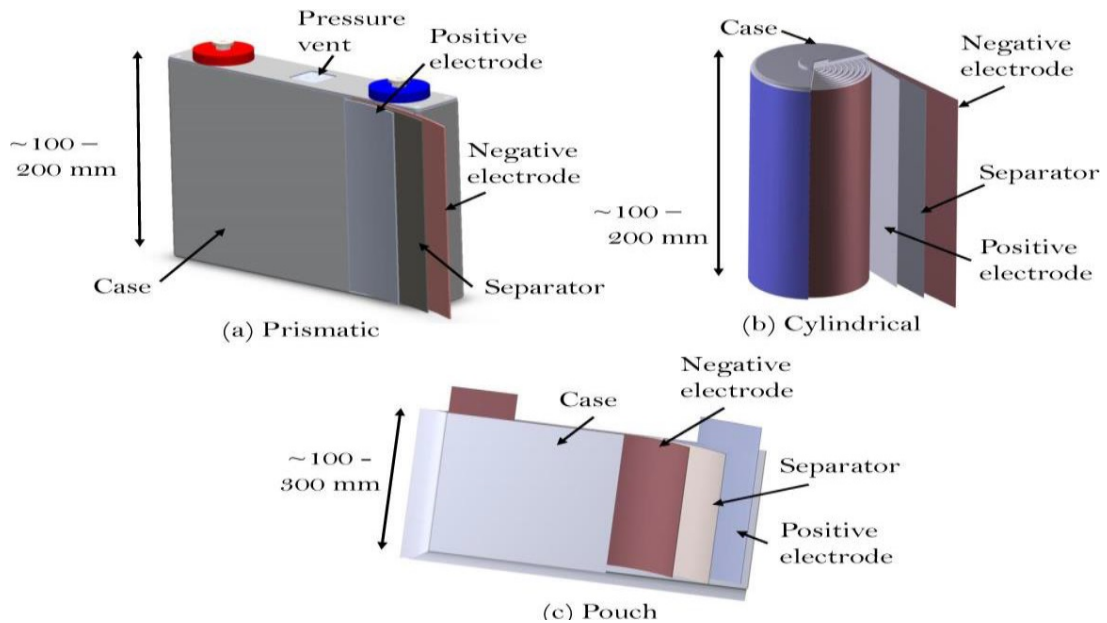


Figure 3: Prismatic Cells- Cylindrical Cells- Pouch Cell [42]

2.3.4. Characterization of a LiFePO₄ battery

This work describes and analyses experimental results gained from a thorough evaluation of a Lithium Iron Phosphate LFP-based Lithium-ion Battery LIB. The experimental tests done specifically examine the electrical properties of the battery while excluding the research of its temperature features. In addition to the standard experiments conducted on lithium-ion batteries, such as rate capability, energy capability, and impedance behavior, the focus here is on investigating the unique attributes displayed by LFP cells. Explicit protocols have been implemented to carefully examine and understand the reported behaviors.

The analysis primarily focuses on battery degradation, aiming to comprehend the effects of examined events on battery performance and how these aspects are modified by the aging process. The study surpasses simple observation by utilizing modeling tools to illustrate and authenticate the occurrences witnessed during experimentation. The research seeks to enhance comprehension of the complex interaction of elements that result in battery degradation using these techniques.

The primary goal of this research is to understand the complex behavior of batteries, providing a detailed understanding of their strengths and weaknesses. These insights are crucial in optimizing battery performance and extending their operating lifespan. Moreover, the obtained knowledge aids in the creation of advanced diagnostic algorithms,

facilitating precise evaluations of a battery's condition. Consequently, this enables the development of efficient battery designs that are based on a comprehensive comprehension of their operational dynamics. This study adopts a comprehensive strategy that establishes the foundation for progress in battery technology, offering useful insights for both practical implementation and theoretical frameworks.

2.4. Battery Cell Parameters and Specifications

This section will examine different criteria and aspects that impact battery technology, specifically focusing on factors associated with battery aging.

SOC and SOH are crucial factors in battery management, representing the present charge level and the overall health condition of the battery, respectively.

$$SOC = \frac{C_c}{C_f} \times 100\% \quad (3)$$

where:

- SOC is the state of charge of the battery
- C_c is the current capacity of the battery
- C_f is the full capacity of the battery

$$SOH = \frac{C_f}{C_n} \times 100\% \quad (4)$$

where:

- SOC is the state of charge of the battery
- C_f is the full capacity of the battery
- C_n is the normal capacity of the battery

The variable "Ccurr" denotes the present battery capacity, "Cfull" denotes the battery capacity when it is completely charged, and "Cnom" is the capacity of a new battery[40].

Two methods can be used to quantify the capacity of a battery: the state of charge SOC and the state of health SOH[40]. Equation 3 defines SOC as the ratio of the current capacity of the battery to its maximum capacity when completely charged. As demonstrated in Equation 3, State of Health SOH, on the other hand, represents the current maximum capacity of the battery in relation to its original capacity. The battery has reached its End Of Life (EOL).

when it reaches 80% of its maximum capacity. The remaining useful life RUL of a battery refers to the number of charge/discharge cycles it can undergo before reaching its end of life EOL. Battery Management Systems BMS can precisely ascertain the State of Charge SOC of Lithium-ion batteries with an accuracy ranging from 0.6% to 6.5%. Nevertheless,

their existing capabilities do not extend to reliably forecasting the State of Health SOH and Remaining Useful Life RUL of batteries[13][79].

Capacity or Nominal Capacity (Ah for a specific C-rate): - The total amp-hours are available when the battery is discharged from 100% SOH to the cut-off voltage with a specified discharge current, expressed as a C-rate. A device's capacity is calculated by multiplying the discharge current in amps by the discharge time (in hours), and it decreases as the C-rate increases[51].

C- and E- rates: - The discharge current is commonly expressed as a C-rate in battery specifications to standardise it with respect to battery capacity, which can vary significantly across different batteries. A C-rate is a quantitative measure of the rate at which a battery is discharged relative to its maximum capacity. A discharge rate of 1C signifies that the current will deplete the battery entirely within one hour. This amounts to a discharge current of 100 amps for a battery with a capacity of 100 amp-hours. This battery would have a 5C rate of 500 amps and a C/2 rate of 50 amps. Similarly, the amount of electricity released is denoted by an E-rate. The discharge power necessary to fully deplete a battery within one hour is referred to as a 1E rate[49].

Charge Current (Recommended): - The optimal current at which the battery is charged under a continuous charging strategy is around 70% SOC after which it transitions to a constant voltage charging method[49].

Cycle Life (a number for a specific DOD): - The maximum number of discharge-charge cycles that a battery may experience before falling short of specified performance requirements. cycle life is estimated for a given charge and discharge circumstances. The actual operating life of the battery is dependent on temperature, humidity, cycle depth, and other factors. As the Depth Of Discharge (DOD) rises, the cycle of life falls[27].

State of Power (SOP): - State Of Power (SOP), is a metric that signifies the present state of charge or discharge of a battery at a particular instant. Typically, it is quantified as a percentage relative to the overall battery capacity. Standard Operating Procedure SOP is used to ascertain the residual charge or the amount of energy now accessible in the battery[73].

Self-Discharge Rate: - The self-discharge rate relates to the speed at which a battery depletes its charge while it is not being utilized. Batteries exhibiting low self-discharge rates possess the ability to maintain their charge for extended durations, making them ideal for applications resulting in extended shelf life[63].

Efficiency: - Efficiency quantifies the capacity of a battery to convert stored chemical energy into electrical energy and inversely[38]. It denotes the proportion of output energy to input energy and is typically stated as a percentage. Greater efficiency implies less energy loss throughout the process of charging and discharging[19].

Internal Resistance: - The resistance within the battery, which normally changes for charging and discharging, is also affected by the state of charge of the battery. The battery effectiveness reduces as internal resistance rises, and thermal stability is adversely affected because more of the charging energy is transferred to heat[49].

State of Charge (SOC, %): - The current battery capacity is expressed as a percentage of the highest capacity. To estimate the change in battery capacity over time, SOC is usually determined using current integration[40].

Voltage, Cut-off: - It is defined as having the lowest voltage that can be used. The "empty" state of the battery is commonly defined by this voltage[49].

Voltage, Nominal: - It is defined as the battery's reference or describing voltage, sometimes referred to as the "normal" voltage of the battery[49].

Voltage, Charge: - It is the voltage at which a battery is fully charged. Charging systems often include constant current charging until the battery voltage reaches the charge voltage, followed by constant voltage charging, which allows the charge current to reduce until it is extremely low[49].

Voltage, Terminal: - With a load applied, the voltage connecting the battery terminals. SOC and discharge/charge current influence terminal voltage[49].

Energy Density (Wh/L): - The "normal" energy density of a battery per unit volume is often known as the volumetric energy density. The specific energy of a battery is determined by its chemistry and manufacturing. It determines the battery size necessary to accomplish a specific electric range, together with the vehicle's energy utilisation[88].

Power Density (W/L): - The highest amount of electricity available per unit of volume. The specific power of a battery is determined by its chemistry, composition, and production procedure. It calculates the battery size needed to meet a specific performance goal[88].

Beginning of Life (BOL): - which denotes the initial condition of a battery when it is initially utilised. It indicates the beginning of the battery's operating lifespan. Typically, the state of health SOH and state of charge SOC of the battery are regarded as 100% at this point[72].

End of Life (EOL): - EOL, refers to the stage when a battery is deemed to have exhausted its usable lifespan or is unable to achieve the necessary performance standards. At the end of its lifespan, the battery experiences a considerable decrease in capacity, resulting in a loss of intended performance[72].

Remaining Useful Life (RUL): - The Remaining Useful Life RUL of a battery is the projected duration during which the battery is anticipated to function adequately before its performance decreases to an unsatisfactory level. Accurately predicting the Remaining Useful Life RUL is of utmost importance to enhance battery management systems, particularly in domains like electric vehicles, renewable energy storage, and portable devices.

Multiple methodologies and models exist for forecasting the Remaining Useful Life RUL of batteries, typically involving the analysis of the battery's progressive deterioration during its lifespan. An established method involves the use of mathematical equations to represent the decline in capacity or other forms of degradation. Below is a concise derivation of a universal Remaining Useful Life RUL equation:

is the rate of change of capacity with respect to time. Practically, the prediction of Remaining Useful Life RUL is a more intricate task that necessitates the consideration of multiple parameters. Several models integrate various characteristics, including temperature, charge/discharge cycles, and operational conditions. Machine learning methodologies, such as regression models and neural networks, are frequently employed to forecast the Remaining Useful Life RUL by leveraging past data.

Battery producers and researchers frequently design specialized models that are customized to the unique properties of the battery chemistry and intended use. The models are trained using empirical data to enhance precision in forecasting the Remaining Useful Life RUL.

It is crucial to acknowledge that forecasting the Remaining Useful Life RUL is intrinsically subject to uncertainty, and outcomes can be affected by fluctuations in operational circumstances and environmental elements. Continuing research endeavours to improve the precision and dependability of Remaining Useful Life RUL forecasts for various battery types and applications[24].

$$RUL(t) = \frac{C_t}{\frac{dC}{dt}} \quad (5)$$

where:

- $RUL(t)$ is the Remaining Useful Life at time(t)
- C_t is the battery capacity at time(t)
- $\frac{dC}{dt}$ is the rate of change of capacity with respect to time.

2.4.1. Methods for Forecasting Remaining Useful Life

Indeed! Researchers have devised two primary methodologies, namely model-based and data-driven methods, to forecast the RUL of lithium-ion batteries. Model-based procedures involve creating electrochemical models, which are made up of specific mathematical models that accurately show how chemical processes work inside the body. This facilitates an exact depiction of internal variables within the model, resulting in a high level of accuracy in forecasting[83]. Nevertheless, these models are exceedingly intricate and necessitate an extensive prior understanding of the batteries under investigation[57].

Indeed, the disassembly of the battery may be necessary for parameters, which poses challenges in incorporating these technologies into practical applications. Predictive maintenance is an anticipatory maintenance approach that seeks to forecast and avert equipment malfunctions through data analysis. An essential goal of predictive maintenance is to accurately estimate the remaining useful life RUL of equipment. There are three prominent approaches to accomplishing this: similarity, survival, and degradation models. The selection of each model is determined by the available data, and similarity models prove to be especially valuable when complete histories from similar machines are accessible[46].

Their functioning involves training a model by utilising degradation profiles of comparable machines to predict the Remaining Useful Life RUL of the specific equipment. This prediction is based on the similarity between the degradation profiles of the equipment in question and those of similar machines. Conversely, data-driven methodologies do not prioritise the examination of the battery's intrinsic electrochemical reactions or failure reasons. Alternatively, they regard the internal functioning of a battery as an unknown entity and utilise past data to construct models that can directly comprehend the decline in battery performance based on the monitoring data. Subsequently, these models undergo further refinement and optimisation to attain elevated levels of precision[36]. While data-driven methods may not attain the same degree of accuracy as model-based methods, they are more straightforward to deploy in practical scenarios.

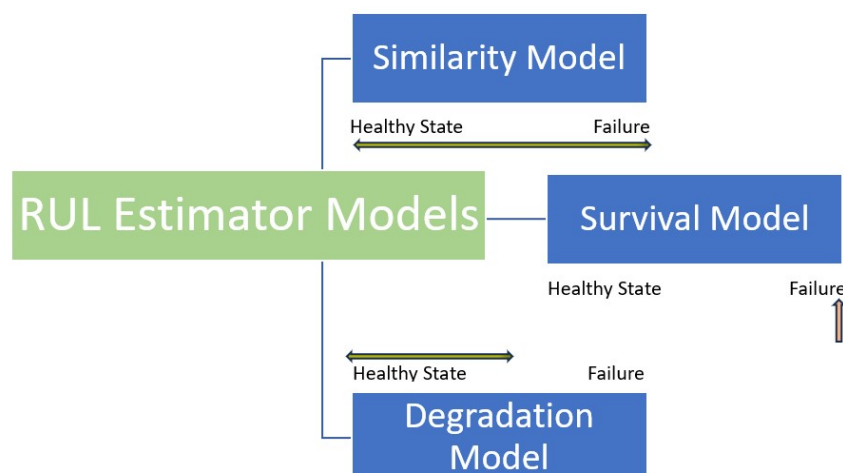


Figure 4: RUL Estimator Models [46]

2.5. Background Research

2.5.1. Battery demand for Electric Vehicles continues to rise

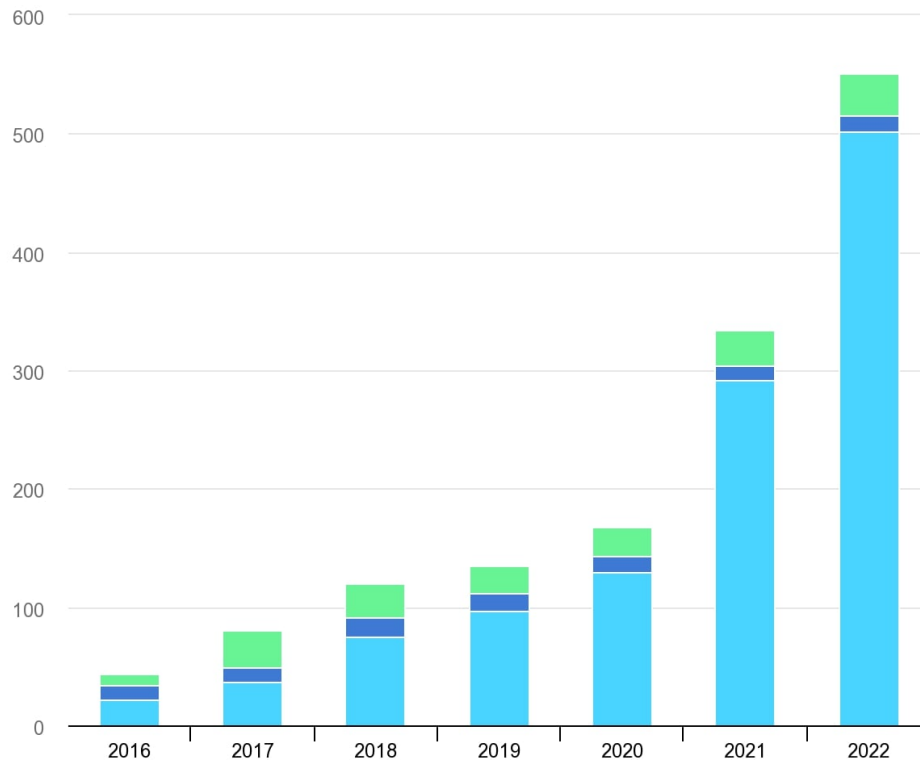


Figure 5: Battery demand by mode, 2016-2022 [22]

The demand for lithium-ion (Li-ion) batteries in the automotive industry grew by around 65% to reach 550 GWh in 2022, compared to 330 GWh in 2021. This increase was mainly driven by the rise in sales of electric passenger cars Light Duty Vehicle (LDV) Light blue in the graph, with new registrations experiencing a 55% spike in 2022 compared to the previous year[22].

In China, the demand for batteries used in vehicles had a growth of over 70% in 2022 compared to 2021. Similarly, the sales of electric cars increased by 80% during the same year. However, the growth in battery demand was significantly reduced due to the rising popularity of Plug-in Hybrid Electric Vehicle (PHEV). The United States had about 80% surge in battery demand for vehicles, even while electric car sales only experienced a modest growth of around 55% in 2022. In 2022, the average battery size for battery electric cars in the United States increased by approximately 7%. However, the average battery size for battery electric cars in the US is still about 40% larger than the global average. This is partly because there is a higher proportion of Sport Utility Vehicle (SUV) in the US electric car market compared to other major markets, and also because manufacturers are focusing on providing longer all-electric driving ranges. The sales of Battery Electric Vehicle (BEV) and Plug-in Hybrid Electric Vehicles PHEV are surpassing the sales of

Electric Vehicle (HEV) worldwide. This trend is driven by the fact that BEV and PHEV have bigger battery capacities, which in turn leads to a higher need for batteries[22].

2.6. Battery Management Systems (BMS)

The Battery Management System BMS is a crucial component of an Electric Vehicle EV as it prevents the batteries from being overcharged or depleted beyond safe limits. Engaging in such behavior can lead to battery degradation, a rise in temperature, diminished longevity, and potentially jeopardize the safety of individuals. The BMS can enhance the vehicle's range by optimizing energy utilization[1]. Figure 6 depicts the conventional block diagram of the Battery Management System BMS.

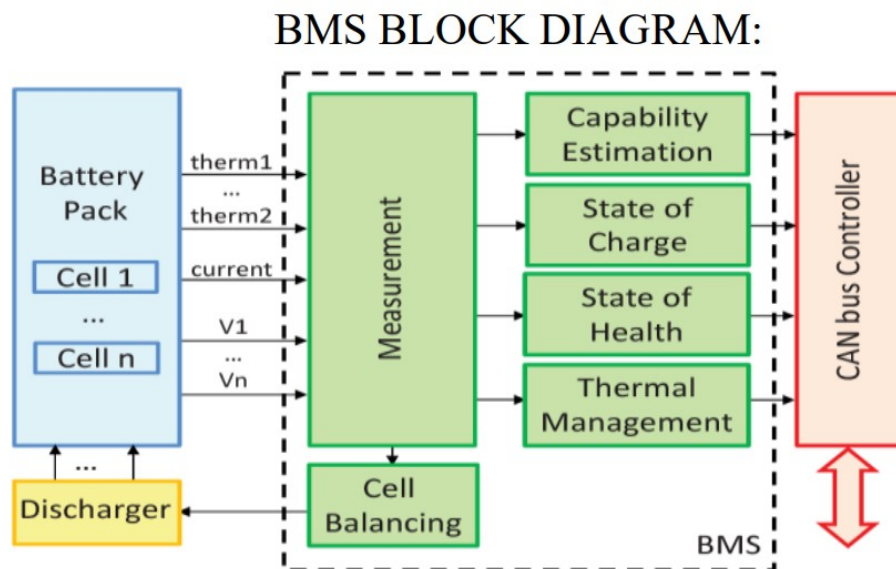


Figure 6: BMS block diagram [1]

The battery management system is essential for the following reasons

1. Maintain the safety and the reliability of the battery
2. Battery state monitoring and evaluation
3. To control the state of charge
4. For balancing cells and controlling the operating temperature
5. Management of regenerative energy

2.6.1. BMS Functionalities

There is a wide variety of Battery Management Systems BMS available in the market. You have the option to either create your own BMS or purchase an Integrated Circuit (IC) that is readily accessible. From a hardware point of view, there are three different kinds of Battery Management Systems BMS based on their topology: Centralised BMS, Distributed BMS, and Modular BMS. Nevertheless, the functionality of these Battery Management Systems BMS is the same. Below is a depiction of a generic Battery Management System[58].

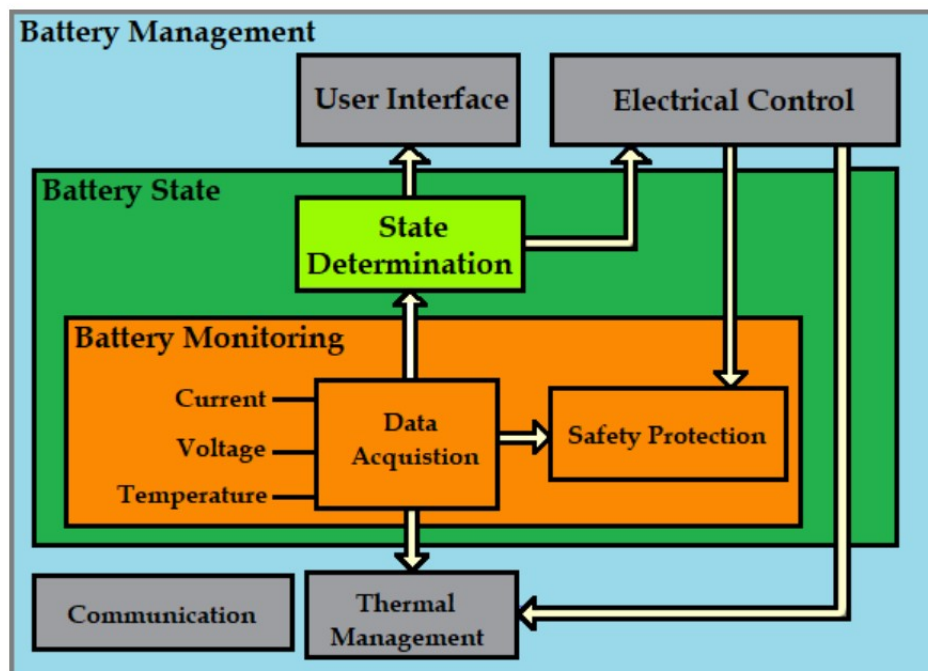


Figure 7: BMS Functionalities [58]

Battery numbers are calculated and standard measurements are taken for the voltages of the individual cells, as well as the current, voltage, and temperature of the battery pack. The BMS utilizes these readings to compute the essential operational parameters of the cells and battery packs, including the SOC, SOH, and DOD. Adaptable These techniques also help to prolong battery life and meet the requirements of the basic power network[58]. An energy management system equipped with a user interface can be utilized to regulate and observe the efficiency of battery systems in various system blocks, hence enhancing the longevity of the batteries. Effective management of the charging and discharging process greatly influences the longevity of a battery. Battery Management Systems BMS offer economic advantages including increased battery longevity, enhanced precision, and reduced costs[69].

2.7. An Overview of Recurrent Neural Network (RNN) and Long Short-Term Memory (LSTM)

The human brain-like algorithms of neural networks detect patterns. A machine perceives sensory data and labels or clusters it. They can recognize vectors of numerical patterns, which must be converted into real-world data like images, sounds, texts, and time series. Artificial neural networks use many densely coupled neurons to solve a problem. A huge number of parallel processors in tiers make up an ANN. The first tier gets raw input, like human optic nerves. In the same way, neurons far from the optic nerve receive signals from those closer to it, each tier receives output from the previous one. The final tier generates system output[52][17].

2.7.1. What is Recurrent Neural Network (RNN)?

A Recurrent Neural Network RNN is an extension of a feedforward neural network that includes an internal memory component. An RNN is characterized by its recurrent nature, meaning that it carries out the same operation for each input of data. However, the output of the current input is influenced by the previous computation. Once the output is generated, it is duplicated and returned to the recurrent network[50]. To make a decision, it takes into account both the current input and the output it has acquired from prior inputs.

RNN, in contrast to feedforward neural networks, can utilize their internal state (memory) for processing sequences of inputs. These capabilities make them suitable for jobs such as unsegmented, continuous handwriting recognition, or speech recognition. In alternative neural networks, each input is considered to be mutually independent. However, in a Recurrent Neural Network RNN, all of the inputs are interconnected[17].

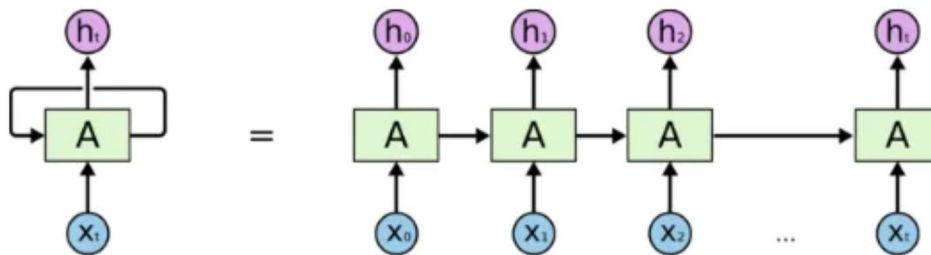


Figure 8: An unrolled recurrent neural network. [52]

Initially, the algorithm extracts the value of $X(0)$ from the input sequence and then generates the output $h(0)$, which, in conjunction with $X(1)$, serves as the input for the subsequent iteration. The values $h(0)$ and $X(1)$ serve as the input for the subsequent phase. Likewise, $h(1)$ obtained from the previous phase serves as the input for $X(2)$ in the subsequent step, and this pattern continues. By doing so, it maintains the ability to retain the context during the training process[50][52]. The formula for the current state and after applying the activation function is shown in equations 6 and 7.

$$h_t = f(h_{t-1}, x_t) \quad (6)$$

$$h_t = \tanh(W_{hh}h_{t-1} + W_{xh}x_t) \quad (7)$$

W is weight, h is the single hidden vector, Whh is the weight at the previous hidden state, Whx is the weight at the current input state, tanh is the Activation function, that implements a Non-linearity that squashes the activations to the range[-1.1]. The output is shown in equation 8 where Yt is the output state. Why is the weight at the output state[50].

$$Y_t = W_{hy}.ht \quad (8)$$

Advantages of Recurrent Neural Network

- Recurrent Neural Networks RNN have the ability to represent a series of data in such a way that each individual data point is considered to be influenced by previous ones.
- In order to expand the effective pixel neighbourhood, convolutional layers are also incorporated into recurrent neural networks.

Disadvantages of Recurrent Neural Network

- Gradient vanishing and exploding problems.
- Training an RNN is a very difficult task.
- It cannot process very long sequences if using tanh or relu as an activation function.

2.7.2. What is (LSTM) and How it Works

It is even possible to employ recurrent neural networks in conjunction with convolutional layers to expand the successful pixel LSTM networks, which stand for long-term short-term memory and are a modified version of recurrent neural networks. These networks make it simpler to remember information from the past. In this case, the vanishing gradient issue that RNN was experiencing is fixed. When given time lags of uncertain duration, LSTM is an excellent choice for classifying, processing, and making predictions about time series. The model is trained through the utilization of backpropagation. There are three gates present in an LSTM network to be found[52][17].

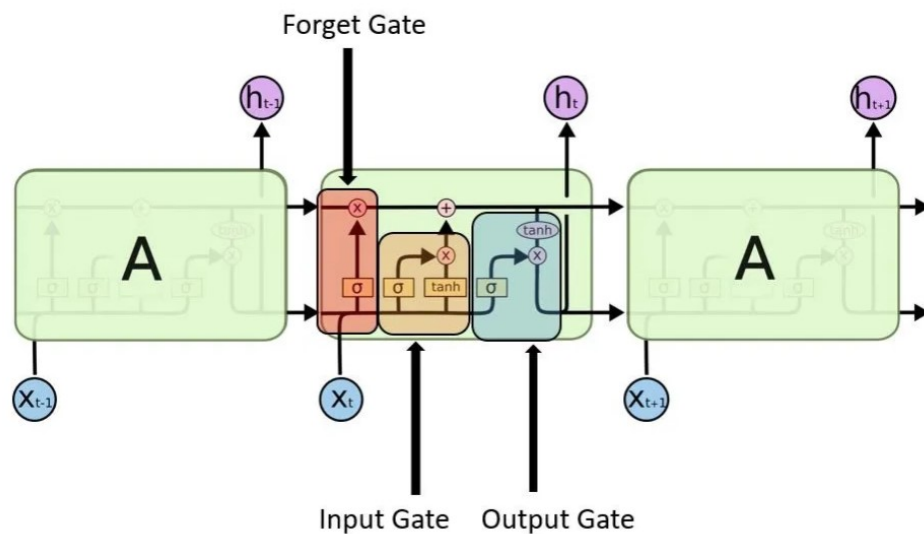


Figure 9: LSTM Gates. [52]

Utilizing a cell state, which is depicted by a horizontal line in the LSTM diagram, is the fundamental concept that underpins long short-term memory LSTM systems. The state of the cell functions like a conveyor belt, enabling information to move along it without undergoing substantial changes. The integration of gates, which regulate the flow of information into and out of the cell state, is the most important characteristic of long short-term memory LSTM devices[50]. An input gate, a forget gate, and an output gate are the three primary components that makeup LSTM cells. This is a high-level overview of the structure. A corresponding weight matrix and bias term are connected to each gate in the system. The LSTM is able to keep or discard information from prior time steps thanks to these gates, which govern the transmission of information and allow it to do so[17].

The key equations involved in an LSTM cell are the following:

1. Forget gate:

Determine which specific details should be eliminated from the block. The sigmoid function determines it. The process examines the prior state (h_{t-1}) and the input content (x_t) and generates a numerical output ranging from 0 (excluded) to 1 (included) for each number in the cell state C_{t-1} [17][50]. The equation for the forget gate is as follows:

$$f_t = \text{sigmoid}(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (9)$$

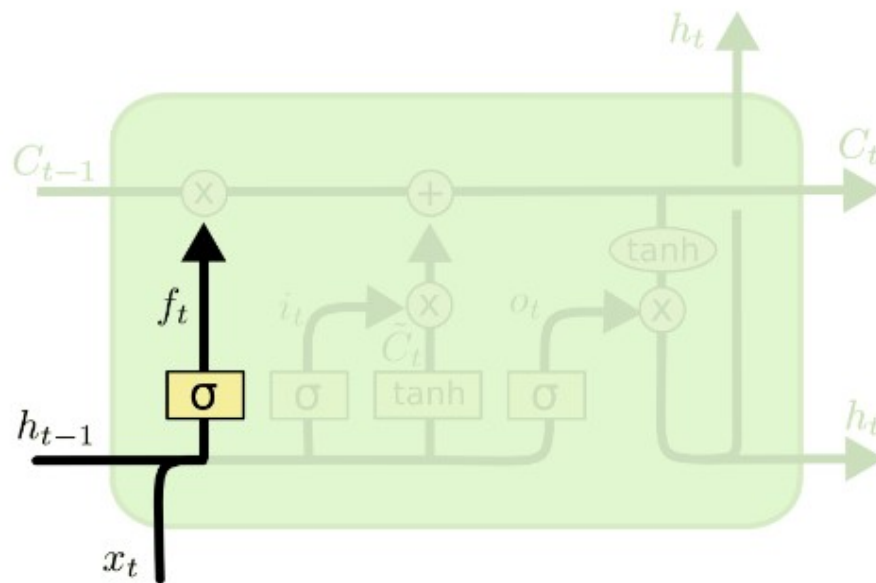


Figure 10: LSTM forget gate. [52]

2. Input gate: Determine the specific value from the input that should be utilized to alter the memory. The sigmoid function determines which values to allow between 0 and 1. The tanh function assigns weights to the input numbers, determining their level of significance on a scale from -1 to 1[50]. The equation for the input gate is as follows:

$$i_t = \text{sigmoid}(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (10)$$

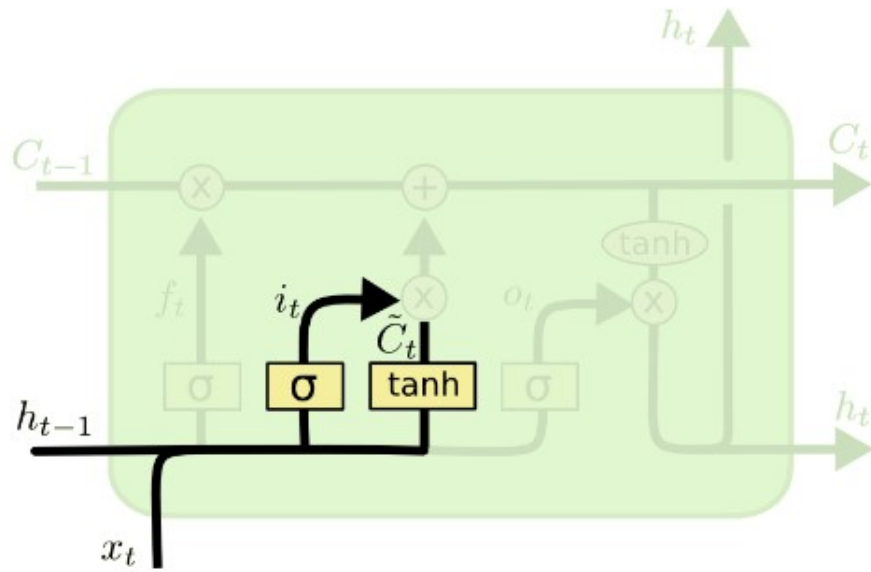


Figure 11: LSTM input gate. [52]

3. Output gate: The output is determined based on the input and the memory of the block. The sigmoid function determines which values to allow between 0 and 1. The tanh function assigns a weight to the input values, determining their level of significance on a scale from -1 to 1. This weight is then multiplied with the result of the sigmoid function[17][52][50]. The equation for the output gate is as follows:

$$o_t = \text{sigmoid}(W_o \cdot [h_{t-1}, x_t] + b_o) \quad (11)$$

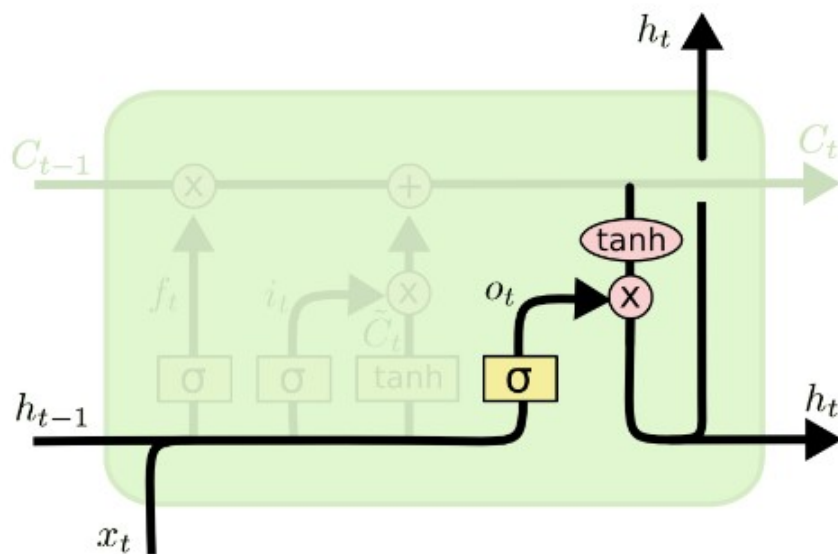


Figure 12: LSTM output gates. [52]

2.7.3. Implementation of a recurrent neural network using Long Short-Term Memory (LSTM)

Machine Learning (ML) models are innovative tools that are transforming the field of data analysis and decision-making in different areas[5]. Their importance stems from their capacity to tackle intricate issues and their adaptable use in vital industries such as healthcare, banking, image identification, and natural language processing. These models facilitate the process of computers acquiring knowledge from extensive information, revealing concealed patterns, and generating precise forecasts or choices. Choosing the right machine learning model that is specifically suited to the situation at hand is crucial, highlighting the significance of a careful and deliberate decision-making process in selecting a model[45].

Machine learning plays a crucial role in addressing complex challenges characterized by extensive amounts of data and sophisticated patterns. The ability of this technology to enable computers to learn from data and uncover hidden insights is especially useful. Within this particular context, the section explores the Long Short-Term Memory LSTM model, which is a specific sort of Recurrent Neural Networks RNN. This exploration highlights the importance of LSTM in capturing long-term relationships in sequential data, providing a detailed knowledge of its ability to handle temporal features. Moreover, this section delves into the practical implementations of LSTM in many fields, providing insights into its versatility and efficacy in real-life situations[6].

The section offers a concise introduction to how to put LSTM models into practice, moving from theoretical talks to practical applications. This feasible observation helps to connect the divide between theoretical comprehension and practical implementation, providing useful perspectives on the utilization of LSTM in data analysis and decision-making procedures. The threefold investigation in this section together presents a complete depiction of the transformation potential of ML models, highlighting their ability to adjust, their versatility, and their crucial function in tackling the complexities of intricate problems across various fields[52].

3. LITERATURE REVIEW

This chapter explores the ageing models utilised in electrochemical cells, focusing specifically on the most recent advancements in machine learning models designed for prognostication. To acquire a thorough comprehension of the topic at hand, an exhaustive and methodical review of the literature was conducted. The principal objective of this study was to investigate the applicability of recent developments in ageing models to the prediction of LIB degradation. Through a comprehensive search of these databases, our objective was to identify the most recent and pioneering research in the respective field.

In the context of electrochemical cells, the importance of ageing models cannot be overstated. Batteries experience a progressive deterioration in performance as a result of enduring numerous charge and discharge cycles; this degradation affects their capacity, efficiency, and overall lifespan. Comprehending and precisely forecasting the ageing process holds significant significance across diverse sectors, including renewable energy systems, portable electronics, and electric vehicles, wherein the dependable and extended functionality of batteries is essential. Battery ageing prediction has witnessed the emergence of machine learning models as potent instruments owing to their capacity to manage intricate data patterns and nonlinear associations. These models use previously collected information during training on battery performance to identify the fundamental ageing patterns and generate dependable forecasts regarding subsequent deterioration. The incorporation of machine learning methodologies has created novel opportunities for the precise evaluation of battery conditions and the enhancement of battery management tactics.

The objective of this literature review is to examine the latest and most auspicious advancements in ageing models that utilise machine learning. Different approaches, such as support vector machines, random forests, neural networks, and recurrent neural networks, which have shown a lot of promise in predicting how batteries will break down, will be looked at. Furthermore, we shall deliberate on the obstacles and prospects associated with the implementation of these models in practical situations, taking into account variables including the accessibility of data, the interpretability of the models, and the computational intricacy. Through an in-depth review and integration of relevant literature, the objective of this chapter is to impart significant perspectives on the present condition of battery ageing prediction. The insights derived from this exhaustive examination will form the foundation for the forthcoming sections of this thesis, in which we shall put forth and assess our model for predicting ageing, which is constructed using machine learning methodologies.

3.1. Battery Aging Modelling

Battery ageing is one of the most important factors that must be taken into consideration for a successful Li-ion battery application. However, the performance of lithium-ion batteries will progressively deteriorate owing to the effects of working conditions and usage duration, which restrict the battery's use to a certain degree. Ageing processes have an impact on battery capacity and resistance, which are the key components of a battery

electrical model, as well as battery longevity, which is directly related to the system's profitability[9].

In addition, the overall operation mechanism of lithium-ion batteries is rather intricate, including electrochemical reactions, energy transfer, charge transfer, heat dissipation, and various other processes. Furthermore, internal characteristics and external factors actively interact when the battery system is working. Hence, constructing a collaborative simulation test platform for electro-thermal and ageing is highly significant to enhance the investigation of the ageing properties of lithium-ion batteries[66].

Generally, battery models fall into the following primary classifications:

- Physics Based Models (PBM)
- Empirical/Semi-Empirical Models
- Equipment Circuit Models (ECM)
- Electrochemical Models (EM)
- Data Driven Models (DDM)
- Hybrid Models (HM)

3.1.1. Physics-Based Models (PBMs)

Physics-based models can be categorised into two distinct categories. Certain of them are considered "phenomenological" because ageing sources, such as an aside reaction that results in the formation of a Solid Electrolyte Interphase (SEI), appear explicitly in the set of governing equations and interact directly with the other model characteristics. The second group compiles mathematical models that are nearly identical to those for a pristine cell, with the addition of isolated empirical relations or curve-fitting procedures to update model parameters such as SEI film resistance or thickness, volume fraction of active material, and so forth. Phenomenological physics-based models represent the most advanced and intricate approach due to the necessity for a comprehensive comprehension of the mechanisms of ageing. After undergoing validation, the model becomes capable of analysing an extensive range of operating conditions and control strategies. Moreover, battery manufacturers may employ this information to enhance the design of their batteries. To date, the majority of physics-based models have incorporated a single ageing source. This is primarily due to the exceedingly complex nature of real-world systems and the limited theoretical foundation that exists in simulating certain ageing sources, such as the structural degradation of active materials[62].

3.1.2. Empirical/Semi-Empirical Models

Empirical models rely solely on data and are constructed using experimental observations. They create correlations between battery performance and degradation characteristics without taking into account the underlying physical and chemical mechanisms[14]. These models are generally uncomplicated to create and execute, rendering them valuable when intricate data regarding the battery's internal functioning is missing or hard to get. Semi-empirical models integrate empirical and physics-based elements. They could integrate certain physical concepts or information regarding the battery's behaviour while still depending on experimental data for certain components of the model. Semi-empirical models provide a compromise between accuracy and complexity and are frequently employed when a comprehensive physics-based depiction is not required, but some scientific understandings are still sought[14].

Empirical and semi-empirical models are highly valuable for making rapid estimations, doing preliminary research, and implementing applications that prioritise simplicity and demand less computational power. Additionally, they may be used for real-time monitoring and regulation of battery systems without necessitating an extensive understanding of the battery's underlying mechanisms[76]. These models frequently include regression analysis or curve fitting techniques to establish correlations between input factors (such as current, voltage, and temperature) and battery performance parameters (such as capacity decrease impedance increase, and cycle life). The efficacy and accuracy of empirical models are significantly influenced by the quality and amount of the dataset employed for training and validation[14].

3.1.3. Equivalent Circuit Models (ECMs)

The equivalent circuit model ECM is commonly employed in various battery applications due to its rapid execution time, simplicity, and relatively high level of accuracy. Nevertheless, the ECM suffers from the drawback of limited model extrapolation when the battery is pushed to its operational limits, resulting in reduced performance. Consequently, this battery model is not frequently employed in applications that need high current rates or operate at extremely low temperatures. Accuracy and complexity are crucial considerations when simulating the dynamic behaviour of a battery. The ECM has the capacity to efficiently fulfil these factors. The model comprises three primary components: one component that represents the thermodynamic characteristics of the battery chemistry, including the open-circuit voltage OCV as a function of the state of charge SOC; another component that represents the kinetic behaviour of the cell's internal impedance; and a source or load that completes the circuit for the charging or discharging processes[74].

ECM parameters commonly consist of internal ohmic resistance, succeeded by one or more Resistor Capacitor (RC) pairs. The most basic model just considers the internal ohmic resistance, failing to adequately depict the battery's dynamics during operation. Therefore, the Thevenin Equivalent Circuit Model ECM, which includes an extra RC pair to be combined with the internal ohmic resistance, is commonly employed due to its favourable trade-off between precision and simplicity. The estimation of ECM parameters

is commonly performed through a Hybrid Pulse Power Characterisation (HPPC) test conducted at various state of charge SOC levels. Temperature and state of charge SOC can influence these parameters. For example, a rise in temperature can result in an elevation of Open Circuit Voltage OCV in specific Lithium-Ion Battery LIB compositions, while a decrease in state of charge SOC can lead to a reduction in charge transfer resistance, ultimately impacting the ECM parameters[74].

3.1.4. Electrochemical Models (EMs)

Electrochemical models are a type of physics-based model that especially aim to represent electrochemical reactions and transport processes occurring in batteries [76]. These models utilise mathematical equations based on battery physics, namely Butler-Volmer kinetics, to accurately depict the electrochemical events occurring at the electrode interfaces [60]. Electrochemical methods offer a more comprehensive knowledge of battery behaviour in comparison to Equivalent Circuit Models. This is due to the ability of EM to provide deep insights into electrode kinetics, concentration gradients, and reaction overpotentials. The use of these models enables researchers and engineers to investigate the impact of various operational parameters, electrode compositions, and temperature on the efficacy of battery performance [4]. While EM provide more precision, they can be more computationally intensive compared to ECM. The computational complexity escalates when incorporating intricate reaction kinetics, multiscale modelling, and spatial dispersion of species [76]. Electron microscopes EM are highly valuable for conducting detailed investigations into specific degradation mechanisms and analysing the properties of materials [60].

3.1.5. Data-Driven Models (DDMs)

Data-driven models depend on empirical data and statistical analysis [47]. They avoid the fundamental principles of battery physics and prefer using machine learning algorithms, such as Neural Network (NN) and other Artificial Intelligence (AI) methods, to establish connections between input data (such as operating conditions, temperature, and state of charge) and battery performance or degradation measurements [25].

DDM are useful in situations when there is a lack of comprehensive physical knowledge or where the underlying degradation mechanisms are complicated to explain analytically [77]. By being exposed to extensive quantities of experimental data, these models can acquire the ability to comprehend intricate patterns and correlations that might be a challenge for traditional physics-based models [25][77].

DDM are frequently used in state-of-health SOH estimation, where they anticipate the battery's health or remaining useful life using real-time readings [81]. Furthermore, DDM can enhance battery management techniques and control algorithms by taking into account actual operating conditions and variability [41].

3.1.6. Hybrid Models (HMs)

Hybrid models HM are highly beneficial in battery modelling because they may effectively combine several modelling techniques, providing a flexible way to accurately represent intricate electrochemical behaviours [80].

The process of amalgamation frequently entails the integration of physics-based models, which explore the complex electrochemical processes at the electrode level, with corresponding circuit models that offer a broader perspective of the complete battery system. Hybrid models achieve a fine equilibrium between computational accuracy and processing efficiency by combining both approaches [82].

Hybrid models have the advantage of adaptability, enabling researchers and engineers to customize the level of intricacy to meet unique needs. This flexibility becomes especially advantageous when examining extensive battery packs or intricate systems, where the computational expense of constructing intricate physics-based models for each cell may become excessively burdensome. Hybrid models are particularly effective in situations like this, as they are able to provide accurate forecasts while also reducing the computational load [81]. Hybrid models are highly important in enhancing our comprehension of battery behavior and maximising their efficiency in many applications, such as electric automobiles and grid-scale energy storage systems.

3.2. Approaches to Predicting the State of Health

A lithium-ion battery's state of health SOH is usually found by comparing the actual values to the initial values of performance parameters that have been calculated directly or indirectly during the battery's operational life. These parameters include internal resistance and battery capacity[32].

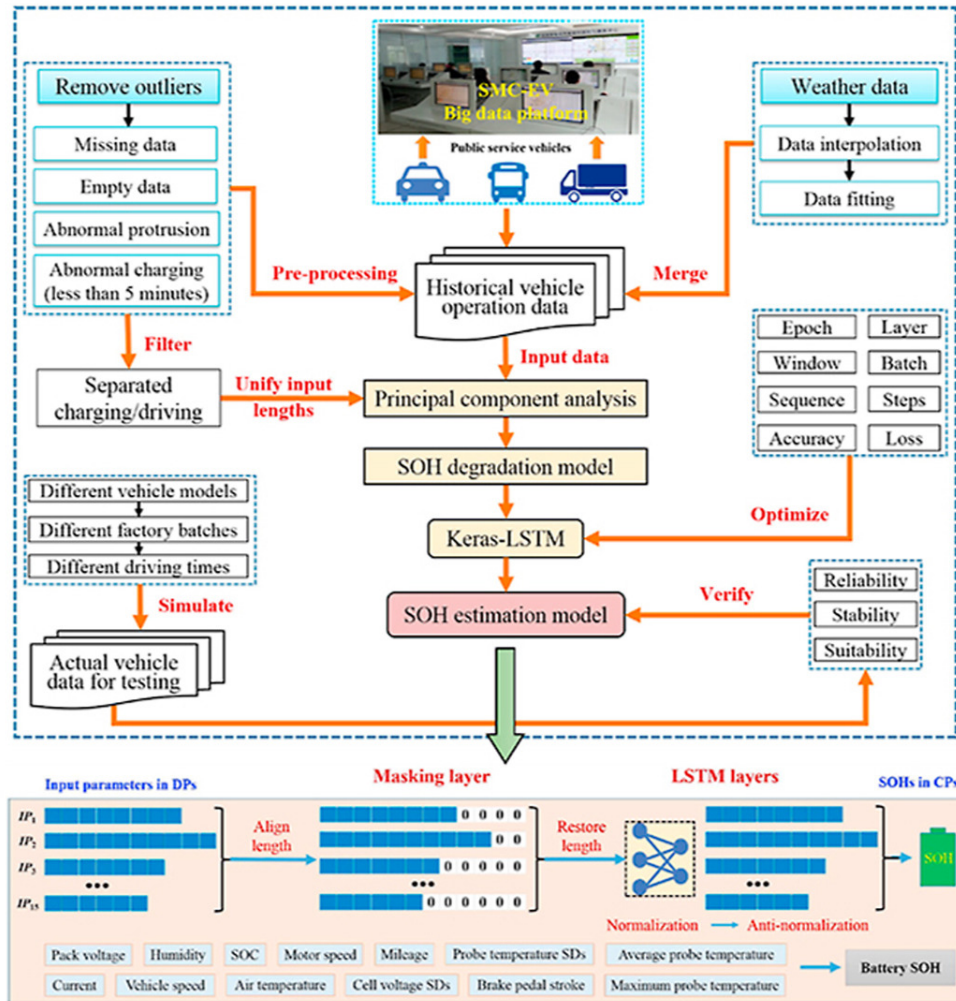


Figure 13: SOH ESTIMATION METHOD [32]

An extensive body of literature has described the various SOH classifications and their attributes. As an illustration, classify the estimation methods for battery SOH into experimental methods and model-based estimation methods. The primary branches and prevalent methodologies are illustrated in Figure 8. Among these, data-driven methods, electrochemical models, indirect analysis methods, and empirical models have been the most developed in recent years [86][87].

By employing data fitting to summarise the correlations between different parameters (e.g., temperature, cycle number, charging/discharging current, etc.) and battery SOH from experimental data, an empirical model for estimating battery SOH can be derived. An approach called "Battery SOH estimation Method Combined Online Model-Based Capacity Estimation and Routine Calibration" was introduced by Han [26].

According to Sebastian, the hold stage temperature, time, and state of charge SOC influence battery life. These three parameters are governed by the renowned Arrhenius rule. In contrast to alternative battery models, empirical models possess more succinct model

structures and greater adaptability in practical implementations. Nevertheless, they continue to rely on experimental data collection methods and the configuration of working conditions significantly affects estimation outcomes[53].

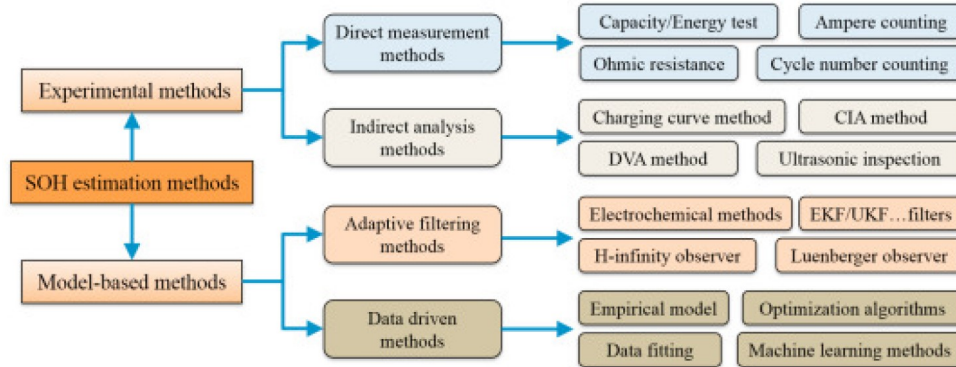


Figure 14: Sorting battery SOH prediction techniques [32]

A variety of methodologies have been employed to evaluate the degree of degradation in lithium-ion batteries. Filter-based approaches, including the equivalent circuit model ECM and electrochemical model EM, have been widely implemented. Approaches such as the equivalent circuit model ECM, Single Particle Model (SPM), and Pseudo 2D Model (P2D) are frequently implemented[84]. Data-driven models DDM that utilize machine learning have emerged as a potentially effective approach for making accurate and timely predictions of battery states, bringing forth a novel viewpoint. The prevalent categories and methodologies are depicted in Figure 14. In recent times, significant progress has been made in the areas of data-driven strategies, electrochemical models, indirect analysis techniques, and empirical models[32].

Several studies have examined the analysis of indirect tests and the electrochemical performance of lithium-ion batteries to get insight into the dynamics of battery capacity variations. To estimate the capacity of onboard batteries approaches such as Capacity Incremental Analysis (CIA) and Differential Voltage Analysis (DVA) have been devised[32]. Moreover, the commonly used method of ampere-hour (Ah) integration requires fully draining the battery after a complete charge using a constant current-constant voltage (CC-CV) protocol. This enables the estimation of the drained capacity by integrating the Ah data. Together, these methods enhance our comprehension of battery capacity patterns and enable precise capacity assessments[85][70].

3.3. Machine Learning for Predicting Battery State

One intriguing approach that has the potential to improve battery management and performance assessment is the incorporation of machine learning techniques into the estimation of battery states, particularly the state of health SOH. These techniques effectively handle large amounts of battery data by utilizing the computing power of algorithms, providing valuable insights for accurate state forecasts.

3.3.1. Overview of Machine Learning Approaches

supervised learning, unsupervised learning, and reinforcement learning are the three main categories of machine learning approaches used in battery state prediction[56]. To enable algorithms to relate input properties to corresponding desired output properties, such as battery states like SOH, supervised learning involves training models using labeled data. Unsupervised learning explores the structure and patterns in unlabeled data, frequently leading to the discovery of new information. Reinforcement learning focuses on teaching models to make sequential decisions based on interactions with an environment, although it is not generally used in battery state prediction[56] [54].

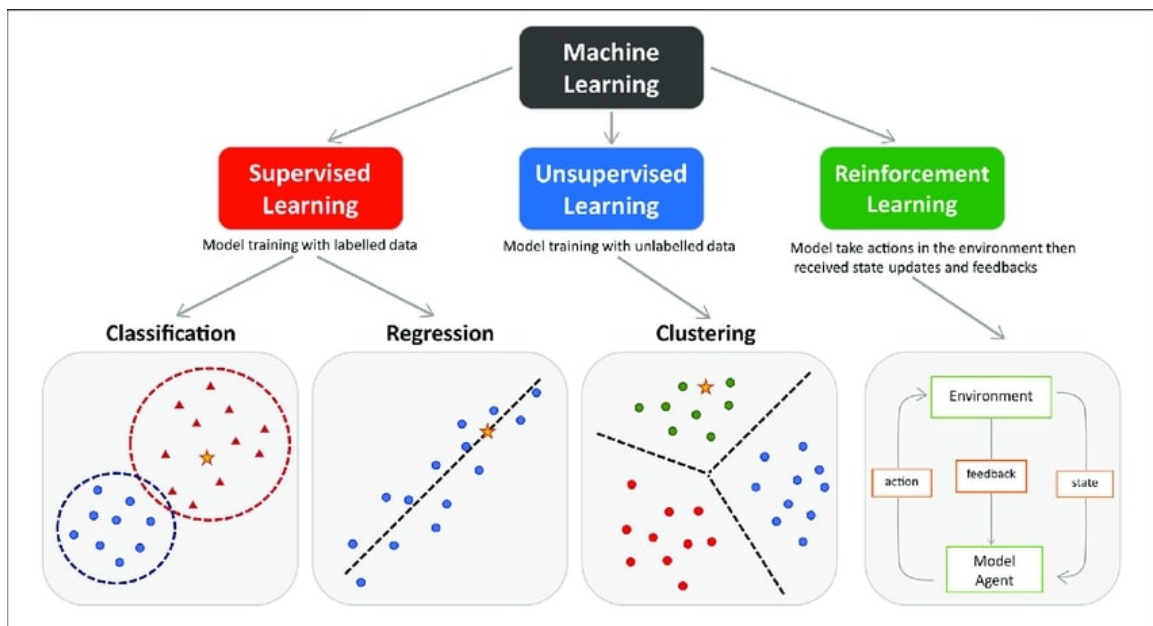


Figure 15: Overview of machine learning approaches [54]

3.3.2. Machine Learning Algorithms for Battery State Prediction

Numerous machine learning methods provide advantageous tools in searching for precise and effective battery health prediction. These algorithms handle various elements of battery behavior and accommodate different data profiles; they range from feedforward neural networks to probabilistic approaches [68].

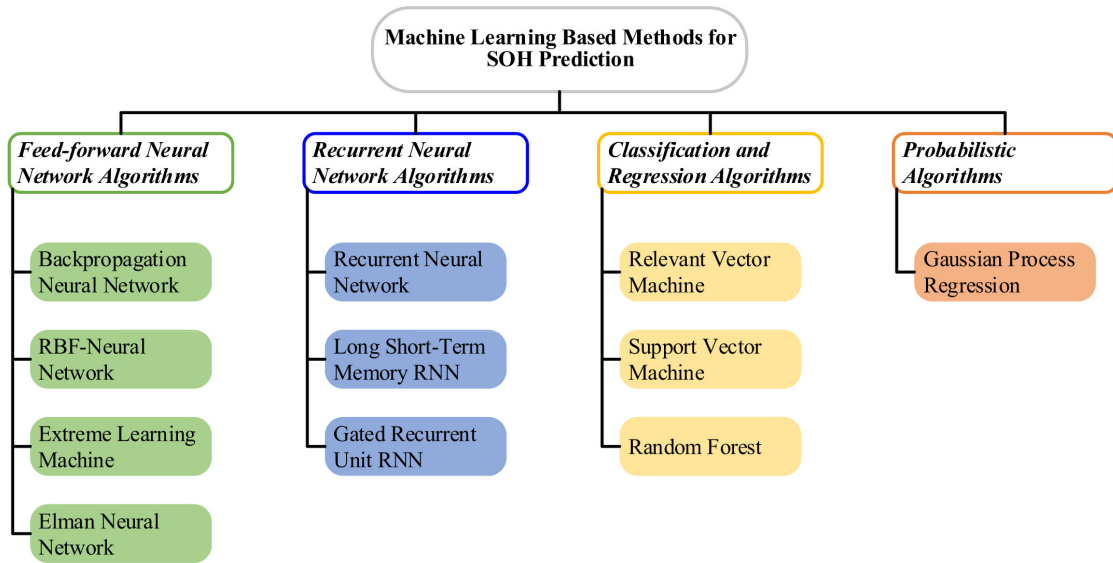


Figure 16: Classification of ML-based SOH prediction algorithms [68]

Feedforward Neural Network Algorithms:

In order to reduce predicted mistakes, Back Propagation Neural Network (BPNN) approach iteratively refines weights while embracing supervised learning. Gradient descent is used in backpropagation to optimize network parameters and increase predicted accuracy[68][61].

The Radial Basis Functions (RBF)-Neural Network, which is based on radial basis functions, excels at approximating complex functions, making it a strong candidate for battery state prediction tasks[67][68].

Extreme Learning Machine (ELM) is a single-layer network that fits well with large battery datasets due to its simplicity and quick training. It has randomly assigned input weights and analytically derived output weights[21].

Elman Neural Network: Elman networks combine recurrent connections with hidden layer links to provide temporal dependency modeling in sequential battery data[67].

Recurrent Neural Network Algorithms:

LSTM and Gated Recurrent Unit (GRU) represent two types of Gated Neural Network (GNN), which are particularly good at solving vanishing gradient problems in conventional recurrent neural networks RNN. Capturing long-term temporal dependencies in battery time series data is their area of expertise[68][67].

RNN: Because of problems with gradient vanishing, RNN which have been embedded in feedback connections for processing sequences, struggle to handle long-term relationships[67].

Long-Short-Term Memory RNN LSTM: By adding memory cells and gating mechanisms to regular RNN, LSTM overcome the shortcomings of its predecessors and adeptly handle the complexities of sequence data[52].

Classification and Regression Algorithms:

Relevant Vector Machine (RVM): This Bayesian variation, which excels in both classification and regression, is modeled after SVM. Its resistance against noise and probabilistic results are its important advantages[68].

Support Vector Machine SVM: SVM is an expert at both classification and regression, even when it comes to battery state prediction. It does this by identifying the best hyperplane for data classification [68][34].

Random Forest: A Random Forest is created by combining many decision trees. Its ensemble method offers significant prediction accuracy and resilience to overfitting[34].

Probabilistic Algorithms:

Gaussian Process Regression: This technique skillfully handles tiny datasets and offers informative uncertainty estimates by adding a probabilistic foundation to predictions[21].

Because every machine learning algorithm has unique benefits and drawbacks, the choice of which one to use depends on the type of data, the need for accuracy, and the requirements for computing efficiency. Understanding the subtleties of these algorithms enables the development of effective battery state prediction models, advancing the field of energy storage management[68].

3.3.3. Comparison of ML algorithms in SOH estimation (LSTM),(GRU) and Regression

In accordance with the introduction of more sophisticated topologies like LSTM and GRU networks, the field of recurrent neural networks RNN has undergone a revolutionary transformation. Traditional recurrent neural networks RNN have limitations that prevent them from efficiently capturing temporal dependencies and long-range patterns within sequential data[61].

These new neural architectures have taken advantage of these limitations and revolutionised the field. More specifically, the combination of these two methods with regression analysis has opened up new dimensions of predictive modeling. This has resulted in the creation of a dynamic toolset that can be used to unravel complex relationships and forecast numerical results.

Method	Advantages	Disadvantage
BP_NN	Provides reasonable accuracy while providing flexibility and ease in execution	When compared to other approaches, operating efficiency is lower
RBF-NN	Shows strong performance and global approximation ability	Operating efficiency is reduced, and there is a proclivity for local optimality.
ELM	It requires less computing and learns faster	The amount of buried neurons is sensitive
Elman-NN	Has a quick-approaching speed and adapts effectively to time-varying properties.	The training process is slow. and there is a risk of settling for local optimality
RIN	Efficient for data with a sequential nature.	Gradient disappearance or explosion is possible.
LSTM	Tracks long-term dependencies and saves data selectively.	Training execution is tough, and training acceleration is challenging.
LSTM	Tracks long-term dependencies and saves data selectively	Training execution is tough, and training acceleration is challenging.
GRU	Captures long-term sequential dependencies and resolves the LSTM gating mechanism.	Requires a huge amount of training data and a large storage tool.
RVM	Better sparsity and no Mercer restriction, avoiding overfitting and underfitting.	High computational load with huge datasets, incompatible with long-term prediction. A lack of stability is also a source of concern.
SVM	Allows for quick and precise estimation and provides adequate accuracy in high-dimensional systems.	It is computationally complex and lacks sparsity.
RE	Increased robustness and efficiency in the processing of complicated data.	Not appropriate for high-dimensional systems, and highly dependent on the amount of random trees.
GPR	Allows for the easy development of uncertainty estimation through the use of covariance. The characteristics are simple to understand.	It is sensitive to kernel functions and has a high computational cost.

Table 1: Vehicle Data sheet [26]

Long-Short-Term Memory (LSTM):

LSTM networks are revolutionary in the field of artificial intelligence due to their amazing capacity to retain knowledge over long periods, making them highly effective at modeling sequential data[61]. The fundamental components of the LSTM architecture include specialized memory cells that are intimately linked to input, forget, and output gates. The unique design of LSTM allows them to effectively capture complex temporal patterns, demonstrating their ability to learn and interpret long-term dependencies present in data.

The key difference between LSTM and regular RNN is that LSTM are skilled at selectively preserving or rejecting information. LSTM are particularly suitable for activities that require a sophisticated comprehension of complicated sequential intricacies[89].

The effectiveness of LSTM networks is seen in research efforts where individuals have utilized their skills to develop models for accurately predicting battery degradation. By utilizing a sophisticated understanding of extended relationships throughout time, Long Short-Term Memory LSTM models demonstrate exceptional proficiency in predicting real-world occurrences that involve intricate temporal connections. Significantly, they establish connections between the aging properties of batteries and the various operational conditions they experience. LSTM networks demonstrate their versatility and serve as useful instruments in predictive modeling. They can be applied in various domains, such as battery technology and the complex dynamics of temporal event forecasting[52].

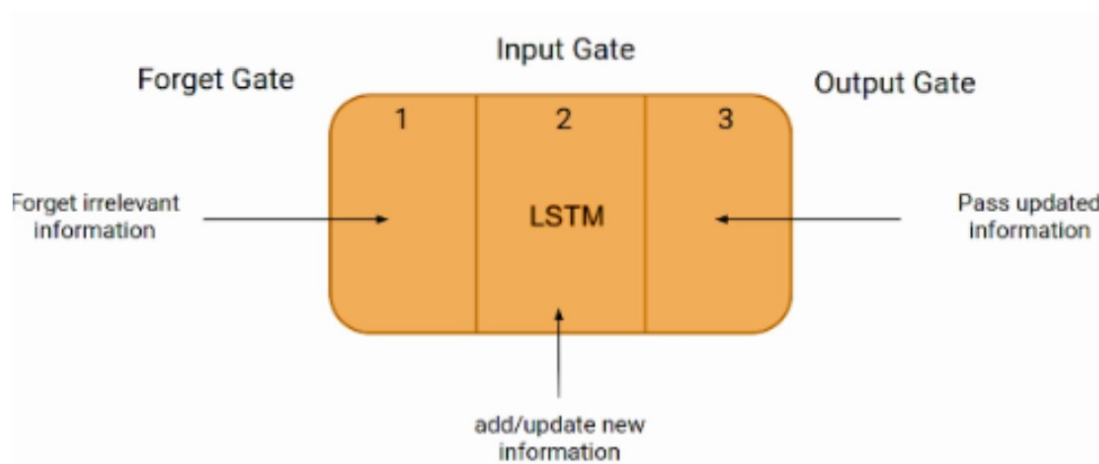


Figure 17: LSTM [59]

Gated Recurrent Unit (GRU):

A Gated Recurrent Unit GRU is a modified version of the Recurrent Neural Network RNN structure that incorporates gating mechanisms to regulate and oversee the transmission of information among the network's cells. GRU were first introduced in 2014 by Cho, et al[12]. and can be regarded as a relatively recent design, particularly when contrasted to the widely-accepted LSTM[31].

The architecture of the GRU enables it to dynamically capture interdependencies from extensive data sequences while retaining knowledge from preceding segments of the sequence. This is accomplished by utilizing gating units, like to those found in LSTM, which effectively address the issue of vanishing or bursting gradients in conventional RNN. These gates have the function of controlling the information that is retained or deleted at each time step. In this post, we will explore the intricacies of how these gates operate and how they effectively address the aforementioned challenges[52].

In addition to its internal gating mechanisms, the GRU operates similarly to an GRU, where the GRU cell processes sequential input data at each time step along with the mem-

ory, also referred to as the hidden state. Subsequently, the concealed state is reintroduced into the recurrent neural network RNN cell alongside the subsequent input data in the sequence. This process operates sequentially, functioning as a relay system to generate the intended outcome.

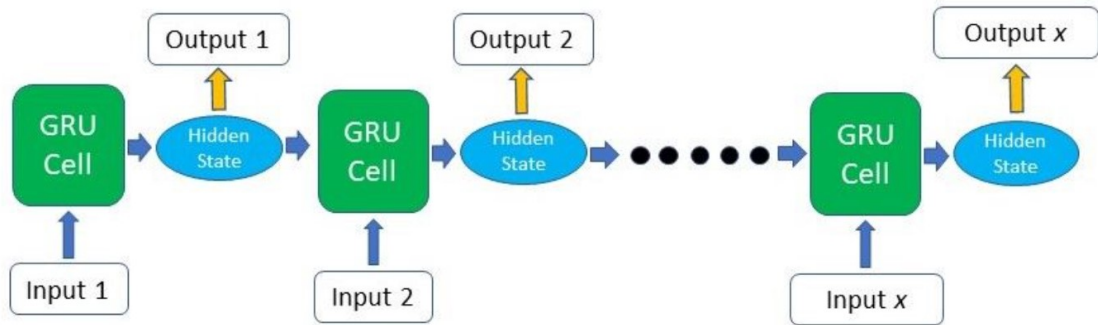


Figure 18: GRU [44]

Regression: Predictive Insights Unveiled

Regression, a fundamental aspect of predictive modeling, holds great importance in numerical prediction tasks. Regression involves creating a mathematical link between input data and a continuous target variable[34].

The primary goal is to obtain a regression function that minimizes the discrepancy between the predicted and actual target values. Regression is widely used in different forms, such as linear, polynomial, and time series regression[35].

Significantly, the combination of LSTM, GRU networks, and regression has fundamentally transformed the process of modeling sequential data. Both LSTM and GRU networks possess an intrinsic ability to comprehend the complexities of time, which makes them particularly skilled at numerical prediction tasks. The study undertaken underscores the effectiveness of LSTM in accurately predicting battery capacity and remaining useful life RUL several steps ahead[28].

4. METHODOLOGY

This chapter's primary goal is to give a summary of the master's thesis's methods and research aims. The chapter also outlines the study approach that was used, covering several topics including feature engineering, data collection, data description, experimental setup, cell description, and the creation of data-driven models to forecast LIB health.

4.1. Test Methodology

The setting up of meticulous testing procedures to gather important experimental data on LIB frequently found in EV is the first step in the study process. These test procedures have been carefully planned out and carried out to guarantee the precision and dependability of the data gathering. The collected data, which serves as a vital basis for the prediction of SOH, may initially include missing values. These are methodically resolved using the proper imputation techniques in order to preserve the dataset's integrity. The data also goes through a rigorous cleaning procedure that removes errors, inconsistencies, and outliers to guarantee that high-quality data is available for further analysis.

4.1.1. State of Health (SOH) Prediction Using Machine Learning Algorithms

Figure 19, which offers a step-by-step breakdown of how ML algorithms are employed for the prediction of SOH, summarizes the complexities of the SOH prediction process. This general implementation flowchart explains the sequential steps in the SOH prediction pipeline and provides thorough visual guidance. Data preparation, feature engineering, method selection, model training, performance evaluation, and prediction are among the phases shown.

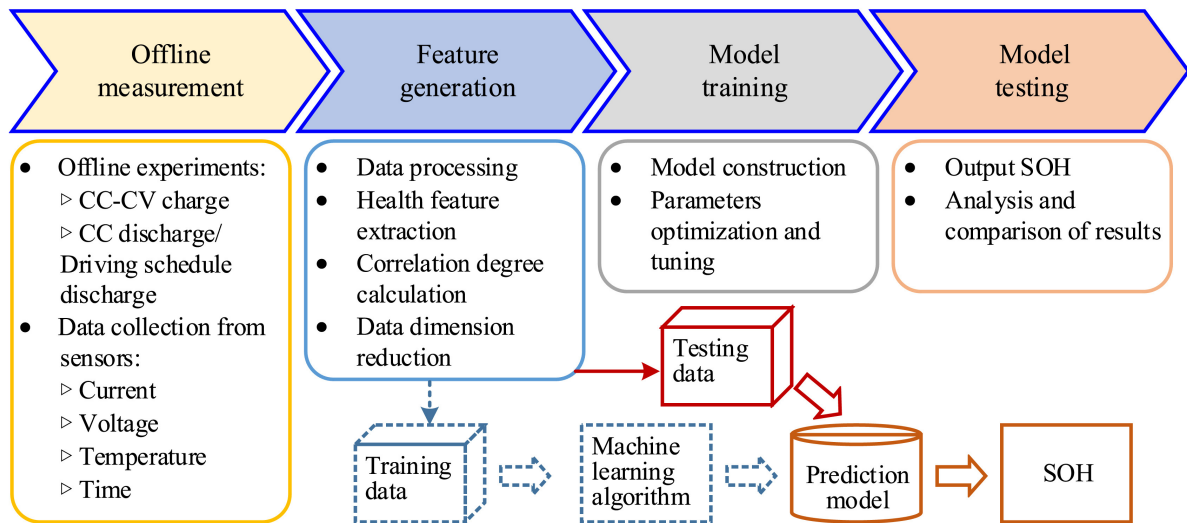


Figure 19: Generic implementation flowchart for SOH prediction using ML algorithms [68]

- **Data Preprocessing:**

Data preparation is the first step in the process, when battery testing data is carefully collected, cleaned, and arranged. This initial data refinement involves filling in missing numbers and cleaning the data to remove mistakes and abnormalities. This procedural step guarantees that the data set is consistent and devoid of anomalies that might interfere with further research.

- **Feature Engineering:**

The next step, feature engineering, is carefully choosing relevant attributes from the data collection that have significant influences on the SOH prediction. Furthermore, domain knowledge could lead to the development of additional features that improve the prediction power of the model. Certain features are undergoing scaling or normalization procedures in order to preserve homogeneity of the data.

- **Model Selection:**

The choice of an appropriate machine learning algorithm is the important step in the forecasting process. The flow chart highlights several possible algorithms, such as support vector machines, complex neural networks, and linear regression and decision trees. The decision is based on a number of variables, including the computational complexity, interpretability, desired predicted precision, and data set characteristics.

- **Model Training:**

The preprocessed dataset is used to rigorously train machine learning algorithms selected in the preceding phase. Usually, the dataset is divided into subsets for testing, validation, and training. The model optimizes its internal parameters to reduce prediction errors as it attempts to identify and internalize patterns from the training data.

- **Model Evaluation:**

The trained model's effectiveness is measured by means of rigorous evaluation methods. The correctness of the model is evaluated using a variety of performance indicators, including as Mean Absolutr Error (MAE), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and R-squared (Coefficient of Determination) (R^2). Methods like cross-validation evaluate the model's flexibility over a range of datasets.

- **Prediction:**

After all of these steps are completed, the trained and validated model is ready for practical use, which is the predictive phase. The model can be fed new and unknown battery data, which will produce an expected SOH value. This provides important information about the state of health and anticipated performance of the battery. Through the use of this flowchart as a framework, we hope to shed light on the intricate combination of approaches that go into SOH prediction. This acts as a guide that directs our empirical investigation and supports the analytical story that will be revealed in the upcoming chapters[68].

4.2. Test Bench Description

The Neware Technology Limited BTS-4000-5V6A cell tester was used for the measurements in this master's thesis. Bidirectional energy transport is made possible by the charging/discharge unit included with this cell tester. Individual battery cells are tested and cycled with the aim of simulating their service life and offering insights into their performance that go beyond battery capacity. The Neware BTS unit consists of an internal server that runs predetermined procedures, an auxiliary channel cabinet with battery or alligator connectors, and the tester itself with an integrated middle machine. Graphical presentation of channel data is possible with the manufacturer's BTS 7.6.1 software, and graphical examination of raw data generated during testing is possible with the BTS application.

The BTS unit's test channels can function within the current and voltage ranges of 0-6 A and 0-5 V, respectively, with a maximum power delivery of ± 30 W per channel. Within the prescribed range, the measurements of voltage and current have an accuracy of $\pm 0.05\%$. The tester comes with universal holders that firmly hold each cell, adapting to the size of the inserted round cells (type 18650, especially). To operate the test bench, process the collected data, and send it to the host computer, and industrial servers are used. This test bench is ideal for the tests that were conducted since the voltage range of 2.5 to 3.6 V corresponds well with the safe operating range of the cells that were utilised. Automation is the most commonly used approach to recording raw data and is heavily utilised to control test profiles.

The development of a frequency scanning approach that maximizes measurement time and information density is required since, as the frequency decreases, the measurement length increases. Because it provides specific parameter settings for steps per frequency decade, frequencies below or above 66 Hz, and the measurement period, which defines the number of averaged sine periods per measurement point, the single-sine mode gives the maximum precision and quality of impedance data. After the measurements are finished, the data collected can be kept and examined.

All tests were performed in the laboratory of the Carrisma Institute Of Electric connected And Secure Mobility (C-ECOS) department of the Research and Test Facility in the Centre of Automotive Research on Integrated Safety Systems and Measurement Area (Carissima). The ambient temperature during the cyclic aging tests can be assumed to be relatively constant with $T = 26.5 \pm 1.5$ °C (room temperature) due to external temperature control in the laboratory.

4.3. An explanation of the test configuration

The individual lithium iron phosphate cells used are clamped between the two battery terminals of the "BTS-4000-5V6A" cell tester test bench and remain there during the subsequent aging tests. For the polarity arrangement, the upper connection is considered positive and the lower connection is negative. Furthermore, there are test bench voltage measuring lines on the battery connection terminals or on the alligator terminals to record

the current battery voltage. Since line and connection resistances would produce too high a resistance value when connected in series with the measuring resistor and thus distort the measurement result, a four-wire measurement is used. Using a four-wire circuit allows negligible resistance through the power lines, meaning the measurements taken are significantly more accurate than with a two-wire circuit.

Overall, it should be noted that good contact must be ensured at all times between the cells and the battery connection terminals for optimal measurement results. A total of 45 individual cells are used for the test series, of which 39 individual cells are attached to the five main modules (06–10) with the associated server 1, and the remaining nine individual cells are attached to two further main modules (01–02) with the associated server 2. In addition, the current temperature within the test laboratory used is recorded with the associated date and time to monitor the room temperature during the tests. This can be used to show that the temperature is in the defined temperature range during the aging tests.

The cell tester monitors the charging/discharging current and the voltage of each cell and can switch off at critical operating points such as Undervoltage or overvoltage, or charging or discharging current that is too high, thereby protecting the individual cells from destruction or damage. To obtain a more precise overview of the real structure of the overall test setup and the integration of the individual cells into the test setup, the following Figures 20 and 21 show the front view. The individual cells shown are tested and aged at room temperature.

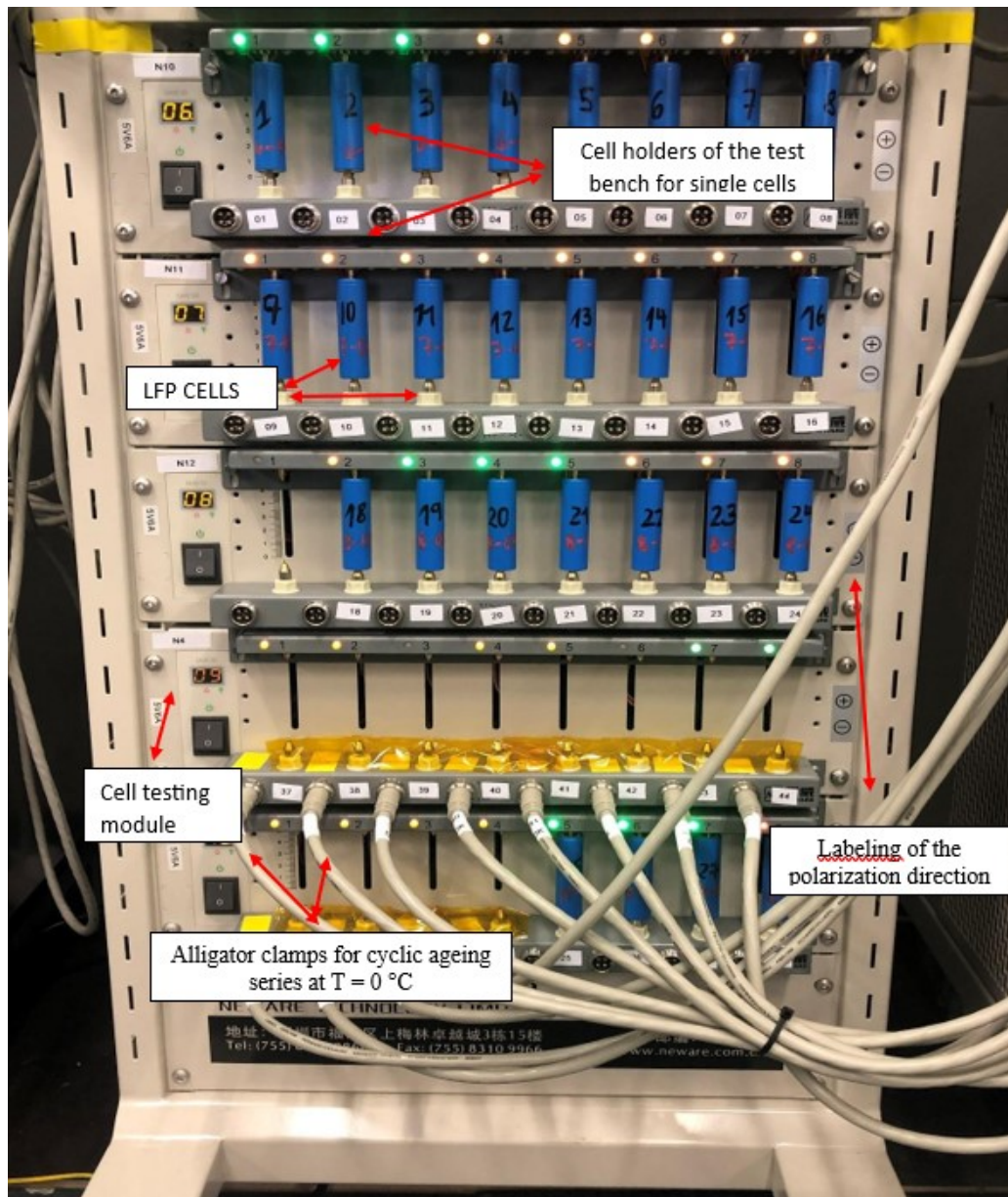


Figure 20: Diagram of the Individual Cell Arrangement and Experimental Test Setup. [44]

Figure 20 shows the experimental setup of the cyclic aging tests. On test module 9 with the measuring channels 01 - 08 and on test module 10 with the measuring channels 01 - 04, the alligator clips with the extension leads are located in the temperature test chamber, with a set temperature of $T = 0\text{ }^{\circ}\text{C}$. For safety reasons, these test channels on the cell tester itself are taped with Kapton tape to prevent accidental insertion of another single cell and thus destruction of the test channel as a result of excessive voltage. The connecting cables are sealed with a sealing compound in the side opening of the temperature chamber to ensure airtight operation within the temperature chamber and to guarantee that a constant temperature is maintained in the temperature chamber

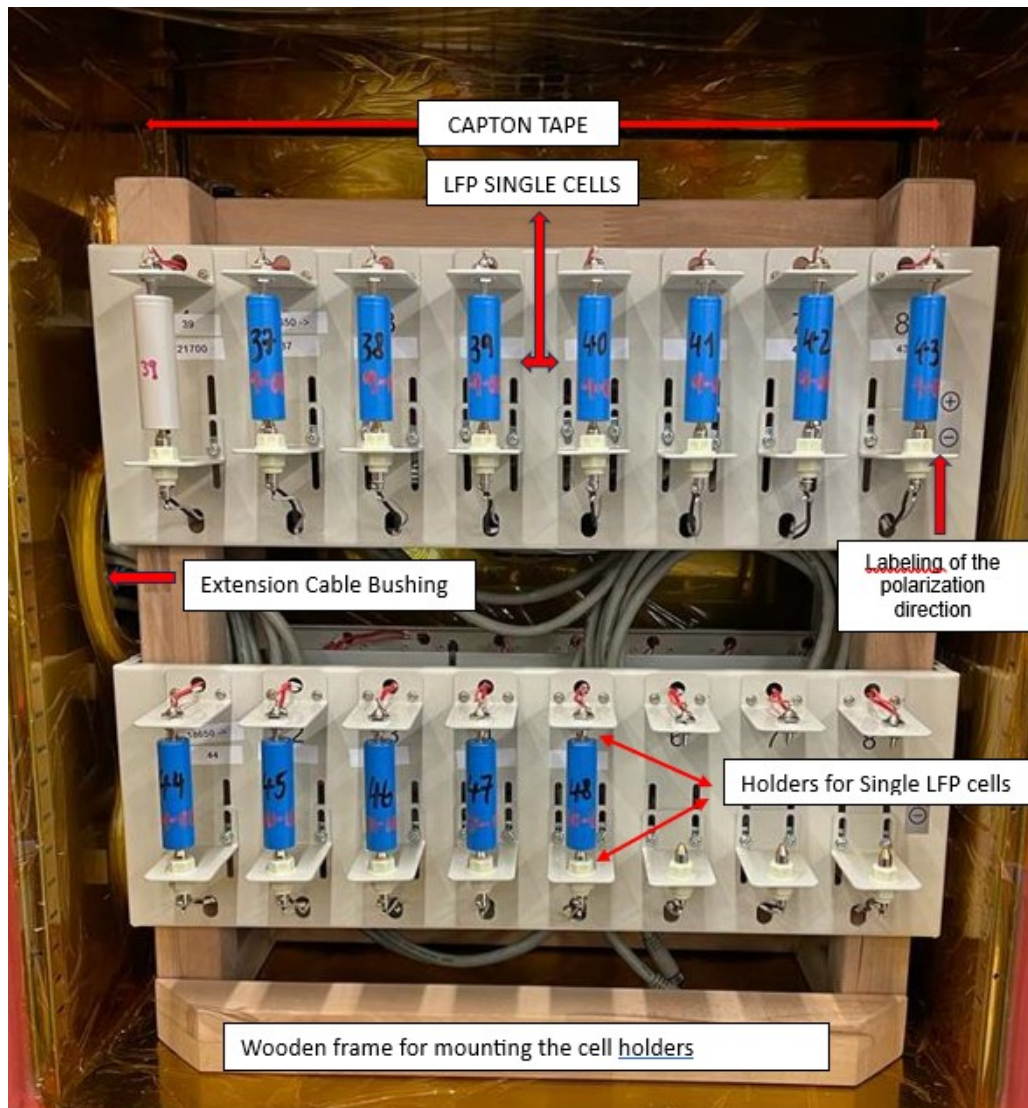


Figure 21: Diagram illustrating the Configuration of a Controlled Temperature Chamber for Cell Testing [44]

Figure 21 describes the integration of the wooden frame, including the cell holders and individual cells inside the temperature test chamber. The temperature test chamber is of the "VT 4011" type from Vötsch Industrietechnik, in which the temperature is set to $T = 0\text{ }^{\circ}\text{C}$ for the cyclic ageing tests on the touch panel. The temperature is constantly set to $T = 0\text{ }^{\circ}\text{C}$, with a maximum spatial deviation of $\pm 0.5\text{ K}$ to $\pm 1.5\text{ K}$. In addition, the inner walls of the temperature test chamber are lined with Kapton tape to prevent potential short circuits. In order to position the test holders of the individual cells well in the climate chamber, a wooden frame is used into which the cell holders are hung. The data sheet for the temperature test chamber described can be found in the appendix. All cells are labeled before the start of the test and coded with the corresponding test locations on the test bench for better clarity. In addition, an incoming inspection is carried out before the LFP cells are installed on the test bench. This consists of a visual inspection for external damage or abnormalities and an incoming inspection using a "UT 139C" digital

multimeter to check the voltage level. All LFP cells meet these quality requirements and have the same initial average voltage level of 3.30 V.

The temperature in the test lab is continuously tracked and recorded, along with the corresponding date and time, for both testing configurations. This rigorous recording procedure makes it possible to monitor room temperature precisely throughout the aging tests and offers solid proof that the temperature regularly stays within the permitted range.

4.4. Description of the LFP single cells

Specification	Characteristics
LiFePO ₄	Cell chemistry
3,2 V	Nominal voltage
1,5 Ah	Nominal capacity on discharge
4,5 A (3 C)	Max. continuous discharge current
1,5 A (1 C)	Max. continuous charging current
3,6 V	End-of-charge voltage
2,5 V	Final discharge voltage
0 °C bis +60 °C	Temperature range during charging
-30 °C bis +60 °C	Temperature range during unloading
-50 °C bis + 60 °C	Storage temperature:
1550mAh - 0.5C (current value of 1500mA at 1C°)	Standard Capacity
40 g ± 4 g	Weight
18,35 mm ± 0,15 mm	Diameter
65 mm ± 0,2 mm	Height

Table 2: Data sheet of the LFP cell used

This study investigates LFP (Lithium Iron Phosphate) single cells produced by the manufacturer "i-tecc⁺." Examining these cells in low-voltage applications within car electrical systems is particularly intriguing. Table 2 provides a comprehensive summary of the essential parameters of the cell, presenting the detailed technical specifications. The LFP single cells have a nominal capacity of 1.5 Ah and their cyclic operation follows the voltage restrictions set by the manufacturer, which range from 2.5 V to 3.6 V. The study constantly standardizes the C-ratios to the nominal capacity to ensure uniformity and comparability. All the LFP single cells used adhere to the cylindrical 18650 design, which is a well-used and commercially available standard in battery technology. This selection guarantees uniformity and harmony with existing battery setups, enhancing the dependability and significance of the research results.



Figure 22: LFP CELL [44]

4.4.1. Selecting the charge/discharge current for capacity and energy measurements

Lithium iron phosphate LFP cells, along with lithium titanate cells, are notable for their extended cycle service life compared to other rechargeable lithium-ion cells. Within this framework, the investigation of the ageing processes of LFP cells under high current loads emerges as a particularly captivating area of study, with a specific emphasis on the currents involved in charging and discharging during cyclic ageing. As per the manufacturer's datasheet, the LFP cell has a maximum charging current of 1.5 A (1 C) and a maximum discharging current of 4.5 A (3 C).

This study focuses on purposely accelerating the cyclic ageing process by using LFP cells at higher currents and discharge depths. The reason for adopting this strategy is based on the recognition that inadequately chosen currents can lead to a higher margin of error in the final estimation of the maximum error in the cell test bench. Thus, to reduce errors in capacity and energy measurement, the continuous charge/discharge currents of the LFP single cells are intentionally selected to be as high as feasible.

The primary rationale behind this decision is based on the observation that measurement inaccuracies tend to increase when the measuring current diminishes. The main cause of this increase is generally ascribed to mistakes in voltage and current measurements. Therefore, the study seeks to reduce potential errors and improve the accuracy of the results by using stronger currents during testing. This research endeavor seeks to enhance the overall understanding of lithium-ion cell behavior and optimize their performance.

4.4.2. Loading Procedure

The process of loading lithium-ion single cells typically involves a Constant Current-Constant Voltage (CC-CV) charging method. This method involves charging the cell with a constant current up to a predetermined voltage level, then proceeding to a constant voltage charging stage with a reduced current until the current reaches a predetermined threshold. A CC-CV charging method is typically used. During the process of charging, it is of the utmost importance to make certain that the charging current and voltage are kept within the limits that have been established.

This is done to prevent irreversible reactions and the breakdown of the electrolyte. It is common for the manufacturer to establish the maximum voltage that a lithium-ion cell is capable of being charged to. Should this voltage be exceeded, the cell may sustain irreparable damage, which may include a reduction in its capacity as well as potential safety issues. In addition, excessive charging can result in the development of lithium metal on the anode of the cell, which can cause the cell to short-circuit and lead to thermal runaway. It is possible to adhere to standardised protocols, such as ISO 12405-4, to guarantee that the testing procedure is accurately and consistently carried out. To accomplish this, you will need to make use of an apparatus that is capable of controlling and monitoring the charging parameters, as well as terminating the charging process once the cell has reached the limitations that have been defined.

To summarise, the loading technique for lithium-ion single cells entails charging the cell using a CC-CV approach while ensuring that the charging current and voltage remain within the limitations that have been established. It is of the utmost importance to avoid overcharging cells since this can result in irreversible damage to the cell as well as safety hazards. It is possible to utilize standardized protocols in order to guarantee that the testing technique yields accurate and consistent results.

4.4.3. Test Plan and Cyclic Aging Data Collection

Under this part, we will cover the test design as well as the tests that were carried out to obtain data on the cyclic aging of the cell.

The plan for the test was developed with the purpose of determining the effects of cyclic ageing on the performance of cells and collecting data that is pertinent for analysis. For the purpose of ensuring that the measurements were accurate and consistent, a methodical methodology was utilized. During the course of the experiment, the cell in question was subjected to a sequence of charge and discharge cycles, during which variables such as current levels, discharge depths, and temperatures were varied. Both the simulation of real-world working conditions and the induction of accelerated aging effects were accomplished through the careful selection of these factors.

The tests were carried out with the assistance of a specialized testing apparatus. It was possible to exercise exact control over the current levels, discharge depths, and temperature conditions during the test cycles because the cell was attached to a testing device that was able to manage and monitor the charging and discharging processes. In the course

of the charge and discharge cycles, several performance characteristics, including voltage, current, capacity, and energy levels, were measured and recorded at predetermined intervals. It is possible that further measurements, like impedance spectroscopy, were carried out to get more depth of understanding regarding the electrochemical behavior of the cell. Extensive analysis was performed on the data that was gathered in order to evaluate the deterioration of cell performance over the course of time as a consequence of cyclic aging. For the purpose of determining the level of aging and its influence on the overall health and performance of the cell, key indicators such as capacity loss, voltage drop, and impedance changes were analyzed.

The results that were acquired from the experiments contributed to a better understanding of the elements that influence deterioration by providing significant insight into the cyclic aging behavior of the cell. To conduct additional research and modeling, the data served as a foundation.

During the testing process, three specific rates of current discharge (1C, 2C, 3C) are examined. Each discharge rate is tested at four different depths of discharge (100% DOD, 75% DOD, 50% DOD, 25% DOD). The tests are conducted at a constant room temperature of $T = 26.5 \text{ }^\circ\text{C} \pm 1.5 \text{ }^\circ\text{C}$. Furthermore, a cyclical aging investigation is conducted for the three existing rates (1 C, 2 C, 3 C) within a temperature chamber, with the battery being fully discharged and maintained at a temperature of $T = 0 \text{ }^\circ\text{C} \pm 1.5 \text{ }^\circ\text{C}$. A total of 27 LFP single cells were included in the cyclic aging research. The utilization of different discharge rates and the grouping of cells aimed to investigate the impact of current rates on aging behaviour and performance degradation. By maintaining a controlled room temperature, the study provided a stable testing environment for accurate data collection and analysis.

4.5. Feature Engineering

In this section, we will examine, describe, and put into practice several data aggregation and data cleaning processes for the dataset that was obtained from the test bench.

4.5.1. Data Description

As part of the examination of the aging process, the data obtained from the test bench is collected. For the purpose of obtaining the data set, forty-five LIBs were subjected to three various current discharge rates (1C, 2C, and 3C), each of which had four different depths of discharge (100% DOD, 75% DOD, 50% DOD, and 25% DOD).

These LIBs were then illuminated and analysed at a room temperature of $26.5 \text{ }^\circ\text{C} \pm 1.5 \text{ }^\circ\text{C}$. In addition, a cyclic ageing research is conducted in the temperature chamber at full depth of discharge and a temperature of $T = 0 \text{ }^\circ\text{C} \pm 1.5 \text{ }^\circ\text{C}$. This study is conducted for the three current rates, which are 1 C, 2 C, and 3 C. The detailed description of the data is presented in Figure 21, which can be found here.

Cycle data, which includes information pertaining to cycles as well as the number of cycles that are executed throughout the test, the number of lines that are formed in this sheet, and finally statis data, which will present the instructions in a step-by-step fashion. Last but not least is the data detail, which displays the entire record of the data.

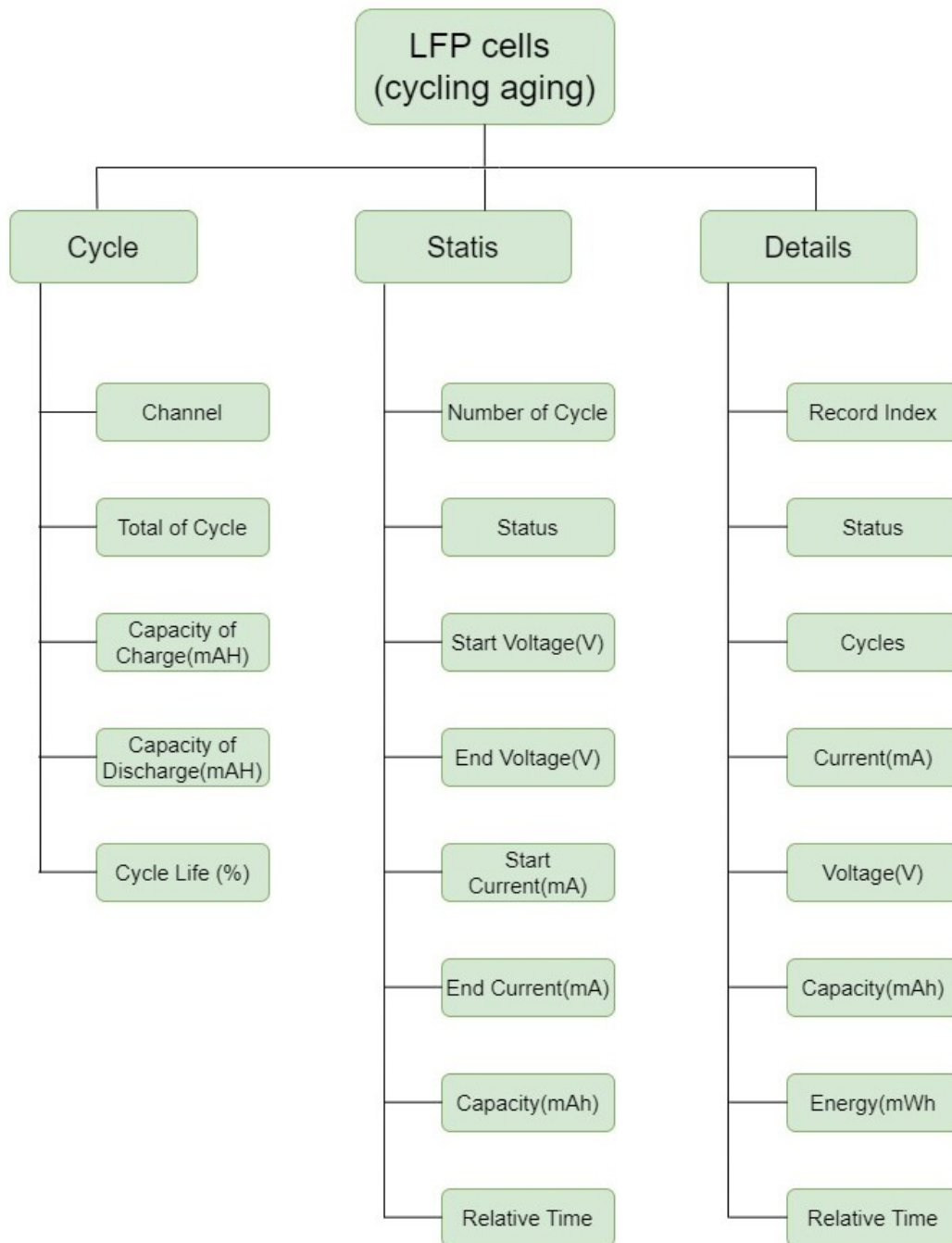


Figure 23: Detailed description of cyclic aging data

For a more detailed overview of the dataset and experimental setup, refer to Figure 24

Cell No	CC Unload Rate	Depth of Discharge (DOD)	Temperature
1	1C	100%	26.5°C ± 1.5°C
2	1C	100%	26.5°C ± 1.5°C
3	1C	100%	26.5°C ± 1.5°C
4	1C	75%	26.5°C ± 1.5°C
5	1C	75%	26.5°C ± 1.5°C
6	1C	75%	26.5°C ± 1.5°C
7	1C	50%	26.5°C ± 1.5°C
8	1C	50%	26.5°C ± 1.5°C
9	1C	50%	26.5°C ± 1.5°C
10	1C	25%	26.5°C ± 1.5°C
11	1C	25%	26.5°C ± 1.5°C
12	1C	25%	26.5°C ± 1.5°C
13	2C	100%	26.5°C ± 1.5°C
14	2C	100%	26.5°C ± 1.5°C
15	2C	100%	26.5°C ± 1.5°C
24	2C	25%	26.5°C ± 1.5°C
25	3C	100%	26.5°C ± 1.5°C
26	3C	100%	26.5°C ± 1.5°C
27	3C	100%	26.5°C ± 1.5°C

Figure 24: Dataset

4.5.2. Data processing and Analysis

This subsection will present a summary of the selected data used for constructing a data-driven model and describe the approaches used for cleaning and aggregating the data. The battery degradation data, as provided in the dataset, was chosen for utilisation in the machine learning model. In order to prepare the data for analysis, engaging in feature engineering, a process that entailed converting and deriving significant characteristics from the raw data.

This subsection describes the thorough methodology that was applied to the subfolders containing the Excel files containing the raw cell testing data in order to process, clean, and analyse it. The set of data consisted of 36 cells, with each cell providing 50 cycles of testing information. Multiple sheets in an Excel document, including "Cycle," "Statis," and "Detail." The aim of this methodology was to extract relevant information from the "Cycle" and "Detail" sheets, with that exception only 19 cells were utilised for additional testing and training due to missing cycles. The extracted data was then consolidated into 19 distinct Excel files to facilitate further analysis. Insights were generated through the visualisation of critical parameters and statistical metrics.

In the course of my research, I developed a customised Python script to simplify the process of extracting and analysing data from many Excel files. This script played a crucial role in efficiently navigating a complicated structure of subfolders and retrieving information from 36 specific cells. By utilising the functionalities of established libraries like 'os' for navigating the file system and 'pandas' for efficient data management, the script effectively processed every subfolder and Excel file.

The main emphasis of the script was on two crucial worksheets within every Excel file: the "Cycle" and "Detail" worksheets. From the "Cycle" sheet, it carefully retrieved crucial parameters that are necessary for comprehending the battery cycles. The parameters encompassed the initial and final voltages, currents, and different time intervals, such as periods of rest, charging, and discharging. Simultaneously, the script accessed the "Detail" tab to collect additional specific data, recording details about the charging, discharging, and resting phases at intervals of one second. The systematic extraction technique led to a thorough consolidation of data. The data from each cell's cycle was meticulously organised into a structured style, facilitating a more thorough and detailed study.

4.5.3. Data Aggregation and Data Cleaning

This subsection of the thesis discusses the critical phases of data aggregation and cleaning, which were instrumental in improving the quality and integrity of our analysis. The preliminary dataset, obtained from the "Detail" and "Cycle" columns of multiple Excel documents, displayed a diverse array of errors and deviations. In order to address these challenges and ensure the analysis remained consistent, accurate, and pertinent, a comprehensive methodology was implemented.

Data Aggregation:

When considering data aggregation, the initial step involved the consolidation of information from the "Cycle" sheet pertaining to each cell and cycle. This involved standardising parameters, including start and end voltages, currents, capacity, and rest intervals, and incorporating them into a solid structure. This was of the utmost importance in assessing the collective behaviour of cells during testing. Furthermore, the data obtained from the "Detail" page, which was recorded at intervals of ten seconds, was carefully synchronised with the corresponding cycle data. This allowed for a thorough analysis of the charging, discharging, and resting periods in connection with each cycle.

Data Cleaning:

The procedure of data cleansing comprised two essential stages. We began by identifying and addressing outliers, which had the potential to introduce bias into the analysis. Sophisticated methodologies such as z-score analysis and visualisation were employed to identify and examine outliers in the values of current, voltage, and dq/dv . Each outlier underwent a meticulous assessment to ascertain its influence on the dataset prior to a determination regarding its inclusion or exclusion. Subsequently, we tackled the concern regarding absent data, an unavoidable characteristic of experimental datasets. Methods such as interpolation were utilised to approximate absent values, thereby guaranteeing

the absence of erroneous conclusions introduced by this strategy. In cases where interpolation was not possible, absent data points were transparently identified and documented. Adherence to this rigorous methodology for data aggregation and cleansing was critical to maintaining the dependability and accuracy of our analysis.

4.5.4. Removing Outliers and Handling Missing Values

This section examines the fundamental processes of outlier exclusion and missing value management, which are indispensable for guaranteeing the precision and reliability of our analyses. Strict protocols were followed to detect and address missing data and outliers, so as to prevent any potential misunderstandings regarding the outcomes of cell assays.

Removing Outliers: In order to eliminate outliers, we initially implemented z-score thresholding on each cycle's current, voltage, and dq/dv values. By employing this method, data points that deviated substantially from the mean were identified in accordance with a predetermined threshold. By employing this methodology, a standardised metric was established to identify data points that may have been outliers, which simplified the exclusion determination process. Then, each identified outlier underwent a comprehensive validation process. Retained were data points that originated from valid experimental conditions or that represented crucial aspects of cellular behaviour. On the contrary, values considered to be artefacts resulting from measurement errors or other irregularities were eliminated.

Handling Absent Values: With respect to the management of absent values, interpolation strategies were executed carefully. The predominant technique employed was linear interpolation, particularly in situations where missing data points were surrounded by legitimate values. In every instance, interpolation was determined following a thorough evaluation of the characteristics of the data and the interpolation's applicability to the particular circumstance. When interpolation was deemed impracticable or unsuitable, those situations were openly recorded, accompanied by a rationale for the corresponding decisions. This methodology guaranteed that the whereabouts of absent data and the rationales for their non-appearance were unambiguously conveyed, thereby preserving complete candour regarding the constraints of our dataset.

4.6. Machine Learning Models

This section provides an in-depth analysis of the complex mechanisms and derivations underlying the machine learning models and algorithms utilised for capacity prediction. In particular, we shall conduct a comprehensive examination of the critical dropout technique and the Long Short-Term Memory LSTM model, all of which contribute to the robustness and precision of our algorithm for estimating capacity.

4.6.1. LSTM Model:

An improved version of the recurrent neural network RNN, the long short-term memory network LSTM not only excels at learning long-term dependent sequence data that RNN is incapable of processing, but it also solves the gradient inflation and gradient disappearance issues that plague conventional RNN. Temporal sequences are associated with both the HI extracted during cyclic charging and battery SOH[43][52]. For precise lithium ion battery SOH prediction, the LSTM is thus implemented. The LSTM network structure is illustrated in Figure 25.

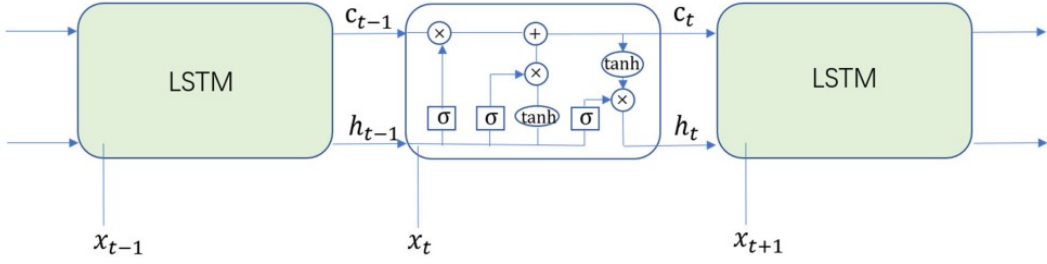


Figure 25: LSTM structure [43]

LSTM networks are characterised from conventional RNN by the incorporation of a cell state, which serves to store data pertaining to long-term dependencies. LSTM networks regulate the flow of data through the utilisation of three gates: the forget gate, the input gate, and the output gate. The sigmoid activation function is employed to designate a value between 0 and 1 to each gate. The formula for these calculations appears below[52].

$$f_t = \text{sigmoid}(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (12)$$

$$i_t = \text{sigmoid}(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (13)$$

$$o_t = \text{sigmoid}(W_o \cdot [h_{t-1}, x_t] + b_o) \quad (14)$$

In LSTM networks, which information from the previous node should be retained or discarded according to its significance is determined by the forget gate. It forgets input that is considered irrelevant to the task at hand. In contrast, the input gate retains only those current inputs that are important to the computation at hand. The output gate subsequently provides the task-required current memory unit value. By utilising these three gate states to update the hidden state and the cell state, the network is capable of performing both short-term and long-term memory operations[43][17].

4.6.2. Dropout Technique:

The method of dropout, in which a unit in a neural network is briefly removed from a network, was initially proposed by Hinton [30]. This method was initially prompted by the role that plays in the process of evolution. Dropout was applied to feed forward neural networks and RBMs by Srivastava [71], who found that a probability of dropout of approximately 0.5 for hidden units and 0.2 for inputs performed well for a variety of tasks.

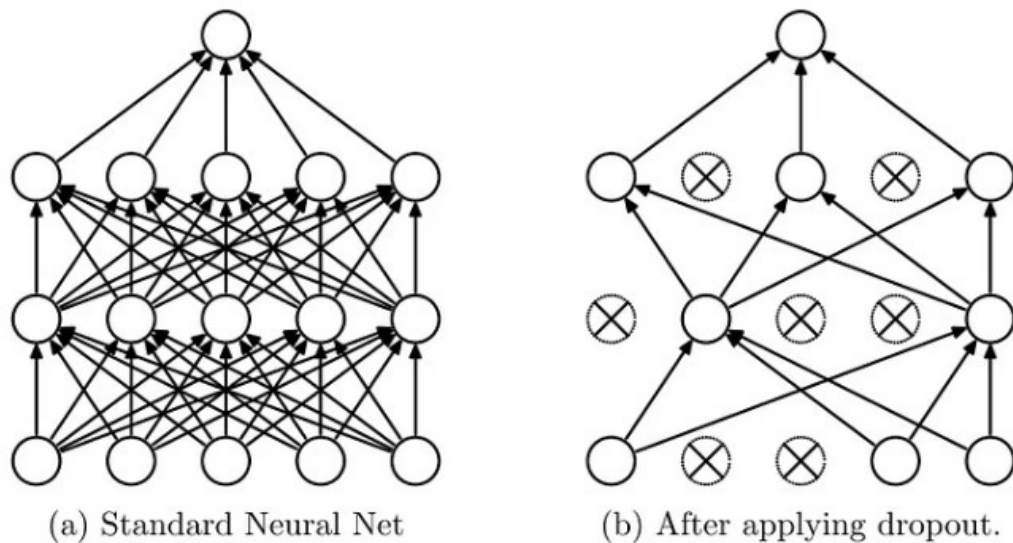


Figure 26: The schematic diagram of dropout neural network model. (a) standard neural network with 2 hidden layer; (b) neural network after applying Dropout technology [20]

Each hidden unit in a neural network trained with dropout must learn to work with a randomly chosen sample of other units." The hidden units should become stronger and be pushed to make useful features on their own, without needing other hidden units to fix their mistakes. "The derivative that each parameter gets tells it how to change so that the final loss function is lower, taking into account what all the other units are doing." Because of this, units may change so that they fix what the other units did wrong. All of this could lead to complicated co-adaptations. This then leads to overfitting because these co-adaptations don't work with data that hasn't been seen before. Srivastava et al. (2014) think that dropout stops co-adaptation of each hidden unit by making it uncertain whether other hidden units are present[71].

In particular, the selection of this specific dropout rate is not arbitrary, but is rooted in thorough consideration and empirical experimentation, all in the pursuit of fine-tuning model performance for accurate capacity prediction.

4.6.3. Model performance metrics:

This section discusses objective criteria for evaluating the extent to which a model accurately represents the fundamental patterns in the data and its dependability in making predictions for the future. Some frequently utilised metrics are Mean Absolute Error MAE, Mean Squared Error MSE, Root Mean Squared Error RMSE, and R-squared. Each of these metrics provides a distinct viewpoint on the performance of the model, covering a range of factors such as the average size of errors and the amount of volatility in the data that can be explained. In the following sections, we will thoroughly examine each of these metrics, including in-depth explanations of how they are calculated, interpreted, and their significance within the framework of regression analysis[15].

Mean Absolute Error (MAE):

- **Definition:**MAE measures the average magnitude of errors in a set of predictions, without considering their direction. It's the average over the test sample of the absolute differences between prediction and actual observation where all individual differences have equal weight[33].
- **Interpretation:**A lower MAE value indicates better model performance. It represents the average error made by the model in predicting the outcome. It's particularly useful because it gives a direct idea of the magnitude of error[15].

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (15)$$

Mean Squared Error (MSE):

- **Definition:**MSE is the average of the squares of the errors. It measures the average squared difference between the estimated values and the actual value[33].
- **Interpretation:** MSE gives more weight to larger errors, as it squares the residuals before averaging, making it useful for identifying models that make large errors. A lower MSE value indicates a better fit[15].

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (16)$$

Root Mean Squared Error (RMSE):

- **Definition:**RMSE is the square root of MSE. It measures the standard deviation of the residuals or prediction errors[10].

- **Interpretation:**RMSE is sensitive to outliers and gives a relatively high weight to large errors. Like MSE, a lower RMSE value indicates a better fit, but it's in the same units as the response variable, making interpretation easier[15][10].

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (17)$$

R-squared (R²):

- **Definition:**R-squared, also known as the coefficient of determination, measures the proportion of the variance in the dependent variable that is predictable from the independent variables.
- **Interpretation:**R² values range from 0 to 1 and are commonly expressed as percentages. A higher R² value indicates that more of the variability in the outcome is explained by the model. However, a high R² does not always mean a good fit, especially in non-linear models[15].

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (18)$$

4.7. Model Running Time

Performance evaluation of machine learning models depends on computational time. Our work focuses on the Long Short-Term Memory LSTM model, which is noted for its predictive accuracy but computationally demanding. These models require a lot of processing power and time for training and prediction, making their efficiency a concern.

In our research, we examined the computational time of running the LSTM model. This included evaluating the models' hardware and software settings and training and inference times. The goal was to balance the model's predicted accuracy and computational economy so it may be used in real-world situations without compromising performance.

We explored numerous ways to optimise the LSTM model's running time. These included parallel computing, model pruning to simplify computations, and hardware acceleration. We are investigating these ways to improve the LSTM model's applicability and scalability, especially for jobs that demand fast predictions, such as electrochemical cell ageing prediction. We seek to provide insights on balancing accuracy and computing efficiency to aid real-time LSTM model deployment decisions.

5. RESULTS AND DISCUSSION

This chapter presents the results obtained from implementing machine learning models specifically tailored to forecast the State of Health SOH of batteries. Prior to analyzing the results, we do a thorough data analysis and present important findings using graphs, as explained in Section 5.1.

In Section 5.2, we thoroughly examine the implementation of three separate models. Long Short-Term Memory LSTM, Gated Recurrent Units GRU, and regression models. The models have undergone meticulous design and optimization to generate the most accurate predictions of SOH. Their performance has been extensively assessed and will be detailed in the subsequent sections.

5.1. Analysis and Visualisation of Data:

This master's thesis thoroughly examines the essential characteristics of the model. It is crucial to assess the impact of these features on the model's performance, as it allows us to decrease the complexity of the data and hence minimize the computing burden associated with the model. This area of data visualisation seeks to examine the efficiency of battery cells in different operational circumstances and to track the degradation of batteries as time progresses.

5.1.1. Cycle Aging with Voltage vs Discharge Capacity

The purpose of this data visualisation phase is to examine the performance of battery cells in different operational scenarios and monitor the deterioration of the batteries over time. Figure 27 displays the various patterns of battery degradation over time under different operational conditions. The batteries' state of health SOH was assessed by testing their capacity retention over some time. The data indicates a statistically significant linear correlation between a battery's health and its charging rate (c-rate) and depth of discharge DOD.

According to this research, the charging and discharging rate of a battery has a substantial influence on its overall health. The findings suggest that the construction rate of battery management systems may significantly impact long-term performance. The subsequent phase of the investigation involves making a deliberate decision regarding the variables to extract from the utilized dataset.

The objective of this stage was to ascertain whether pairs of variables exhibit a significant correlation, either positive or negative, suggesting the presence of a linear association between the variables. Due to this connection, the model's input may become repetitive, leading to a higher computational workload for the model with only a marginal improvement in its discriminatory capacity. The primary goal of the variable selection stage is to optimize the model's performance by eliminating unneeded or redundant characteristics from the dataset.

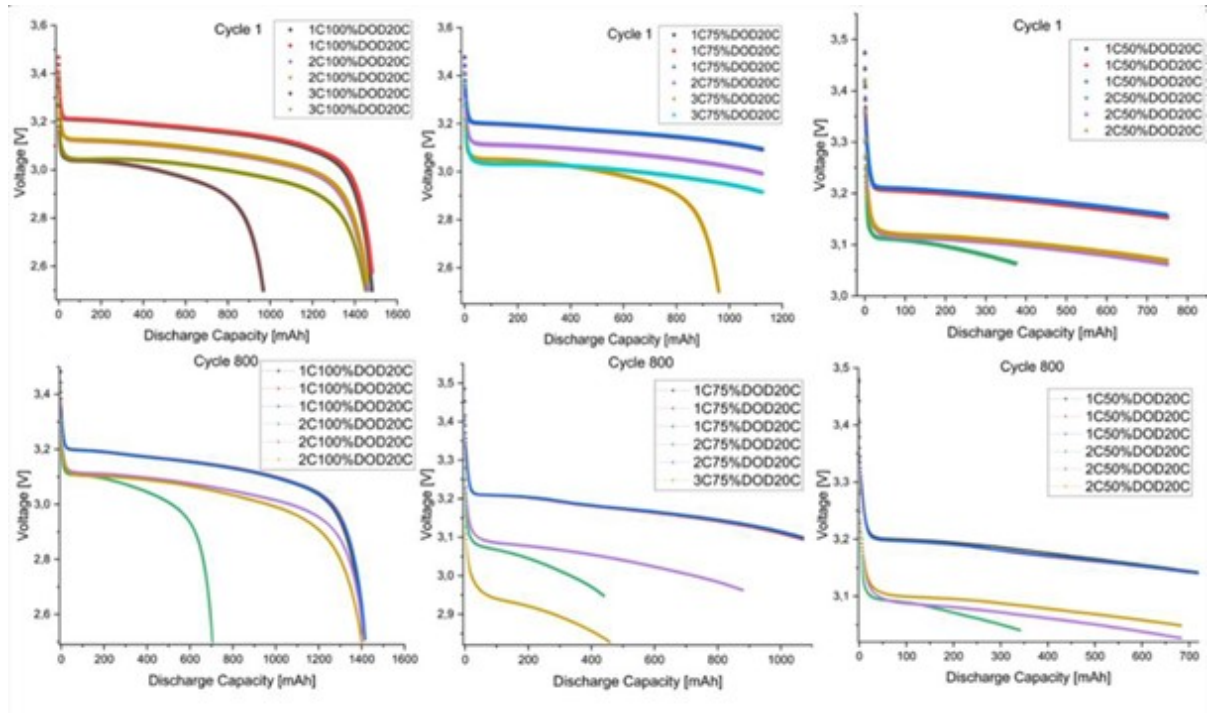


Figure 27: Cycling aging of cells for different DOD and c-rates

5.1.2. Discharge Voltage Analysis

A systematic and methodical approach was employed to generate graphs that effectively illustrate the correlation between Discharge Voltage and Time for each individual battery cell. The aim was to create comprehensive and easily understandable visual representations. To accomplish this, we decided to collect data points at regular intervals, particularly selecting data from every 100th cycle to create our graphs.

The intentional sampling method offered several significant advantages. Firstly, it significantly reduced the quantity of data points, which was crucial in preventing the graphs from becoming excessively crowded and challenging to interpret. By graphing the data at regular intervals of 100 cycles, we successfully recorded significant variations in the discharge voltage as time progressed. These alterations often indicate significant shifts in battery behavior, which are essential for comprehending the overall performance and condition of the battery.

In addition, the selected intervals enabled us to focus on the most notable trends and patterns in the discharge voltage over time. This level of specificity was crucial in identifying pivotal points in the battery's lifespan, such as when its performance starts to noticeably decline or when the battery undergoes accelerated deterioration.

The application of this technique in graphing produced transparent and instructive visual depictions of the discharge voltage patterns of individual cells during the cycling trials. The graphs had a dual purpose: providing a clear visual assessment of degradation trends and establishing the groundwork for a more comprehensive and analytical comparison

of performance among other cells. The meticulous methodology employed in data visualisation played a crucial role in uncovering significant revelations on the degradation of battery performance over time. This yielded useful insights for comprehending and perhaps enhancing battery technology.

Here is an example specifically for understanding the discharge voltage behavior of Cell 01 and Cell 25. How it undergoes to a behaviour in proportional to time.

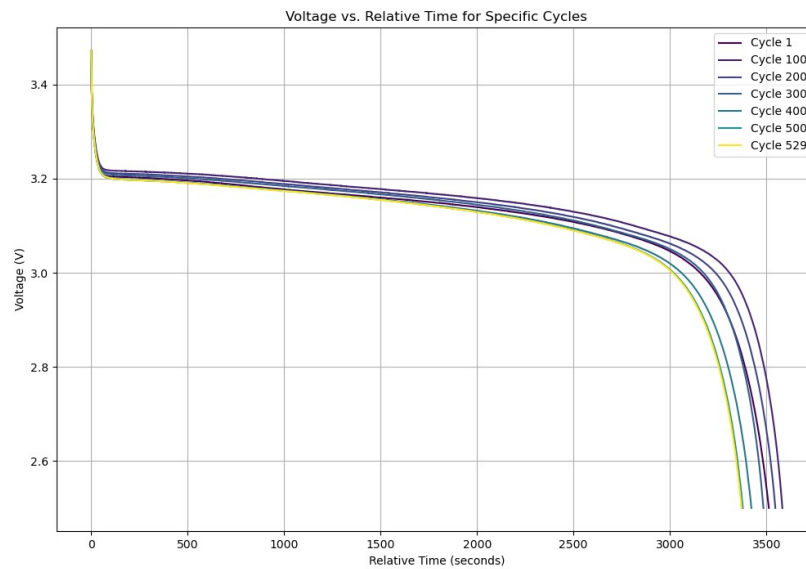


Figure 28: Cell 01

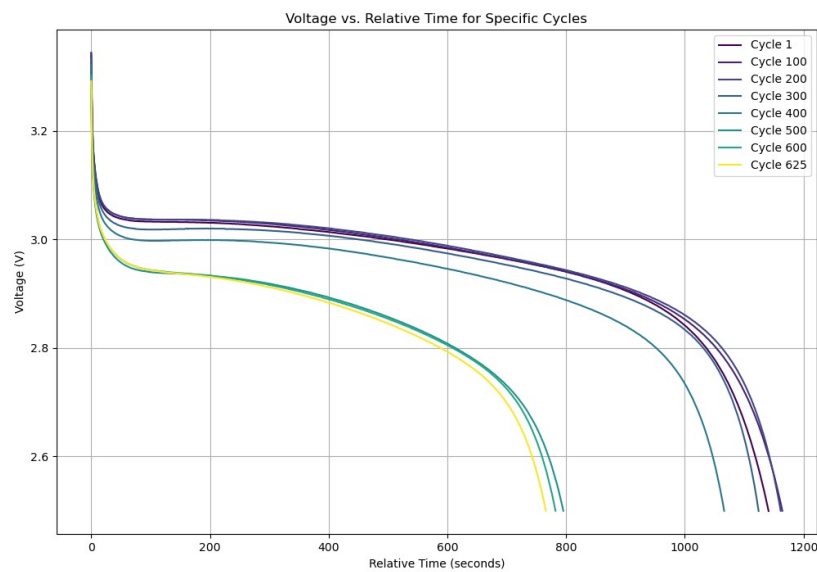


Figure 29: Cell 25

5.2. Model Implementation

This section focuses on the implementation details and outcomes of our research utilising three different models to estimate the State of Health SOH of batteries: the Long Short-Term Memory LSTM Model, the Gated Recurrent Unit GRU Model, and the Regression Model.

5.2.1. LSTM Model

Our research primarily aimed to forecast the State of Health SOH of batteries by utilising an LSTM (Long Short-Term Memory) model, which is a type of Recurrent Neural Network. The approach we employed encompassed:

Optimising Learning Rates: We conducted experiments with various learning rates to improve the model's learning efficiency and performance.

Exploring Different Activation Functions: In particular, the sigmoid and tanh functions were examined in order to enhance the accuracy of SOH predictions.

Adjusting the Number of Hidden Layers: This allowed us to identify the optimal model architecture, striking a harmonious balance between complexity and learning capacity.

During the entire procedure, we continuously monitored the model's performance, thereby preventing overfitting and assuring dependable learning. The optimal model configuration, determined through thorough experimentation, entailed training for a total of 1000 epochs. This configuration yielded the utmost accuracy in calculating SOH.

Additional validation was performed on a separate dataset, which proved the model's capacity to apply to fresh data. The model achieved the following performance metrics on the test dataset: MAE, MSE, RMSE and R-squared (Coefficient of Determination) (R²). During the evaluation, the RMSE for the test dataset was calculated, and the results revealed that the LSTM had a mean RMSE of 0.0896.

Train Cells	Test Cells	MAE	MSE	RMSE	R^2
[2,3]	1	0.011705088	0.000265847	0.016304817	0.903198700
[1,3]	2	0.047639822	0.010119943	0.100597926	0.696146680
[1,2]	3	0.017572331	0.000855068	0.029241543	0.686601076
[5,6]	4	0.015811831	0.000364690	0.019098665	0.922664997
[4,6]	5	0.012040499	0.001532602	0.039148462	0.771259021
[4,5]	6	0.035698464	0.002668902	0.051661419	0.537625133
[8,9]	7	0.314061283	0.137998895	0.371482026	-0.897261660
[7,9]	8	0.066355687	0.009098701	0.095387110	0.912041946
[7,8]	9	0.128806238	0.026127578	0.161642075	0.640511410
[11,12]	10	0.084407611	0.015050414	0.122680127	0.826728163
[10,12]	11	0.024937193	0.001241111	0.035229407	0.985705255
[10,11]	12	0.051666264	0.006176525	0.078590869	0.941595947
[14,15]	13	0.137220526	0.034085341	0.184622158	0.442535299
[13,15]	14	0.057933271	0.008147129	0.090261450	0.911369059
[13,14]	15	0.097966363	0.016151324	0.127078761	0.849168348
[23]	24	0.027544456	0.001284119	0.035834611	0.704191198
[25,26]	27	0.009314132	0.000240842	0.015519085	0.713026930
[26,27]	25	0.056921515	0.012070852	0.109867429	0.887303483
[27,25]	26	0.012529988	0.000342394	0.018503892	0.951855082

Table 3: LSTM Model Result

5.2.2. GRU Model

Our investigations used both the LSTM Model and the GRU Model, a recurrent neural network variation, for SOH estimation. Implementing the GRU model involves comparable considerations as the LSTM model. To optimize performance, we conducted trials with various learning rates and activation functions. By experimenting with different hidden layer counts, we determined the optimal model depth. Similar to the LSTM Model, we closely watched the GRU Model's training performance, observing consistent estimation mistakes without overfitting. After thorough study, we identified the most effective GRU configuration.

In Figure 29, the model achieved the following performance metrics in the test data set: MAE, MSE, RMSE, and R-squared (Coefficient of Determination). The model's generalization ability was tested on a different dataset, demonstrating accurate SOH estimation for unknown battery cells with a competitive RMSE. Evaluation findings showed a mean RMSE of 0.8866 for the GRU in the test dataset.

Train Cells	Test Cell	MAE	MSE	RMSE	R^2
[2, 3]	1	0.238085428	0.099259158	0.315054214	0.952187582
[1, 3]	2	0.449058533	0.971038041	0.985412625	0.756601125
[1, 2]	3	1.275124483	4.832380490	2.196196824	-1.353276704
[5, 6]	4	0.239488475	0.102715896	0.320493207	0.958434098
[4, 6]	5	0.226213585	0.620596472	0.787779456	0.782970405
[4, 5]	6	0.500841147	0.817013146	0.903887795	0.706497567
[8, 9]	7	0.917731171	1.154596449	1.074521498	0.130017536
[7, 9]	8	0.293227180	0.177784874	0.421645437	0.876962415
[7, 8]	9	0.511323700	0.414962346	0.644175171	0.687087219
[11, 12]	10	0.224528466	0.110127242	0.331854248	0.945119972
[10, 12]	11	0.102927978	0.022841945	0.151135519	0.986812225
[10, 11]	12	0.229401742	0.108015447	0.328657036	0.950490543
[14, 15]	13	0.560821917	0.520511117	0.721464564	0.462615948
[13, 15]	14	0.239321163	0.153709582	0.392058137	0.894443076
[13, 14]	15	0.447688320	0.313397328	0.559819013	0.807546684
[23]	24	0.358074977	0.259238749	0.509154936	0.930991290
[25, 26]	27	0.514439001	0.494065119	0.702897659	0.931660483
[26, 27]	25	1.233073498	14.888784601	3.858598787	0.909436952
[27, 25]	26	0.938728022	2.688892773	1.639734868	0.956108890

Table 4: GRU Model Result

5.2.3. Regression Model

To augment our recurrent neural network approach, we also developed a Regression Model for estimating SOH. This approach enabled us to delve into various regression techniques, such as Linear regression, Polynomial regression, Regression with decision trees, and Random forest regression. Additionally, we employed feature engineering strategies to refine the input features, thereby improving the Regression Model's precision in SOH prediction. Similar to our other models, this one underwent extensive training and validation to evaluate its effectiveness.

Train Cells	Cell	MAE	MSE	RMSE	R^2
[2, 3]	1	0.535928828	0.355938885	0.596660139	0.844199796
[1, 3]	2	0.55188521	1.504879938	1.226735480	0.618026828
[1, 2]	3	0.561883937	1.498163259	1.223994795	0.460869293
[5, 6]	4	0.257916326	0.147528327	0.384094164	0.932677503
[4, 6]	5	0.409104975	3.362550482	1.833752847	0.341491875
[4, 5]	6	0.398978535	2.961144410	1.720776909	0.373509138
[8, 9]	7	0.300391326	0.126114018	0.355152356	0.902430513
[7, 9]	8	0.272264725	0.113018129	0.336186189	0.923608570
[7, 8]	9	0.299819132	0.125837237	0.354735446	0.902592054
[11, 12]	10	0.311906124	0.166598007	0.408164269	0.920823669
[10, 12]	11	0.311183154	0.165973124	0.407399793	0.921050228
[10, 11]	12	0.300380063	0.160714213	0.400891772	0.929501984
[14, 15]	13	0.421253423	0.246648525	0.496637216	0.767898942
[13, 15]	14	0.331158653	0.162306833	0.402940235	0.885467076
[13, 14]	15	0.412809247	0.246998707	0.496989645	0.857444109
[23]	24	0.975903984	1.798043205	1.340911334	0.563011800
[25, 26]	25	5.119517737	32.81997052	5.728871662	0.794935289
[26, 27]	26	3.243908364	17.113854884	4.136889518	0.713661434
[27, 25]	27	1.366522283	2.2540622420	1.501353470	0.738751676

Table 5: Linear Regression Model Result

Train Cells	Cell	MAE	MSE	RMSE	R^2
[2, 3]	1	0.205569759	0.077508538	0.278403552	0.96607326
[1, 3]	2	0.326281534	1.318692874	1.148354886	0.658974314
[1, 2]	3	0.290988550	1.229121305	1.108657434	0.553785031
[5, 6]	4	0.260738021	0.149172736	0.386228865	0.931927001
[4, 6]	5	0.411862717	3.362304918	1.833658888	0.341539965
[4, 5]	6	0.401742040	2.961026952	1.720763479	0.373533989
[8, 9]	7	0.254546512	0.092390615	0.303958246	0.928520992
[7, 9]	8	0.254345189	0.095064471	0.308325628	0.935743841
[7, 8]	9	0.254318918	0.092159719	0.303578193	0.928661108
[11, 12]	10	0.28998915	0.139811044	0.373913151	0.933554299
[10, 12]	11	0.289730351	0.139501868	0.373499489	0.933642024
[10, 11]	12	0.288026634	0.140842124	0.375289837	0.938218966
[14, 15]	13	0.210109048	0.106182931	0.325857226	0.900079797
[13, 15]	14	0.201312200	0.085320424	0.292069601	0.939318085
[13, 14]	15	0.266856335	0.133991504	0.366046850	0.922666485
[23]	24	0.668170991	1.428292723	1.195112013	0.652874266
[25, 26]	25	2.269723412	10.345779187	3.214685533	0.935357827
[26, 27]	26	2.279664302	10.522807424	3.243887702	0.823938814
[27, 25]	27	0.485051366	0.359617759	0.599681381	0.958319901

Table 6: Polynomial Regression Model Result

Train Cells	Cell	MAE	MSE	RMSE	R^2
[2, 3]	1	0.071839435	0.02863872	0.169229783	0.987464369
[1, 3]	2	0.181575372	1.23772477	1.112530798	0.679914694
[1, 2]	3	0.179380297	1.167566511	1.080619251	0.580126421
[5, 6]	4	0.087103296	0.256444702	0.504427418	0.883886647
[4, 6]	5	0.238312742	3.615206064	1.901369523	0.292012840
[4, 5]	6	0.227953421	3.203608149	1.789862606	0.322210959
[8, 9]	7	0.002626341	0.000119869	0.010948487	0.999907262
[7, 9]	8	0.005920331	0.002737523	0.052331342	0.998149648
[7, 8]	9	0.002303299	0.000121785	0.011036563	0.999905729
[11, 12]	10	0.002983195	1.837796755	0.001428686	0.999991266
[10, 12]	11	0.002511783	1.334567345	0.003651019	0.999993659
[10, 11]	12	0.004249141	0.000408867	0.020220464	0.999820649
[14, 15]	13	0.095085034	0.068830825	0.262304832	0.935254263
[13, 15]	14	0.103193851	0.036952768	0.192231028	0.973339327
[13, 14]	15	0.133209029	0.067617825	0.260034277	0.960974211
[23]	24	0.108824488	0.422244667	0.649803561	0.897379586
[25, 26]	25	0.203886913	0.236361689	0.486170432	0.998523172
[26, 27]	26	0.149926631	0.089021677	0.298365006	0.998510544
[27, 25]	27	0.086192996	0.049160748	0.22172232	0.994302215

Table 7: Decision Tree Regression Model Result.

Train Cells	Cell	MAE	MSE	RMSE	R^2
[2, 3]	1	0.06351235	0.016962359	0.130329623	0.992575301
[1, 3]	2	0.235617994	1.421667086	1.193219124	0.632356015
[1, 2]	3	0.175659557	1.160437424	1.07961741	0.58089683
[5, 6]	4	0.065432447	0.112198293	0.334960136	0.948799872
[4, 6]	5	0.216600927	3.424775429	1.850614879	0.329305996
[4, 5]	6	0.206209142	3.016237446	1.737631289	0.361853076
[8, 9]	7	0.005142591	0.000655866	0.025609881	0.999492582
[7, 9]	8	0.006703569	0.001256866	0.035452306	0.99910547
[7, 8]	9	0.004980341	0.000667628	0.025838506	0.999483203
[11, 12]	10	0.00420677	0.916100005	0.008897225	0.999962379
[10, 12]	11	0.003718821	0.7213e005	0.008780709	0.999963271
[10, 11]	12	0.004249141	0.000721707	0.026866461	0.999608342
[14, 15]	13	0.008043257	0.042545245	0.206249488	0.959964097
[13, 15]	14	0.095098394	0.026237174	0.16107983	0.981491717
[13, 14]	15	0.112255765	0.044362926	0.210652084	0.974395832
[23]	24	0.170325659	0.406764704	0.637781078	0.901141766
[25, 26]	25	0.308150381	2.469212483	1.571737202	0.984571944
[26, 27]	26	0.149596505	0.090606992	0.301009953	0.998484019
[27, 25]	27	0.073173988	0.028147822	0.167773126	0.996737636

Table 8: Random Forest Regression Model Result.

The experimental results revealed that the Regression Model used for estimating the SOH exhibited different performance metrics depending on the specific regression technique used. The metrics encompassed in this analysis are Mean Absolute Error MAE, Mean Squared Error MSE, Root Mean Squared Error RMSE, and the Coefficient of Determination (R-squared, R^2). The findings for each regression procedure are given in a systematic manner in Figures 30 (Linear Regression), 31 (Polynomial Regression), 32 (Decision Tree Regression), and 33 (Random Forest Regression). The diverse performance observed across these models highlights the versatility and possible practicality of the Regression Model in real-world State of Health SOH estimate settings.

Consistent with our methodology for the LSTM and GRU Models, we conducted a thorough assessment of the Regression Model's ability to generalise. The assessment was performed on a separate test dataset, guaranteeing an unbiased evaluation of the model's ability to handle novel data.

The RMSE values were of special importance during this evaluation phase. The results showed that Linear Regression had a RMSE of 1.229, Polynomial Regression had a mean RMSE of 0.934, Decision Tree Regression had a mean RMSE of 0.475, and Random Forest Regression had a mean RMSE of 0.511. The results yielded a thorough comprehension of the predicted accuracy and resilience of each algorithm.

To summarise, this research involved a thorough investigation of three separate prediction models: the LSTM Model, the GRU Model, and the Regression Model. Every model was rigorously designed with precise architectural concerns and underwent rigorous training and validation procedures. The combined knowledge gained from these models greatly enhances the developing field of State of Health SOH estimation for battery health monitoring. Together, they exhibit impressive precision and ability to apply to many situations, making them well-suited for real-world use. However, it is crucial to perform additional research to assess and enhance the computational efficiency and scalability of these models, particularly in light of their potential implementation in diverse and ever-changing real-world situations. Ensuring their practical utility in the broader context of battery health monitoring systems is of utmost importance.

6. CONCLUSION AND FUTURE WORK

6.1. Conclusion

In addressing the vital need for long-lasting, efficient, and cost-effective battery systems, the significance of accurate prediction and evaluation techniques for the SOH of these systems is paramount. A key challenge in this area is accurately determining the SOH of lithium-ion batteries LIB, particularly given the scarcity of extensive long-term data, especially for batteries undergoing numerous cycles with limited aging information. This thesis addresses this challenge by developing a predictive model for assessing the health of LIB, utilizing data derived from battery cycling experiments. A focal point of this research is the model's ability to predict the health of battery cells not included in the training data set.

Throughout this research, we employed various data-driven models, subjecting them to extensive testing. The findings demonstrate that LSTM networks are particularly effective in estimating the SOH of batteries. Remarkably, the model maintained an accuracy level above 98% in predicting the health status of the batteries tested, highlighting its efficiency and reliability. However, it's crucial to note that reducing the sample size needed for training these models would require more robust testing methodologies. Moreover, the computational demands of these models pose a significant challenge to their real-world application.

To sum up, this study makes a substantial contribution to the field of battery maintenance by offering a dependable forecasting model for determining the SOH of LIB. While the outcomes of this research are encouraging, it's important to recognize its limitations, such as the necessity for broader testing and enhancing computational efficiency. Additionally, this research sheds light on the impact of temperature on battery performance, presenting new opportunities for future investigations and practical implementations in battery health management. This finding not only adds to the depth of our understanding but also opens new possibilities for optimizing battery usage and longevity in various applications.

6.2. Future Work

The findings from our investigations highlight the presence of discrepancies in data gathering, attributed to the nuances of the experimental setup. To ensure the integrity of future research, it is crucial to establish protocols that will significantly reduce these inaccuracies. Our analysis also revealed that health indicators derived during the study show promising capabilities in mirroring the wear and tear of batteries, thus improving the precision of SOH estimations. An important avenue for advancing our understanding of battery health involves adapting our methodologies to accommodate instances where complete charging data may not be readily available, due to a variety of realistic constraints. Crafting strategies to estimate battery health with partial data sets is therefore an essential step forward.

While the LSTM model has proven its merit in forecasting the progression of battery health, its applicability is somewhat constrained by its performance over batteries with extensive cycle histories. For batteries subjected to numerous cycles, incorporating mathematical modeling techniques to trace the trajectory of capacity fading could offer a more nuanced understanding. It's also pertinent to mention that the current scope of the LSTM model is primarily centered on estimating a battery's lifespan rather than providing an all-encompassing evaluation of its health or its appropriateness for certain uses. Future research should delve into categorizing and assessing batteries across different stages of health. This broader approach will likely yield insights that are vital for optimizing battery management strategies and maximizing their application potential.

LITERATURE REFERENCES

- [1] A. Hariprasad , I. Priyanka , R. Sandeep , V. Ravi, O. Shekar. Battery management system in electric vehicles, 2020. URL <https://www.ijert.org/battery-management-system-in-electric-vehicles>. Last updated: 05-05-2020 Accessed on: 30.12.2023.
- [2] A. Accardo, G. Dotelli, M. L. Musa, and E. Spessa. Life cycle assessment of an nmc battery for application to electric light-duty commercial vehicles and comparison with a sodium-nickel-chloride battery. *Applied Sciences*, 11(3), 2021. doi: 10.3390/app11031160. URL <https://www.mdpi.com/2076-3417/11/3/1160>.
- [3] M. Al-Gabalawy, N. S. Hosny, and S. A. Hussien. Lithium-ion battery modeling including degradation based on single-particle approximations. *Batteries*, 6(3), 2020. doi: 10.3390/batteries6030037. URL <https://www.mdpi.com/2313-0105/6/3/37>.
- [4] A. Allam and S. Onori. Online capacity estimation for lithium-ion battery cells via an electrochemical model-based adaptive interconnected observer. *IEEE Transactions on Control Systems Technology*, 29(4):1636–1651, 2021. doi: 10.1109/TCST.2020.3017566.
- [5] M. L. Aston Zhang, Zachary Lipton and A. Smola. Dive into deep learning, 2020. URL <https://alex.smola.org/projects.html>. Last updated: 24-09-2022 Accessed on: 30.12.2023.
- [6] T. O. Ayodele. Machine learning overview. In Y. Zhang, editor, *New Advances in Machine Learning*, chapter 2. IntechOpen, Rijeka, 2010. doi: 10.5772/9374. URL [10.5772/9374](https://doi.org/10.5772/9374).
- [7] A. Barré, B. Deguilhem, S. Grolleau, M. Gérard, F. Suard, and D. Riu. A review on lithium-ion battery ageing mechanisms and estimations for automotive applications. *Journal of Power Sources*, 241:680–689, 2013. ISSN 0378-7753. doi: 10.1016/j.jpowsour.2013.05.040. URL <https://www.sciencedirect.com/science/article/pii/S0378775313008185>.
- [8] Battery University. Types of lithium-ion batteries, 2023. URL <https://batteryuniversity.com/article/bu-205-types-of-lithium-ion>. Last updated: 08.12.2023 Accessed on: 30.12.2023.
- [9] A. Berrueta, J. Pascual, I. S. Martín, P. Sanchis, and A. Ursúa. Influence of the aging model of lithium-ion batteries on the management of pv self-consumption systems. In *2018 IEEE International Conference on Environment and Electrical Engineering and 2018 IEEE Industrial and Commercial Power Systems Europe (EEEIC / I&CPS Europe)*, pages 1–5, 2018. doi: 10.1109/EEEIC.2018.8493778.
- [10] T. Chai and R. R. Draxler. Root mean square error (rmse) or mean absolute error (mae)? – arguments against avoiding rmse in the literature. *Geoscientific Model Development*, 7(3):1247–1250, 2014. doi: 10.5194/gmd-7-1247-2014. URL <https://gmd.copernicus.org/articles/7/1247/2014/>.

-
- [11] J. Cho, T.-J. Kim, Y. Kimb, and B. Parkb. Complete blocking of mn^{3+} ion dissolution from a limn_2o_4 spinel intercalation compound by co_3o_4 coating. *Chemical Communications - CHEM COMMUN*, pages 1074–1075, 06 2001. doi: 10.1039/b101677f.
- [12] K. Cho, B. van Merriënboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk, and Y. Bengio. Learning phrase representations using rnn encoder-decoder for statistical machine translation, 2014.
- [13] M. U. Cuma and T. Koroglu. A comprehensive review on estimation strategies used in hybrid and battery electric vehicles. *Renewable and Sustainable Energy Reviews*, 42(C):517–531, 2015. doi: 10.1016/j.rser.2014.10.04. URL <https://ideas.repec.org/a/eee/rensus/v42y2015icp517-531.html>.
- [14] H. Dai, X. Zhang, W. Gu, X. Wei, and Z. Sun. A semi-empirical capacity degradation model of ev li-ion batteries based on eyring equation. In *2013 IEEE Vehicle Power and Propulsion Conference (VPPC)*, pages 1–5, 2013. doi: 10.1109/VPPC.2013.6671660.
- [15] DataTechNotes. Regression model accuracy (mae, mse, rmse, r-squared) check in r, 2019. URL <https://www.datatechnotes.com/2019/02/regression-model-accuracy-mae-mse-rmse.html>. Last updated: 01-02-2019 Accessed on: 30.12.2023.
- [16] Y. Ding, Z. P. Cano, A. Yu, J. Lu, and Z. Chen. Automotive li-ion batteries: Current status and future perspectives. *Electrochemical Energy Reviews*, 2(1), 1 2019. ISSN 2520-8489. doi: 10.1007/s41918-018-0022-z. URL <https://www.osti.gov/biblio/1561559>.
- [17] N. Donges. A guide to recurrent neural networks: Understanding rnn and lstm networks, 2019. URL <https://builtin.com/data-science/recurrent-neural-networks-and-lstm>. Last updated: 12-10-2019 Accessed on: 30.12.2023.
- [18] M. El-Dalahmeh, J. Lillystone, M. Al-Greer, and M. El-Dalahmeh. State of health estimation of lithium-ion batteries based on data-driven techniques. In *2021 56th International Universities Power Engineering Conference (UPEC)*, pages 1–6, 2021. doi: 10.1109/UPEC50034.2021.9548209.
- [19] S. Farhad and A. Nazari. Introducing the energy efficiency map of lithium-ion batteries. *International Journal of Energy Research*, 43, 12 2018. doi: 10.1002/er.4332.
- [20] A. G. A review of dropout as applied to rnns, 2018. URL <https://adriangcoder.medium.com/a-review-of-dropout-as-applied-to-rnns-72e79ecd5b7b>. Last updated: 22-07-2018 Accessed on: 30.12.2023.
- [21] T. Gao and W. Lu. Machine learning toward advanced energy storage devices and systems. *iScience*, 24(1):101936, 2021. ISSN 2589-0042. doi: 10.1016/j.isci.2020.101936. URL <https://www.sciencedirect.com/science/article/pii/S2589004220311330>.
- [22] Global EV Outlook 2023 Catching up with climate ambition. Battery demand for evs continues to rise, 2008. URL <https://www.iea.org/reports/global-ev-outlook-2023>. Last updated: 12.2008 Accessed on: 30.12.2023.

-
- [23] J. B. Goodenough and K.-S. Park. The li-ion rechargeable battery: A perspective. *Journal of the American Chemical Society*, 135(4):1167–1176, 2013. doi: 10.1021/ja3091438.
- [24] J. S. Goud, R. Kalpana, and B. Singh. Modeling and estimation of remaining useful life of single cell li-ion battery. In *2018 IEEE International Conference on Power Electronics, Drives and Energy Systems (PEDES)*, pages 1–5, 2018. doi: 10.1109/PEDES.2018.8707554.
- [25] W. Guo, Z. Sun, S. B. Vilsen, J. Meng, and D. I. Stroe. Review of “grey box” lifetime modeling for lithium-ion battery: Combining physics and data-driven methods. *Journal of Energy Storage*, 56:105992, 2022. ISSN 2352-152X. doi: 10.1016/j.est.2022.105992. URL <https://www.sciencedirect.com/science/article/pii/S2352152X22019806>.
- [26] X. Han, M. Ouyang, L. Lu, and J. Li. A comparative study of commercial lithium ion battery cycle life in electric vehicle: Capacity loss estimation. *Journal of Power Sources*, 268:658–669, 2014. ISSN 0378-7753. doi: 10.1016/j.jpowsour.2014.06.111. URL <https://www.sciencedirect.com/science/article/pii/S0378775314009756>.
- [27] M. H. S. M. Haram, J. W. Lee, G. Ramasamy, E. E. Ngu, S. P. Thiagarajah, and Y. H. Lee. Feasibility of utilising second life ev batteries: Applications, lifespan, economics, environmental impact, assessment, and challenges. *Alexandria Engineering Journal*, 60(5):4517–4536, 2021. ISSN 1110-0168. doi: 10.1016/j.aej.2021.03.021. URL <https://www.sciencedirect.com/science/article/pii/S1110016821001757>.
- [28] S. M. Hell and C. D. Kim. Development of a data-driven method for online battery remaining-useful-life prediction. *Batteries*, 8(10), 2022. doi: 10.3390/batteries8100192. URL <https://www.mdpi.com/2313-0105/8/10/192>.
- [29] J. Hemdani, L. Degaa, N. Rizoug, and A. Chaari. State of health prediction of lithium-ion battery using machine learning algorithms. In *2023 9th International Conference on Control, Decision and Information Technologies (CoDIT)*, pages 2729–2733, 2023. doi: 10.1109/CoDIT58514.2023.10284459.
- [30] G. E. Hinton, N. Srivastava, A. Krizhevsky, I. Sutskever, and R. R. Salakhutdinov. Improving neural networks by preventing co-adaptation of feature detectors, 2012.
- [31] S. Hochreiter and J. Schmidhuber. Long short-term memory. *Neural computation*, 9:1735–80, 12 1997. doi: 10.1162/neco.1997.9.8.1735.
- [32] J. Hong, Z. Wang, W. Chen, L. Wang, P. Lin, and C. Qu. Online accurate state of health estimation for battery systems on real-world electric vehicles with variable driving conditions considered. *Journal of Cleaner Production*, 294:125814, 2021. ISSN 0959-6526. doi: 10.1016/j.jclepro.2021.125814. URL <https://www.sciencedirect.com/science/article/pii/S0959652621000342>.
- [33] R. J. Hyndman and A. B. Koehler. Another look at measures of forecast accuracy. *International Journal of Forecasting*, 22(4):679–688, 2006. ISSN 0169-2070. doi:

- 10.1016/j.ijforecast.2006.03.001. URL <https://www.sciencedirect.com/science/article/pii/S0169207006000239>.
- [34] A. Jain, J. Mao, and K. Mohiuddin. Artificial neural networks: a tutorial. *Computer*, 29(3):31–44, 1996. doi: 10.1109/2.485891.
- [35] J. Jia, J. Liang, Y. Shi, J. Wen, X. Pang, and J. Zeng. Soh and rul prediction of lithium-ion batteries based on gaussian process regression with indirect health indicators. *Energies*, 13(2), 2020. doi: 10.3390/en13020375. URL <https://www.mdpi.com/1996-1073/13/2/375>.
- [36] S. Jin, X. Sui, X. Huang, S. Wang, R. Teodorescu, and D.-I. Stroe. Overview of machine learning methods for lithium-ion battery remaining useful lifetime prediction. *Electronics*, 10(24), 2021. doi: 10.3390/electronics10243126. URL <https://www.mdpi.com/2079-9292/10/24/3126>.
- [37] S. K. JOHNSON. Here’s what tesla will put in its new batteries, 2020. URL <https://arstechnica.com/cars/2020/09/heres-what-tesla-will-put-in-its-new-batteries/>. Last updated: 24-09-2020 Accessed on: 30.12.2023.
- [38] K. Li and K. J. Tseng. Energy efficiency of lithium-ion battery used as energy storage devices in micro-grid. In *IECON 2015 - 41st Annual Conference of the IEEE Industrial Electronics Society*, pages 005235–005240, 2015. doi: 10.1109/IECON.2015.7392923.
- [39] Q. Li. Technological evolution of lithium batteries for new energy vehicles. In *2022 International Conference on Industrial IoT, Big Data and Supply Chain (IIoTBDS)*, pages 19–23, 2022. doi: 10.1109/IIoTBDS57192.2022.00015.
- [40] Y. Li, K. Liu, A. M. Foley, A. Zülke, M. Bercibar, E. Nanini-Maury, J. Van Mierlo, and H. E. Hoster. Data-driven health estimation and lifetime prediction of lithium-ion batteries: A review. *Renewable and Sustainable Energy Reviews*, 113:109254, 2019. ISSN 1364-0321. doi: 10.1016/j.rser.2019.109254. URL <https://www.sciencedirect.com/science/article/pii/S136403211930454X>.
- [41] Y. Li, K. Liu, A. M. Foley, A. Zülke, M. Bercibar, E. Nanini-Maury, J. Van Mierlo, and H. E. Hoster. Data-driven health estimation and lifetime prediction of lithium-ion batteries: A review. *Renewable and Sustainable Energy Reviews*, 113:109254, 2019. ISSN 1364-0321. doi: 10.1016/j.rser.2019.109254. URL <https://www.sciencedirect.com/science/article/pii/S136403211930454X>.
- [42] A. Lidbeck and K. Syed. Experimental characterization of li-ion battery cells for thermal management in heavy duty hybrid applications. Master’s thesis, Chalmers University of Technology, Sweden, 2017. URL <https://publications.lib.chalmers.se/records/fulltext/252994/252994.pdf>.
- [43] J. Lin, G. Yan, and C. Wang. Li-ion battery state of health prediction based on long short-term memory recurrent neural network. *Journal of Physics: Conference Series*, 2010(1):012133, sep 2021. doi: 10.1088/1742-6596/2010/1/012133. URL <https://dx.doi.org/10.1088/1742-6596/2010/1/012133>.

-
- [44] G. Loye. Gated recurrent unit (gru) with pytorch, 2019. URL <https://blog.floydhub.com/gru-with-pytorch/>. Last updated: 22-07-2019 Accessed on: 30.12.2023.
- [45] B. . Mahesh. Machine learning algorithms, 2020. URL <https://www.scirp.org/reference/referencespapers?referenceid=3168174>. Last updated: 24-09-2020 Accessed on: 30.12.2023.
- [46] Matlab Channel Youtube. Estimating remaining useful life (rul) | predictive maintenance, 2020. URL https://www.youtube.com/watch?v=Dd_4rbWYgI4&t=202s. Last updated: 01-01-2020 Accessed on: 30.12.2023.
- [47] H. Meng and Y.-F. Li. A review on prognostics and health management (phm) methods of lithium-ion batteries. *Renewable and Sustainable Energy Reviews*, 116: 109405, 2019. ISSN 1364-0321. doi: 10.1016/j.rser.2019.109405. URL <https://www.sciencedirect.com/science/article/pii/S1364032119306136>.
- [48] Y. Miao, P. Hynan, A. von Jouanne, and A. Yokochi. Current li-ion battery technologies in electric vehicles and opportunities for advancements. *Energies*, 12:1074–1094, 03 2019. doi: 10.3390/en12061074.
- [49] MIT Electric Vehicle Team. A guide to understanding battery specifications, 2008. URL http://web.mit.edu/evt/summary_battery_specifications.pdf. Last updated: 12.2008 Accessed on: 30.12.2023.
- [50] A. Mittal. Understanding rnn and lstm, 2019. URL <https://aditi-mittal.medium.com/understanding-rnn-and-lstm-f7cdf6dfc14e>. Last updated: 12-10-2019 Accessed on: 30.12.2023.
- [51] X. Nansi. *Design and Optimization of Lithium-Ion Batteries for Electric-Vehicle Applications*. PhD thesis, University of Michigan, 2014.
- [52] C. Olah. Understanding lstm networks, 2015. URL <https://colah.github.io/posts/2015-08-Understanding-LSTMs/>. Last updated: 27-08-2015 Accessed on: 30.12.2023.
- [53] S. Paul, C. Diegelmann, H. Kabza, and W. Tillmetz. Analysis of ageing inhomogeneities in lithium-ion battery systems. *Journal of Power Sources*, 239:642–650, 2013. ISSN 0378-7753. doi: 10.1016/j.jpowsour.2013.01.068. URL <https://www.sciencedirect.com/science/article/pii/S037877531300116X>.
- [54] J. Peng, E. Jury, P. Dönnies, and C. Ciurtin. Machine learning techniques for personalised medicine approaches in immune-mediated chronic inflammatory diseases: Applications and challenges. *Frontiers in Pharmacology*, 12, 09 2021. doi: 10.3389/fphar.2021.720694.
- [55] H. Popp, M. Koller, M. Jahn, and A. Bergmann. Mechanical methods for state determination of lithium-ion secondary batteries: A review. *Journal of Energy Storage*, 32:101859, 2020. ISSN 2352-152X. doi: 10.1016/j.est.2020.101859. URL <https://www.sciencedirect.com/science/article/pii/S2352152X20316960>.

-
- [56] R. Pugliese, S. Regondi, and R. Marini. Machine learning-based approach: global trends, research directions, and regulatory standpoints. *Data Science and Management*, 4:19–29, 2021. ISSN 2666-7649. doi: 10.1016/j.dsm.2021.12.002. URL <https://www.sciencedirect.com/science/article/pii/S2666764921000485>.
- [57] J. Qu, F. Liu, Y. Ma, and J. Fan. A neural-network-based method for rul prediction and soh monitoring of lithium-ion battery. *IEEE Access*, 7:87178–87191, 2019. doi: 10.1109/ACCESS.2019.2925468.
- [58] A. Raj. Battery management system (bms) for electric vehicles, 2018. URL <https://circuitdigest.com/article/battery-management-system-bms-for-electric-vehicles>. Last updated: 05-12-2018 Accessed on: 30.12.2023.
- [59] K. Rao. How are inputs fed into a lstm recurrent neural network, 2023. URL <https://www.quora.com/How-can-I-determinate-the-number-of-connections-the-weight-matrix-in-an-LSTM-network-for-each-gate-and-each-input-of-each-LSTM-bloc>. Last updated: 31-10-2023 Accessed on: 30.12.2023.
- [60] S. K. Rechkemmer, X. Zang, W. Zhang, and O. Sawodny. Empirical li-ion aging model derived from single particle model. *Journal of Energy Storage*, 21:773–786, 2019. ISSN 2352-152X. doi: 10.1016/j.est.2019.01.005. URL <https://www.sciencedirect.com/science/article/pii/S2352152X18307588>.
- [61] Z. Ren and C. Du. A review of machine learning state-of-charge and state-of-health estimation algorithms for lithium-ion batteries. *Energy Reports*, 9:2993–3021, 2023. ISSN 2352-4847. doi: 10.1016/j.egy.2023.01.108. URL <https://www.sciencedirect.com/science/article/pii/S235248472300118X>.
- [62] M. Safari, M. Morcrette, A. Teysot, and C. Delacourt. Multimodal physics-based aging model for life prediction of li-ion batteries. *Journal of The Electrochemical Society - J ELECTROCHEM SOC*, 156, 01 2009. doi: 10.1149/1.3043429.
- [63] J. A. Sanguesa, V. Torres-Sanz, P. Garrido, F. J. Martinez, and J. M. Marquez-Barja. A review on electric vehicles: Technologies and challenges. *Smart Cities*, 4 (1):372–404, 2021. ISSN 2624-6511. doi: 10.3390/smartcities4010022. URL <https://www.mdpi.com/2624-6511/4/1/22>.
- [64] R. Schröder, M. Aydemir, and G. Seliger. Comparatively assessing different shapes of lithium-ion battery cells. *Procedia Manufacturing*, 8:104–111, 2017. ISSN 2351-9789. doi: 10.1016/j.promfg.2017.02.013. URL <https://www.sciencedirect.com/science/article/pii/S2351978917300173>. 14th Global Conference on Sustainable Manufacturing, GCSM 3-5 October 2016, Stellenbosch, South Africa.
- [65] Y. Shang, G. Lu, Y. Kang, Z. Zhou, B. Duan, and C. Zhang. A multi-fault diagnosis method based on modified sample entropy for lithium-ion battery strings. *Journal of Power Sources*, 446:227275, 2020. ISSN 0378-7753. doi: 10.1016/j.jpowsour.2019.227275. URL <https://www.sciencedirect.com/science/article/pii/S0378775319312686>.

-
- [66] M. Shi, X. Shi, Z. Li, X. Wang, S. Ren, and F. Di. Study on the aging characteristics of li-ion battery based on the electro-thermal and aging joint simulation platform. In *2020 IEEE/IAS Industrial and Commercial Power System Asia (ICPS Asia)*, pages 257–261, 2020. doi: 10.1109/ICPSAsia48933.2020.9208456.
- [67] A. Shrestha and A. Mahmood. Review of deep learning algorithms and architectures. *IEEE Access*, 7:53040–53065, 2019. doi: 10.1109/ACCESS.2019.2912200.
- [68] X. Shu, S. Shen, J. Shen, Y. Zhang, G. Li, Z. Chen, and Y. Liu. State of health prediction of lithium-ion batteries based on machine learning: Advances and perspectives. *iScience*, 24(11):103265, 2021. ISSN 2589-0042. doi: 10.1016/j.isci.2021.103265. URL <https://www.sciencedirect.com/science/article/pii/S2589004221012347>.
- [69] Y. Song, W. Liu, H. Li, Y. Zhou, Z. Huang, and F. Jiang. Robust and accurate state-of-charge estimation for lithium-ion batteries using generalized extended state observer. In *2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, pages 2146–2151, 2017. doi: 10.1109/SMC.2017.8122937.
- [70] Y. Song, Y. Peng, and D. Liu. Model-based health diagnosis for lithium-ion battery pack in space applications. *IEEE Transactions on Industrial Electronics*, 68(12):12375–12384, 2021. doi: 10.1109/TIE.2020.3045745.
- [71] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov. Dropout: a simple way to prevent neural networks from overfitting. *J. Mach. Learn. Res.*, 15(1):1929–1958, jan 2014. ISSN 1532-4435.
- [72] A.-I. Stroe, V. Knap, and D.-I. Stroe. Comparison of lithium-ion battery performance at beginning-of-life and end-of-life. *Microelectronics Reliability*, 88-90:1251–1255, 2018. ISSN 0026-2714. doi: 10.1016/j.microrel.2018.07.077. URL <https://www.sciencedirect.com/science/article/pii/S0026271418306292>. 29th European Symposium on Reliability of Electron Devices, Failure Physics and Analysis (ESREF 2018).
- [73] M. Theiler, D. Schneider, and C. Endisch. Experimental investigation of state and parameter estimation within reconfigurable battery systems. *Batteries*, 9(3), 2023. ISSN 2313-0105. doi: 10.3390/batteries9030145. URL <https://www.mdpi.com/2313-0105/9/3/145>.
- [74] M.-K. Tran, M. Mathew, S. Janhunen, S. Panchal, K. Raahemifar, R. Fraser, and M. Fowler. A comprehensive equivalent circuit model for lithium-ion batteries, incorporating the effects of state of health, state of charge, and temperature on model parameters. *Journal of Energy Storage*, 43:103252, 2021. ISSN 2352-152X. doi: 10.1016/j.est.2021.103252. URL <https://www.sciencedirect.com/science/article/pii/S2352152X2100949X>.
- [75] V. V. Vardhan, S. K. Singh, R. k. Gatla, P. Sridhar, and S. Chatterjee. Choice of batteries and battery management for new generation electric vehicles. In *2022 International Conference on Augmented Intelligence and Sustainable Systems (ICAISS)*, pages 1279–1282, 2022. doi: 10.1109/ICAISS55157.2022.10010890.

-
- [76] M. Varini, P. E. Campana, and G. Lindbergh. A semi-empirical, electrochemistry-based model for li-ion battery performance prediction over lifetime. *Journal of Energy Storage*, 25:100819, 2019. ISSN 2352-152X. doi: 10.1016/j.est.2019.100819. URL <https://www.sciencedirect.com/science/article/pii/S2352152X19300568>.
- [77] W. Vermeer, G. R. Chandra Mouli, and P. Bauer. A comprehensive review on the characteristics and modeling of lithium-ion battery aging. *IEEE Transactions on Transportation Electrification*, 8(2):2205–2232, 2022. doi: 10.1109/TTE.2021.3138357.
- [78] M. Vinkhuyzen. Tesla switching to lfp batteries for standard range model 3 & model y, 2022. URL <https://cleantechnica.com/2021/08/21/tesla-switching-to-lfp-batteries-for-standard-range-model-3-model-y/>. Last updated: 05-01-2022 Accessed on: 30.12.2023.
- [79] W. Waag, C. Fleischer, and D. U. Sauer. Critical review of the methods for monitoring of lithium-ion batteries in electric and hybrid vehicles. *Journal of Power Sources*, 258:321–339, 2014. ISSN 0378-7753. doi: 10.1016/j.jpowsour.2014.02.064. URL <https://www.sciencedirect.com/science/article/pii/S0378775314002572>.
- [80] F.-K. Wang and T. Mamo. A hybrid model based on support vector regression and differential evolution for remaining useful lifetime prediction of lithium-ion batteries. *Journal of Power Sources*, 401:49–54, 2018. ISSN 0378-7753. doi: 10.1016/j.jpowsour.2018.08.073. URL <https://www.sciencedirect.com/science/article/pii/S0378775318309418>.
- [81] S. Wang, S. Jin, D. Deng, and C. Fernandez. A critical review of online battery remaining useful lifetime prediction methods. *Frontiers in Mechanical Engineering*, 7, 2021. doi: 10.3389/fmech.2021.719718. URL <https://www.frontiersin.org/articles/10.3389/fmech.2021.719718>.
- [82] Y. Wang, J. Tian, Z. Sun, L. Wang, R. Xu, M. Li, and Z. Chen. A comprehensive review of battery modeling and state estimation approaches for advanced battery management systems. *Renewable and Sustainable Energy Reviews*, 131: 110015, 2020. ISSN 1364-0321. doi: 10.1016/j.rser.2020.110015. URL <https://www.sciencedirect.com/science/article/pii/S1364032120303063>.
- [83] Y. Wang, J. Tian, Z. Sun, L. Wang, R. Xu, M. Li, and Z. Chen. A comprehensive review of battery modeling and state estimation approaches for advanced battery management systems. *Renewable and Sustainable Energy Reviews*, 131: 110015, 2020. ISSN 1364-0321. doi: 10.1016/j.rser.2020.110015. URL <https://www.sciencedirect.com/science/article/pii/S1364032120303063>.
- [84] L. Wu, K. Liu, and H. Pang. Evaluation and observability analysis of an improved reduced-order electrochemical model for lithium-ion battery. *Electrochimica Acta*, 368:137604, 2021. ISSN 0013-4686. doi: 10.1016/j.electacta.2020.137604. URL <https://www.sciencedirect.com/science/article/pii/S0013468620319976>.

-
- [85] D. Xiao, G. Fang, S. Liu, S. Yuan, R. Ahmed, S. Habibi, and A. Emadi. Reduced-coupling coestimation of soc and soh for lithium-ion batteries based on convex optimization. *IEEE Transactions on Power Electronics*, 35(11):12332–12346, 2020. doi: 10.1109/TPEL.2020.2984248.
- [86] R. Xiong, L. Li, Z. Li, Q. Yu, and H. Mu. An electrochemical model based degradation state identification method of lithium-ion battery for all-climate electric vehicles application. *Applied Energy*, 219:264–275, 2018. ISSN 0306-2619. doi: 10.1016/j.apenergy.2018.03.053. URL <https://www.sciencedirect.com/science/article/pii/S0306261918303829>.
- [87] R. Xiong, L. Li, and J. Tian. Towards a smarter battery management system: A critical review on battery state of health monitoring methods. *Journal of Power Sources*, 405:18–29, 2018. ISSN 0378-7753. doi: 10.1016/j.jpowsour.2018.10.019. URL <https://www.sciencedirect.com/science/article/pii/S037877531831111X>.
- [88] D. Yoo, J. Park, J. Moon, and C. Kim. Reliability-based design optimization for reducing the performance failure and maximizing the specific energy of lithium-ion batteries considering manufacturing uncertainty of porous electrodes. *Energies*, 14(19), 2021. ISSN 1996-1073. URL <https://www.mdpi.com/1996-1073/14/19/6100>.
- [89] Y. Zhang, R. Xiong, H. He, and M. G. Pecht. Long short-term memory recurrent neural network for remaining useful life prediction of lithium-ion batteries. *IEEE Transactions on Vehicular Technology*, 67(7):5695–5705, 2018. doi: 10.1109/TVT.2018.2805189.
- [90] J. Zhao and A. Burke. Electric vehicle batteries: Status and perspectives of data-driven diagnosis and prognosis. *Batteries*, 8:142, 09 2022. doi: 10.3390/batteries8100142.
- [91] J. Zhao, X. Feng, J. Wang, Y. Lian, M. Ouyang, and A. F. Burke. Battery fault diagnosis and failure prognosis for electric vehicles using spatio-temporal transformer networks. *Applied Energy*, 352:121949, 2023. ISSN 0306-2619. doi: 10.1016/j.apenergy.2023.121949. URL <https://www.sciencedirect.com/science/article/pii/S0306261923013132>.

A. APPENDIX

A.1. System Configuration

A.1.1. Hardware Specifications:

1. Processor: Intel Core i7 8th Generation (8 cores)
2. Operating System: Windows 11
3. RAM: 16 GB
4. Graphics Card: intel iRIS

A.1.2. Software utilized:

1. Jupyter Lab
2. Python

A.2. Neware Battery Testing System Specification



Manufacturer of Battery Testing Equipment
Technology Specification

BTS-5V6A		Battery Testing System
Model: BTS-4008T-5V6A-S1		
Description		Product Specification
AC Input		AC 220V±10% 50Hz
Power		425W
Resolution		AD: 16bit; DA: 16bit
Input Impedance		≥1MΩ
Voltage	Measuring Range	Charge: 25mV~5V Discharge: 25mV~5V
	Discharge Min Voltage	1.5V
	Accuracy	± 0.05% of FS
	Stability	± 0.05% of FS
Current	Range	Range1: 0.5mA~0.1A Range2: 0.1A~3A Range3: 3A~6A
	Accuracy	± 0.05% of FS
	Stability	± 0.05% of FS
	Constant Voltage stop current	Range1: 0.2mA; Range2: 6mA; Range3: 12mA
Power	Output Power Per Channel	30W
	Stability	± 0.1% of FS
Time	Rise Time	1ms (0~Full Range)
	Step Time	≤ (365*24) hour/step Time Format 00: 00: 00: 000 (h:min:s:ms)
Data Acquisition	Intervals	Time interval Δt: 100ms
		Voltage interval ΔU: 10mV
		Current interval ΔI: (Range1: 0.2mA; Range2: 6mA; Range3: 12mA)
Frequency	10Hz	
Charge	Mode of Operation	CCC, CVC, CC & CVC, CPC
	End Conditions	Voltage, Current, Test Time, Capacity
Discharge	Mode of Operation	CCD, CPD, CRD
	End Conditions	Voltage, Current, Test Time, Capacity
Pulse mode	Charge	CR mode、CP mode
	Discharge	CC mode、CP mode
	Min pulse width	500ms





Manufacturer of Battery Testing Equipment
Technology Specification

	Automated Switch	Automated switch from charge to discharge for each pulse
	End Condition	Voltage、Relative Time
DCIR	Support for the calculation of the point of the DCIR	
Cycle	Cycles	65535
	Steps	254
	Nested Function	Max three levels of loops
Protective Function	Software	<ul style="list-style-type: none"> ● Power-off data protection ● Off-line Operating ● User-defined safety(upper and lower)tolerance of current, voltage and delay time
Channel Features		Independent pairs of closed loop for constant current source and constant voltage source
Channels		Independent control
Detection and Sampling		4-wire Connecting
Noise Density		<85dB
Data Management		MYSQL Database
Communication Means		TCP/IP Protocol
Export Formats		EXCEL2003/2010, TXT, Graph
Communication Interface		Ethernet Port
Number of Channels Per Cabinet		8
Operating Environment		
Description		Product Specification
Operating Temperature		0℃~40℃
Storage Temperature		-10℃~50℃
Operating Humidity		≤70% RH
Storage Humidity		≤80% RH
Description		Product Specification
Battery Holder		Universal Holder



Manufacturer of Battery Testing Equipment
Technology Specification

Holder Picture		
	Polymer Clip	Alligator Clip
Picture for reference, please confirm the actual product		
Dimensions (W*D*H) (mm)	3U (19) , 480*330*130	
Equipment Picture		
	Picture for reference, please confirm the actual product	

