HAL
open science

Data-driven strategy for state of health prediction and anomaly detection in lithium-ion batteries

Slimane Arbaoui, Ahmed Samet, Ali Ayadi, Tedjani Mesbahi, Romuald Boné

To cite this version:

Slimane Arbaoui, Ahmed Samet, Ali Ayadi, Tedjani Mesbahi, Romuald Boné. Data-driven strategy for state of health prediction and anomaly detection in lithium-ion batteries. Energy and IA, 2024, 17, pp.100413. 10.1016/j.egyai.2024.100413 . hal-04673182

HAL Id: hal-04673182

https://hal.science/hal-04673182v1

Submitted on 19 Aug 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Energy and AI 17 (2024) 100413

Contents lists available at ScienceDirect

Energy and AI

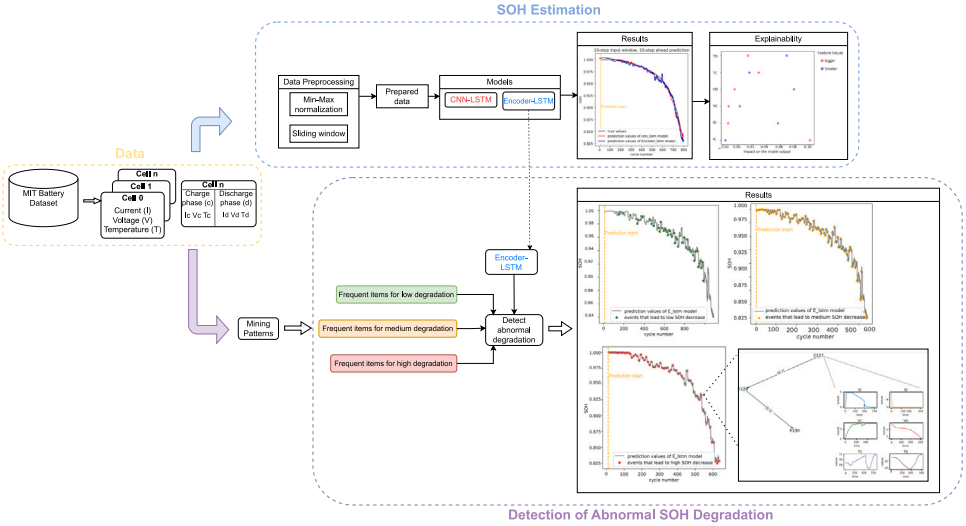
journal homepage: www.elsevier.com/locate/egyai

Data-driven strategy for state of health prediction and anomaly detection in lithium-ion batteries

Slimane Arbaoui [a,*], Ahmed Samet [a], Ali Ayadi [a], Tedjani Mesbahi [b], Romuald Boné [a]

[a] SDC team, ICube laboratory, INSA Strasbourg, University of Strasbourg, 24 Bd de la Victoire, Strasbourg, 67000, France
[b] SMH team, ICube laboratory, INSA Strasbourg, University of Strasbourg, 24 Bd de la Victoire, Strasbourg, 67000, France

GRAPHICAL ABSTRACT



HIGHLIGHTS

- The paper presents E-LSTM and CNN-LSTM models for forecasting Lithium-Ion Battery SOH.
- Research emphasizes model explainability using SHapley Additive exPlanations (Shap).
- The paper integrates pattern mining to detect abnormal degradation in battery SOH.
- Experiments use the battery dataset from the Massachusetts Institute of Technology (MIT).

ARTICLE INFO

Keywords:
Lithium-ion batteries
State of health
LSTM
CNN

ABSTRACT

This study addresses the crucial challenge of monitoring the State of Health (SOH) of Lithium-Ion Batteries (LIBs) in response to the escalating demand for renewable energy systems and the imperative to reduce CO2 emissions. The research introduces deep learning (DL) models, namely Encoder-Long Short-Term Memory (E-LSTM) and Convolutional Neural Network-LSTM (CNN–LSTM), each designed to forecast battery SOH. E-LSTM integrates an encoder for dimensionality reduction and an LSTM model to capture data dependencies.

* Corresponding author.
E-mail addresses: slimane.arbaoui@insa-strasbourg.fr (S. Arbaoui), ahmed.samet@insa-strabourg.fr (A. Samet), ali.ayadi@unistra.fr (A. Ayadi), tedjani.mesbahi@insa-strasbourg.fr (T. Mesbahi), romuald.bone@insa-strasbourg.fr (R. Boné).

https://doi.org/10.1016/j.egyai.2024.100413
Received 7 June 2024; Received in revised form 4 August 2024; Accepted 5 August 2024
Available online 9 August 2024
2666-5468/© 2024 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

Auto-encoders
Pattern mining
Explainable artificial intelligence

CNN–LSTM, on the other hand, employs CNN layers for encoding followed by LSTM layers for precise SOH estimation. Significantly, we prioritize model explainability by employing a game-theoretic approach known as SHapley Additive exPlanations (SHAP) to elucidate the output of our models. Furthermore, a method based on pattern mining was developed, synergizing with the model, to identify patterns contributing to abnormal SOH decrease. These insights are presented through informative plots. The proposed approach relies on the battery dataset from the Massachusetts Institute of Technology (MIT) and showcases promising results in accurately estimating SOH values, in which the E-LSTM model outperformed the CNN–LSTM model with a Mean Absolute Error (MAE) of less than 1%.

1. Introduction

As the world seeks to transition towards a renewable energy system, the importance of energy storage technologies cannot be overstated. Among the frontrunners in this field are Lithium-Ion Batteries (LIBs), which have emerged as a crucial component in the race towards sustainability. This transition has gained significant momentum, with governments and international organizations recognizing the urgent need to combat climate change.

The European Union's Fit for 55 package [1] represents a landmark agreement aimed at reducing greenhouse gas emissions by at least 55% by 2030. This comprehensive strategy encompasses various sectors with strong emphasis on deploying renewable energy sources and electrifying key industries. Simultaneously, the Electric 2035 vision [2] sets forth an ambitious target of electrifying all vehicles sold by 2035, promoting the widespread adoption of electric mobility as a means to decarbonize the transportation sector.

Central to the success of these initiatives is the efficient and reliable storage of renewable energy. LIBs have emerged as the most promising solution, with their high energy density, long lifespan, and versatility making them ideal for a wide range of applications [3,4].

However, ensuring the optimal performance and longevity of LIBs poses a critical challenge. Over time, these batteries can experience degradation due to a multitude of factors, including operating conditions, charging cycles, and manufacturing inconsistencies [5,6]. To address this issue, accurate estimation of the State of Health (SOH) of LIBs is crucial. By understanding and detecting early signs of degradation, maintenance and optimization efforts can be implemented, ultimately extending battery life and reducing the environmental impact associated with frequent replacements.

In the domain of battery SOH estimation and the detection of abnormal decreases, a noticeable gap exists in current research. While some studies focus on accurate SOH estimation or offer insights into predictive models, there is a deficiency in research that simultaneously achieves both objectives and effectively addresses the crucial aspect of identifying abnormal decreases in SOH. To bridge this gap, this article proposes an innovative twofold approach. The initial facet of our approach concentrates on constructing a robust prediction model capable of effectively estimating SOH values. By employing advanced deep learning (DL) algorithms, this model analyzes historical time series data encompassing current, voltage, and temperature during both charge and discharge phases. The model forecasts SOH for the subsequent ten cycles with precision. Recognizing that DL models, while adept at low-error prediction compared to machine learning (ML) models, are often considered as black boxes, we address explainability using a post hoc explainable technique called SHapley Additive exPlanations (SHAP). SHAP elucidates the features upon which the model relies for predictions, clarifying the impact of these features on overall accuracy. This enables us to achieve highly accurate SOH estimations and provides insights into the model's decision-making process. The second facet of our approach employs pattern mining techniques to extract patterns associated with abnormal decreases in SOH values of lithium-ion phosphate (LFP)/graphite batteries, the specific battery type utilized in our dataset. These patterns, in conjunction with our model, form a method enabling the estimation of SOH values and the detection of abnormal decreases.

The rest of this paper is structured as follows: Section 2 provides a brief review of the current state-of-the-art on techniques used to address the problem under consideration. Section 3 defines several key concepts utilized in our approach. Sections 4 and 5 detail the proposed approach, outlining its methodology and steps. Section 6 presents a comprehensive analysis of the results obtained through our approach. In Section 7, a conclusion is drawn based on the findings and implications discussed in the paper.

2. Related work

In recent years, the field of battery health estimation has seen remarkable advancements, driven by the increasing demand for efficient, reliable, and long-lasting energy storage solutions. Researchers have explored various cutting-edge approaches and techniques to enhance the accuracy and explainability of SOH estimation models [5]. First, we need to clearly define what we mean by SOH:

The SOH is a measure of the current health or condition of a battery. It is often expressed as a percentage that reflects the battery's remaining capacity relative to its original capacity when it was new. The formula to calculate SOH is defined by Eq. (1) as follows:

$$SOH(\%) = \left(\frac{Current\ capacity}{Original\ capacity}\right)x100 \tag{1}$$

In this article, we focus on two main parts, models used to estimate the values of the SOH and methods used to explain how these models make their prediction.

2.1. SOH estimation through deep learning

In the realm of battery prognostics, two predominant methodologies have emerged: model-based approaches and data-driven methods [7]. This article focuses its attention on the latter category, specifically delving into data-based methods. These techniques harness raw measurements derived from real-world experiments, sidestepping the intricate exploration of electrochemical reactions integral to the aging process. Instead, the focus is on deciphering the dynamic behavior of the system through empirical data.

For instance, Wen et al. [8] proposed a novel model for predicting battery SOH at different ambient temperatures using Incremental Capacity Analysis (ICA) and a Back Propagation (BP) neural network. By establishing a mapping relationship between temperature and IC curve characteristics, the model achieved accurate predictions with an average error of 1.16%. Gu et al. [9] introduced a hybrid model, integrating Convolutional Neural Network (CNN) and Transformer architectures, to estimate the next SOH value. They utilized Principal Component Analysis (PCA) to derive three features from an initial set of four (temperature, voltage, current, and capacity). The resulting model achieved remarkable accuracy on the NASA dataset, with an error rate of less than 0.55%. On the other hand, Fan et al. [10] introduced a model named T-LSTM, which combines a Transformer and Long Short-Term Memory (LSTM) layer to estimate the SOH of a cycle based on historical voltage data. Their model was trained and evaluated on a dataset developed by the authors. Results demonstrate a significant reduction in error by more than 20% in comparison to employing a simple LSTM model. Feng et al. [11] introduced a method for accurately predicting SOH and Remaining Useful Life (RUL) of LIBs using a

Gaussian Process Regression (GPR) model. They improved the model by incorporating cyclic charging current features highly correlated with battery capacity. The proposed SOH estimation model achieved the highest accuracy compared to other models and demonstrated robustness using random walk battery data. Audin et al. [12] proposed an approach for predicting SOH using auto-encoders and LSTM neural networks applied to usage data. The model is tested on various battery aging datasets and aims to improve energy management strategies and predictive maintenance for LIBs. Jorge et al. [13] presented an approach to predict the evolution of SOH, using an LSTM network, based on features extracted from current, voltage, and temperature curves: average, maximum, minimum, total energy and fundamental frequency, etc. Their study is based on datasets from the Massachusetts Institute of Technology (MIT) and NASA. Liu et al. [14] proposed an approach to estimate the SOH of rechargeable LIBs by leveraging the charging process and LSTM Recurrent Neural Network (LSTM–RNN). Their method enabled online estimation of SOH using health indicators extracted from the charging process. Saxena et al. [15] developed a CNN model to accurately predict battery capacity fade curves, indicating performance degradation. The model successfully identified the rollover cycle, indicating rapid decay onset. This research has implications for monitoring battery health and estimating the remaining lifespan. A novel method for predicting the RUL of LIBs [16] was introduced, utilizing an Encoding Net to capture global degradation tendencies and Bayesian model averaging for integrated estimation, enhancing RUL estimation accuracy. Ardeshiri et al. [17] defined a DL approach for accurately predicting RUL using a stacked bidirectional LSTM (SBLSTM) model with extreme gradient boosting (XGBoost). This combination effectively predicted battery capacity degradation trajectories. Toughzaou et al. [18] combine CNN with LSTM networks to estimate SOH and predict the RUL of LIBs. The model used features, extracted from the voltage and temperature characteristics (max-value, min-value, mean-value) using the K-means algorithm, in addition to the current values. Their hybrid CNN–LSTM method outperformed LSTM, offering an accurate and efficient solution for battery assessment and prognostics. Gong et al. [19] tested several ML models, including Linear Regression (LR), Support Vector Machines (SVM), GPR, and Elastic Net. Elastic Net is a LR regularization technique that automatically selects variables and scales them to enhance prediction accuracy. The models were trained using two different modes: the first mode involved training with data from one cell and testing the model on the remaining cells, and the second mode involved training with 50% of the data from each cell and testing on the remaining 50%. The input for the models consisted of health indicators such as the energy of the Constant Current (CC) charging stage, the energy of the Constant Voltage (CV) charging phase, and the energy of the Equal Discharge Voltage Interval (EDVI). They utilized multiple datasets for training and testing their models, including those from MIT, NASA, CALCE, and Oxford. The models achieved RMSE, MAPE, and MAE of less than 0.5%. Van et al. [20] presented a method for estimating the SOH and internal resistances of LIBs using an LSTM network. Based on experimental data, including voltage, current, and temperature. The first LSTM network is designed to estimate SOH. Then, the data, including current, voltage, temperature, and the estimated SOH, is used by a second LSTM network to estimate the solid electrolyte interphase (SEI) resistance. The model's performance was compared with a Feed Forward neural network (FNN) and achieved a MAE of 0.3%. Chen et al. [21] proposed a hybrid model in which they combine an equivalent circuit model with an improved vision transformer network (VIT) to estimate the SOH. The approach extracts Health Indicator (HI) from battery data (current, voltage and time) using correlation analysis and a fractional-order RC circuit model. These HI will be then fed into a VIT to estimate the SOH. The models were trained and tested using NASA and CALCE datasets and compared with simple LSTM and CNN models, achieving a MAE of 0.36% in both datasets. Chen et al. [22] summarized the majority of the works conducted to estimate SOH values from 2018 to 2022, focusing on the models used, feature extraction methods, and datasets utilized. They compared several state-of-the-art models, including RNN, LSTM, LR, CNN, Transfer learning (TL), and Random Forest, using the MIT battery dataset. The results showcased particularly good performance for the last three models: CNN, TL, and Random Forest. Wang et al. [23] proposed an open-source code library implementing five deep learning models to estimate the SOH. This library can be tested on either the MIT battery dataset or a dataset provided by the authors, which includes 55 batteries cycled under six charging and discharging strategies. The study also compares the impact of three input types (complete charging data, partial charging data, and handcrafted features, which are features calculated based on integrating the current and voltage) and three normalization methods (min–max, z-score, and min–max with range −1 to 1). Qi et al. [24] introduced a novel approach integrating variational mode decomposition (VMD) with a BiLSTM neural network to forecast the RUL of supercapacitors. Aging experiments were conducted under various temperatures and voltages to obtain aging data. VMD was applied to decompose the aging data, eliminating disturbances such as capacity recovery and test errors. The hyperparameters of BiLSTM were optimized using the sparrow search algorithm (SSA) to enhance the alignment between input data and network structure. The decomposed aging data were then fed into BiLSTM for prediction. Experimental results indicated that the VMD-SSA-BiLSTM model achieved high prediction accuracy and robustness, with an average RMSE of 0.11, a 44.3% reduction compared to BiLSTM alone, and a minimum RMSE of 0.031.

2.2. Explainable models for SOH estimation

In the landscape of DL models for SOH estimation, there is a growing recognition of the need for transparency and explainability in model outcomes. As these sophisticated models become integral to prognostic applications, their inherent complexity often renders their decision-making processes opaque [25]. In response to this challenge, the emerging field of explainable artificial intelligence (XAI) seeks to bridge the gap between predictive accuracy and model explainability. XAI employs two methodologies to make complex models more explainable, post hoc and intrinsic methods [25].

Post hoc techniques aim to interpret and explain the predictions of already-trained models. Two prominent examples in this category are SHAP and Local Interpretable Model-agnostic Explanations (LIME). SHAP values provide a mathematical foundation for assigning the contribution of each feature to a model's output, offering a comprehensive understanding of feature importance. On the other hand, LIME generates locally faithful explanations for specific predictions by approximating the behavior of the model around a particular instance.

Intrinsic techniques, on the other hand, focus on imbuing models with interpretability from the outset of their design. Rather than explaining an existing model after training, intrinsic methods prioritize building models that are inherently understandable. While these techniques are still evolving, they hold promise for seamlessly integrating interpretability into the core of DL models.

The demand for explainable models in SOH estimation arises from the necessity to comprehend and trust the insights provided by DL algorithms. A number of studies have been carried out to improve the interpretation of SOH estimation models for LIBs.

One notable work by Wang et al. [26] focuses on enhancing both the accuracy and explainability of SOH estimation models. They proposed an innovative approach that combines ML techniques with explainability methods. By collecting battery data, including voltage, current, and temperature, they trained ML models to predict SOH. They applied explainability techniques, such as feature importance analysis and rule extraction, to identify the most influential features and derive interpretable rules for SOH estimation.

Another significant contribution by Kim et al. [27] introduced a prognostics framework based on DL for estimating both SOH and RUL

of LIBs. Their framework consisted of three phases. In the first phase, impedance-related features were extracted from discharge curves, which were reliable and real-time accessible. These features served as inputs for the subsequent phases. The second phase involved a knowledge-infused RNN, combining an empirical model with a deep neural network. This hybrid model leveraged the collected impedance-related features to make accurate SOH and RUL estimations. In the third phase, the authors used the Monte Carlo dropout, a DL method that provides probabilistic predictions from neural networks, to enhance the reliability and robustness of the estimations.

Shi et al. [28] propose a Physics-infused (PI-LSTM) model integrating physics-based modeling with LSTM techniques to estimate degradation and forecast RUL of LIBs. The study showcases the model's capability to accurately capture battery aging behavior and make reliable predictions, offering valuable insights for battery health management and maintenance.

Table 1 presents a comprehensive comparison of several related works in the field, focusing on key aspects such as the models used, evaluation results, model explainability methods, datasets and features, and SOH abnormal degradation detection. By examining these critical factors, the table provides an in-depth overview of the current state-of-the-art approaches and their strengths and weaknesses. Several conclusions can be drawn from this detailed comparison:

- *Diverse approaches*: The table highlights a wide range of approaches for SOH estimation.
- *Datasets*: The studies make use of diverse datasets from sources such as Sandia National Laboratories, MIT, NASA Prognostics Center of Excellence (PCoE), and the Center for Advanced Life Cycle Engineering (CALCE), among others. These datasets encompass a range of battery types and operational conditions, enhancing the robustness and generalizability of the research findings.
- *Explainability methods*: A few studies incorporate explainability methods to shed light on the factors influencing SOH estimation. These methods aim to improve transparency by identifying key features, patterns, or variables contributing to the predictions. However, the majority of the reviewed works do not explicitly address explainability, indicating an area for further development.
- *SOH degradation detection*: None of the reviewed works specifically focus on the real-time detection of abnormal degradation in SOH. The primary emphasis is on SOH prediction rather than on timely anomaly detection. This reveals a potential research gap, suggesting a need for dedicated efforts to develop techniques for the early detection of anomalies in battery health.

3. Background

This section is intended to describe the three different types of neural networks used in the present work.

3.1. LSTM model

LSTM is a specialized type of RNN designed to address the limitations of traditional RNNs in capturing long-term dependencies in sequential data. Traditional RNNs often suffer from the vanishing gradient problem, which hinders learning across long sequences. LSTMs introduce a memory cell and specific gating mechanisms to mitigate this issue, as depicted in Fig. 1.

The key components of an LSTM cell are as follows:

- *Memory cell*: Stores and updates information over time, serving as a conduit for information from one time step to the next.
- *Forget gate*: Determines the relevance of information from the previous memory cell state and decides what should be discarded.

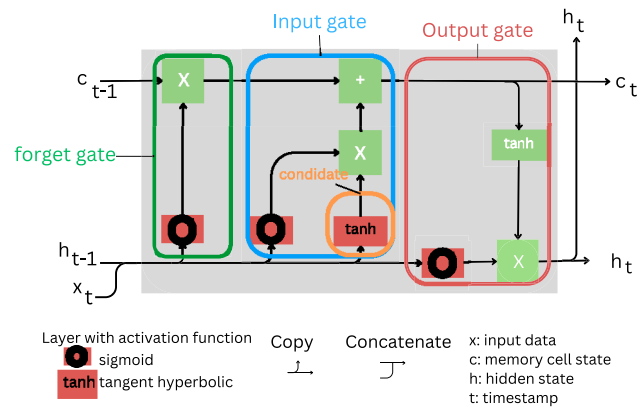


Fig. 1. Detailed architecture of an LSTM cell.

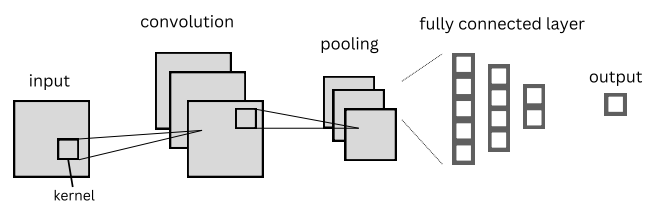


Fig. 2. The typical CNN model architecture.

- *Input gate*: Decides which new information is important and should be stored in the memory cell state.
- *Candidate layer*: Generates potential new candidate values that could be added to the memory cell state.
- *Update mechanism*: Combines the input gate and candidate layer to update the memory cell state.
- *Output gate*: Identifies which part of the memory cell state should be output as the hidden state at the current time step.

These components work together to enable LSTMs to capture and manage long-term dependencies in sequential data, making them particularly effective in various applications, including time series analysis.

3.2. CNN model

CNNs are specialized DL models designed for processing grid-like data, such as images and videos. Inspired by the human visual cortex, CNNs exhibit an hierarchical feature extraction [29]. The CNN model as illustrated in Fig. 2 includes [30]:

- *Convolutional layers*: The fundamental building blocks of CNNs that use filters to detect patterns and local features within the input.
- *Pooling layers*: These layers reduce the spatial dimensionality of feature maps, preserving important features while simplifying complexity.
- *Activation functions*: Introduce non-linearity to neurons, enabling the modeling of intricate relationships within the data.
- *Fully connected layers*: These layers map learned features to the desired output, facilitating final predictions.

CNNs excel in tasks like image classification, object detection, and image segmentation. Their proficiency lies in their capacity to extract pertinent and hierarchical features from images, making them an ideal choice for such applications.

Table 1
Comparison table of related work.

Approach	Model	Evaluation results	Dataset	Features and window size	Explainability methods	SOH degradation detection
Wen et al. [8]	BP neural network	max-error = 3.09, mae = 1.18, mse = 0.22	Five 18 650 ternary lithium batteries are selected	Temperature and IC curve characteristics	✘	✘
Audin et al. [12]	Auto-encoder LSTM	mae = 1.03×10^{-2}, std = 0.93×10^{-2}, rmse = 1.32×10^{-2}, mape = 1.31	Sandia National Laboratories	Current, temperature and voltage; window: 1, 25, 50, 100, 200	✘	✘
Jorge et al. [13]	LSTM	mae = 0.77×10^{-2}, std = 0.7×10^{-2}, rmse = 1.04×10^{-2}, rmspe = 1.14%, nmse = 0.48×10^{-2}	MIT and NASA	Current, temperature and voltage; window: 25	✘	✘
Liu et al. [14]	LSTM–RNN	rmse = 0.56×10^{-2}	NASA and CALCE	Voltage; window: 10	✘	✘
Wang et al. [26]	LRP-driven model	mae = 0.71×10^{-2}, mape = 0.75×10^{-2}, rmse = 0.99×10^{-2}	MIT	Current, and voltage curve during charging process	✓	✘
Saxena et al. [15]	CNN	mae = 4.0%	Commercial graphite/LiFePO4 cylindrical battery (A123 Systems) cycling dataset	Discharge curves capacity and voltage; window: 100	✘	✘
Wang et al. [16]	Bayesian model averaging	rmse = 2.16, mape = 12.54%	Data-driven prediction of battery cycle life before capacity degradation and data of LiFePO4 (LFP)	Current, temperature and voltage	✘	✘
Ardeshiri et al. [17]	SBLSTM and XGBoost	rmse = 1.94×10^{-2}	NASA and MIT	Current, temperature and voltage; window: 50	✘	✘
Toughzaoui et al. [18]	CNN–LSTM and LSTM	mse = 0.02×10^{-2}, rmse = 1.40×10^{-2}, mae = 0.76×10^{-2}	NASA	Current, temperature and voltage	✘	✘
Shi et al. [28]	PI-LSTM	mape = 1.72%	NASA	Cycle time, rest time, environmental temperature, SOC level, and loading condition	✓	✘
Gu et al. [9]	CNN-Transformer	error rate < 0.55%	NASA	PCA features from temperature, voltage, current, and capacity	✘	✘
Fan et al. [10]	T-LSTM	Significant reduction in error by > 20% compared to simple LSTM	Dataset developed by authors	Historical voltage data	✘	✘
Hailin et al. [11]	GPR	Highest accuracy compared to other models	Random walk battery data	Cyclic charging current features	✘	✘
Gong et al. [19]	Elastic Net, LR, SVM, GPR	rmse, mape, mae < 0.5%	MIT, NASA, CALCE and Oxford	Health indicators from charging stages	✘	✘
Van et al. [20]	LSTM	mae = 0.3%	Experimental data	Voltage, current, temperature, SEI resistance	✘	✘
Chen et al. [21]	Hybrid model (ECM+VIT)	mae = 0.36%	NASA, CALCE	Health indicators from battery data	✘	✘
Chen et al. [22]	Various (RNN, LSTM, LR, CNN, TL, Random Forest)	Comparison study	MIT	Various extracted features	✘	✘

3.3. Auto-encoders models

Auto-encoders are a type of neural network architecture used for unsupervised learning and dimensionality reduction. They encode input data into a lower-dimensional latent space and then decode it to reconstruct the original input [31].

Auto-encoder as illustrated in Fig. 3 include:

- *Encoder*: This component maps input data to a lower-dimensional latent space through several layers [32].
- *Latent space*: It represents a lower-dimensional representation capturing essential features of the input data.

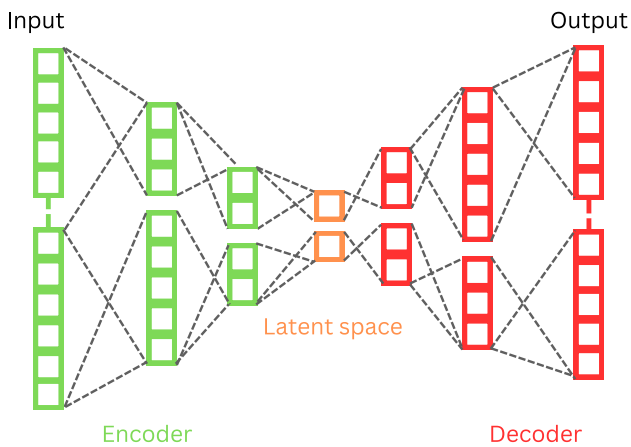


Fig. 3. Structure of an auto-encoder.

- *Decoder*: The decoder's role is to reconstruct the original input from the encoded representation [32].
- *Reconstruction loss*: This metric quantifies the dissimilarity between the original input and the reconstructed output, guiding the learning process [32].

Variations of auto-encoders include Variational Auto-encoders (VAEs), Denoising Auto-encoders, and Sparse Auto-encoders, each variation is suited for specific tasks [33–35].

Auto-encoders find applications in data compression, anomaly detection, feature learning, denoising, and generative modeling, especially when labeled data is scarce [31].

4. Encoder-LSTM and CNN-LSTM models for SOH estimation

Our primary approach revolves around developing a highly accurate SOH estimation model. To achieve this objective, we will harness the power of DL algorithms, using the MIT battery dataset [36] as a training dataset. The model will take into account various influential factors affecting battery health, including temperature, current and voltage in both charge and discharge cycles. By incorporating this comprehensive set of variables, we aim to create a robust and reliable prediction model (See Fig. 4) that can effectively assess the state of battery health and anticipate its performance over time.

4.1. Data set

To develop the proposed DL model an initial dataset derived from 124 LIBs was used, each possessing a capacity of 1.1 Ah and a nominal voltage of 3.3 V. The data is organized into three batches, each one containing around 1000 charge and discharge cycles, drawn from a total of 48 different cells.

The data was recorded using a 48-channel Arbin discharge bench, with an initial temperature of 30°. During the charging cycles, different charging profiles were used for different cells. Some cells were charged in two stages. They were first charged at a rate of $8C$ until they reached a State of charge (SOC) of 35%, after which the charge rate was reduced to $3.6C$ until they reached an SOC of 80%. For other cells, a different approach was taken, with an initial charging rate of $5.4C$ until the SOC reached 40% followed by a charging rate of $3.6C$. The final charging phase involved a more gradual charge rate of $1C$ from 80% to full charge. Full charge and discharge were determined by voltage thresholds of 3.6 V and 2 V, respectively, in which all cells were discharged at $4C$ in a CC mode.

The dataset provided detailed information on these diverse charging cycles, presented in the format of $C1(Q1) - C2$, where $C1$ and $C2$ represent the charging slot capacities applied, and $Q1$ represents the SOC at which the charge switches from $C1$ to $C2$. This dataset covers a range of battery cycles, providing our model with the flexibility to adapt to different real-world scenarios. The dataset is available in format, which can be loaded into Python using the h5py package.[1] Three batches are relatively extensive, collectively amounting to 8 GB.

To extract data for current (I), voltage (V), and temperature (T) during both the charge (c) and discharge (d) phases, we use the current values as a reference point. As defined in Eq. (2), if the current value is greater than or equal to zero we infer that we are in the charge phase; otherwise, we identify it as the discharge phase.

$$\begin{array}{ll} Charge, & \text{if } I \geq 0 \\ Discharge, & \text{if } I < 0 \end{array} \tag{2}$$

4.2. Data preprocessing

During the dataset analysis, we observed that each cell had a varying number of cycles, and the length of features differed from cycle to cycle. To address this issue, we fix the length of each feature to a specific value (L), ensuring uniformity in the input data, and processing by the DL model. Fixing feature length is crucial because many DL models require inputs of consistent lengths to ensure optimal performance. Moreover, it simplifies the data by eliminating any disparities in feature lengths, making the results easier to analyze and interpret.

For model training preparation, we applied a common normalization technique known as min–max normalization as defined by Eq. (3), this technique scales the feature values in the training set to a range between 0 and 1, based on the minimum and maximum values of each feature. We adopted the same minimum and maximum values to normalize the test set.

$$normalize\ value = \frac{data - min(data)}{max(data) - min(data)} \tag{3}$$

This standardization ensures that each feature contributes equally during training, preventing certain features from dominating the training process.

The subsequent phase involves the application of the sliding window technique [13] to the normalized data. In this context, we have opted for a consistent window size of 10 cycles for both input and output sequences. The normalized data undergoes segmentation, where in each window encapsulates the features of a specific time span. This approach facilitates the generation of coherent input–output pairs that capture the temporal dynamics of the battery system. For clarity, Fig. 5 provides a visual depiction of this sequentialization process, elucidating the step-by-step transformation from the raw data to the organized input–output sequences.

After completing the data preprocessing, the next step was to split the dataset into a training set (60%) and a test set (40%). This division is specifically designed for the auto-encoder model, enabling it to learn patterns from the training set and evaluate its performance on unseen data during testing. By allocating a larger proportion of the data to the testing set, we ensure that the model is not overfitting to the training data and can generalize well to new samples

4.3. Model generating

We have introduced two models, E-LSTM and CNN–LSTM, designed for the estimation of SOH values in LIBs. Both models were developed using the Tensorflow framework.

[1] h5py: https://www.h5py.org/.

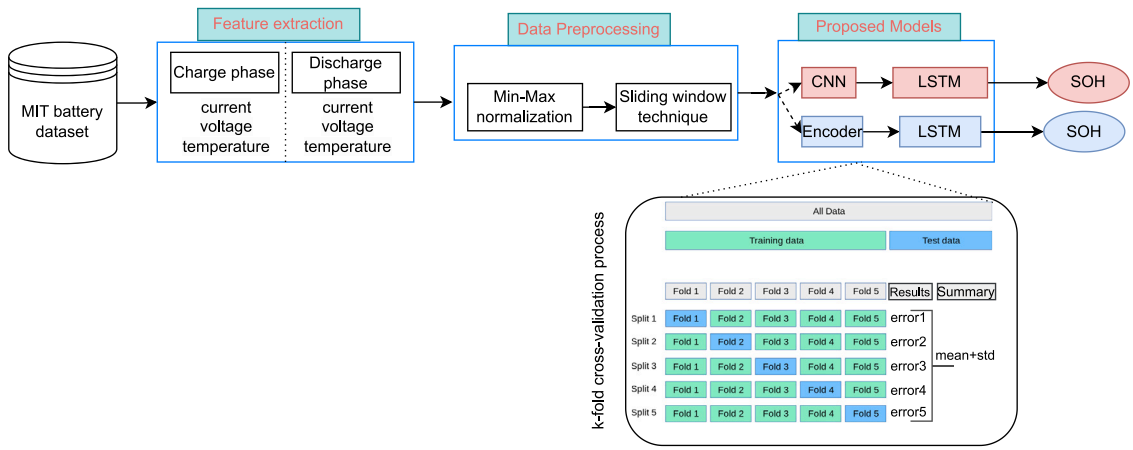


Fig. 4. Flow chart of the proposed SOH estimation approach.

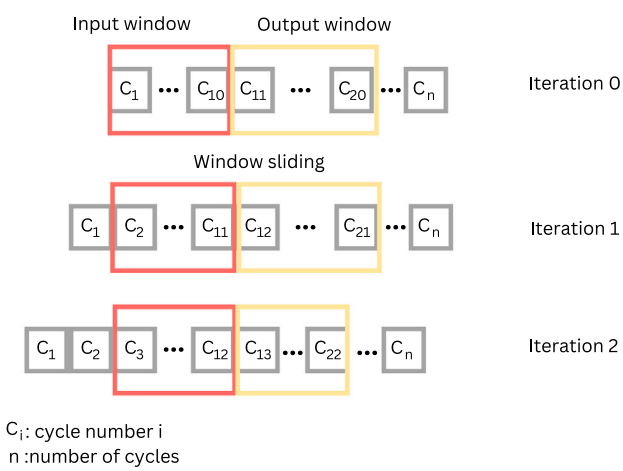


Fig. 5. Sliding window technique.

4.3.1. E-LSTM model

The model comprises two key components: an encoder and an LSTM model. The encoder takes an input of shape $(10, 6, L)$ and compresses it into a flattened representation of shape (640,), thereby reducing dimensionality while retaining vital features. The output from the encoder is then reshaped into (10, 64) and provided as input to LSTM model. As presented in Fig. 6, the LSTM is trained to forecast SOH values for the subsequent ten cycles based on historical data. At each time step, the LSTM receives a vector of shape (64,), representing the encoded information for that cycle. The combination of the encoder and LSTM model in our approach allows us to efficiently capture critical battery performance features and use them for precise SOH value estimations in future cycles.

Encoder part: within the encoder component of the proposed E-LSTM model, we have implemented two distinct architectures, the dense encoder and the CNN encoder. The dense encoder consists of multiple fully connected layers, making it suitable for processing structured data with a fixed size, such as numerical data. This architecture excels in capturing complex relationships between input features, although it may have limitations when dealing with large volumes of data. Conversely, the CNN encoder integrates convolutional layers to extract relevant features from the input data. These layers apply filters to the input, enabling the network to capture local patterns and spatial information. Subsequently, the output of each convolutional layer is then passed through a max pooling layer, reducing data dimensionality while retaining essential features. We conducted a thorough evaluation of both encoder architectures on the dataset and the results, which are summarized in Table 2, provide valuable insights into their suitability for the SOH estimation approach. The dense encoder demonstrated superiority over the CNN encoder in terms of its simplicity and reduced parameter count. Its leaner architecture made it easier to train and less prone to overfitting, which is crucial for achieving better generalization on unseen data.

Table 2
The mean square error (MSE) for both encoders within the training and testing.

MSE	Dense encoder	CNN encoder
Training	0.0020	0.0276
Testing	0.0083	0.0410

LSTM part: The LSTM model underwent training and evaluation using k-fold cross-validation, a widely adopted technique for assessing model performance and curbing overfitting [37]. The same technique was subsequently applied to the CNN–LSTM model. The dataset was partitioned into five equally sized folds. Throughout each iteration of the k-fold cross-validation process, four folds were employed for training the LSTM model, while the remaining fold served as the test set. This process was repeated five times, ensuring each fold had its opportunity as the test set. The Mean Absolute Error (MAE), Mean Squared Error (MSE) and Mean Absolute Percentage Error (MAPE) were computed for each fold to assess the model's performance on the test data. In each iteration, the model weights were initialized using the Glorot (Xavier) uniform initializer,[2] which consistently yielded superior results compared to other initializers. This initialization method involves drawing samples from a uniform distribution within the range $[-\text{limit}, \text{limit}]$, where the limit is determined by Eq. (4):

$$\text{limit} = \sqrt{\frac{6}{\text{fanin} + \text{fanout}}} \tag{4}$$

here, fanin represents the number of input units, and fanout represents the number of output units. Additionally, biases were initialized to zeros. The incorporation of k-fold cross-validation provides a more robust estimate of the model's generalization ability to new data and mitigates the risk of overfitting.

4.3.2. CNN-LSTM model

The CNN–LSTM model serves as an extension of our E-LSTM model, with the specific aim of estimating SOH values for the next ten cycles based on input data of shape (10, 6, L). This indicates that the model

[2] https://www.tensorflow.org/api_docs/python/tf/keras/initializers/GlorotUniform.

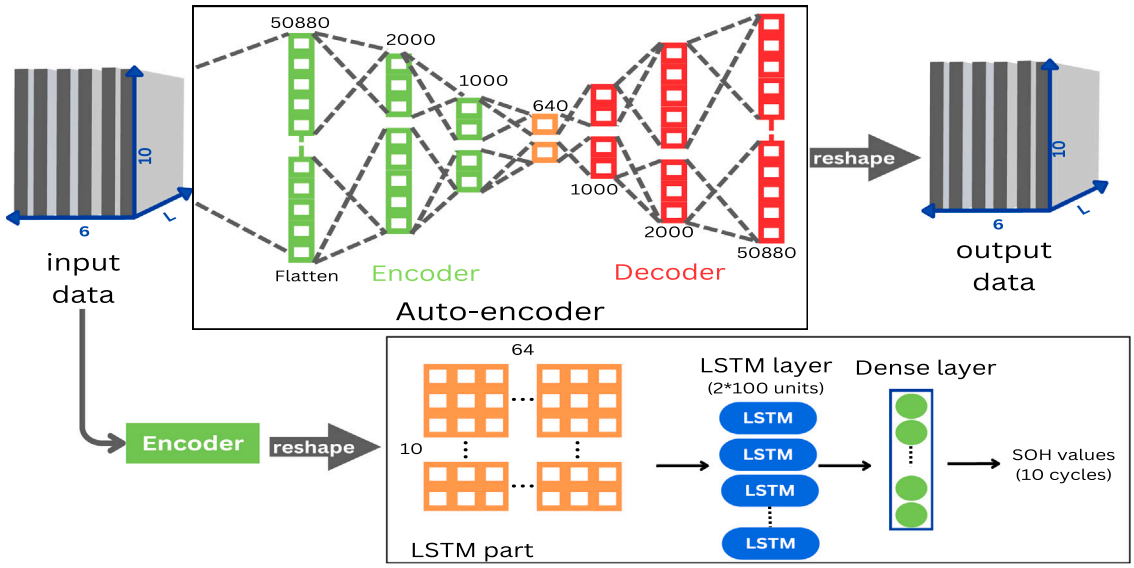


Fig. 6. The E-LSTM model architecture.

processes a sequence of ten-time steps, each containing six features (Ic, Id, Vc, Vd, Tc, Td), and each of these features possesses L dimensions, representing the maximum length of each feature in the dataset. Shorter features are standardized to this length using zero-padding. We introduced this second model to facilitate a comparative analysis with our first approach, which uses an encoder to reduce dimensionality. In contrast, the CNN–LSTM model does not employ an encoder. This comparison enables us to assess the impact of the encoder on estimation outcomes.

Within the CNN–LSTM model, we integrate convolutional layers to extract spatial features from the input data. These layers apply filters to the input, allowing the network to capture local patterns and spatial information effectively. The output from the convolutional layers is then passed into LSTM layers, which capture temporal dependencies in the data. These LSTM layers enable the model to learn how the input features evolve over time, subsequently making estimations for future time steps, as illustrated in Fig. 7. By combining both convolutional and LSTM layers, the CNN–LSTM model leverages spatial and temporal information, making it capable of handling complex patterns and relationships within the input data.

4.4. Model explainability

To gain insights into the factors influencing our model's predictions, we developed a specialized function to calculate SHAP values. SHAP values are a valuable technique for quantifying the contribution of each feature to the prediction, enhancing model explainability.

Given the complexity of our input data, we opted not to use the SHAP library and instead implemented the approach proposed by Strumbelj et al. [38]. This method utilizes Monte-Carlo sampling to approximate the SHapley value for each feature.

The function takes the model and the input data X as input. It then performs Monte-Carlo sampling by iterating M times. In each iteration, a sample is randomly selected from the input data, and a modified sample without the jth feature is created. For both the original and modified samples, the function calculates the model's predictions. Subsequently, it computes the difference between the predictions and averages them. This process is repeated for all six features (Ic, Id, Vc, Vd, Tc, Td), resulting in a list of SHAP values quantifying the impact of each feature on the model's predictions.

With this function, we can gain valuable insights into the factors influencing our model's predictions, aiding our understanding of the underlying patterns and feature importance in predicting SOH values.

5. Pattern mining for SOH abnormal decrease

To uncover patterns leading to abnormal SOH decrease in LIBs, we present an approach centered on the comparative analysis of a reference cell identified within the MIT battery dataset. This particular cell exhibits distinctive characteristics, notably boasting the longest lifespan in the dataset and having reached its end of life. In contrast, a substantial number of atypical cells are considered in our analysis. The viability of this comparison is facilitated by the standardized composition of the MIT battery dataset, where a singular type of LIB is used, specifically LFP /graphite cells and only the charge policy is changed. The objective is to discern sequences within the dataset that are associated with a decline in SOH. By contrasting the behavior of the reference cell with that of the atypical cells, our approach aims to pinpoint patterns indicative of abnormal SOH deterioration, contributing to a deeper understanding of the factors influencing LIB health degradation. To achieve this, we have categorized the SOH drops into three levels, as illustrated in Eq. (5).

$$\text{SOH_categories} = \begin{cases} \text{Low}, & \text{if } 1\% \leq \text{SOH_Drop} < 9\% \\ \text{Medium}, & \text{if } 9\% \leq \text{SOH_Drop} < 14\% \\ \text{High}, & \text{if SOH_Drop} \geq 14\% \end{cases} \tag{5}$$

These categories serve as reference points for evaluating the severity and extent of SOH declines observed in the dataset. By classifying the declines into different levels, we can systematically analyze and compare the patterns and characteristics associated with each category. This allows us to identify specific sequences of events that contribute to SOH decreases and gain insights into the underlying factors impacting battery health.

Furthermore, this approach allows for adaptability, as we can adjust the categories as needed to better capture the diverse range and severity of SOH declines encountered in real-world scenarios. We aim to identify events that contribute to abnormal drops in the SOH values of LIBs and use them, combined with our model, to present a predictive SOH maintenance approach for LIBs.

We have developed a series of algorithms to extract, identify, and transform the relevant data. Here's an overview of the steps involved:

- **Sequence extraction (Algorithm 1):** The initial stage of this approach involves the computation of the derivative of SOH values at each data point for both typical and untypical cells. This derivative represents the rate of change of SOH with respect to cycle number.

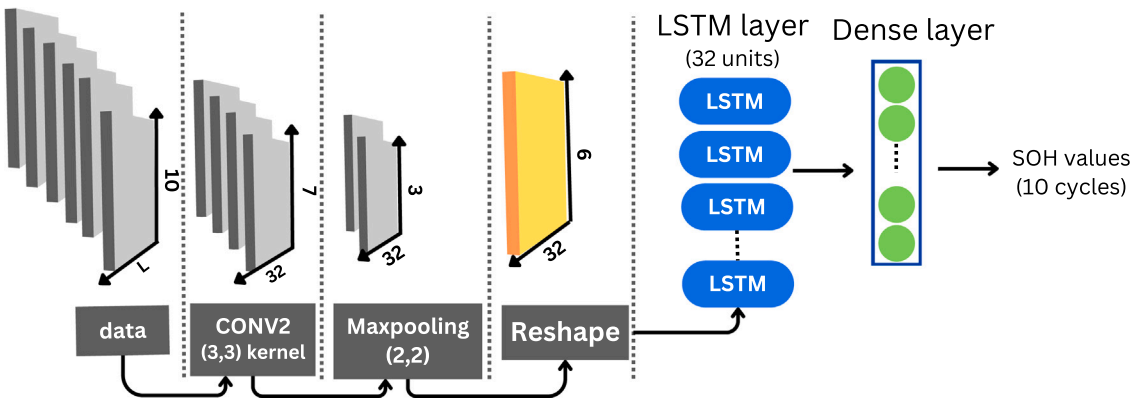


Fig. 7. CNN–LSTM model architecture.

Subsequently, a comparative analysis is conducted between these derivatives. In instances where the disparity between the two values surpasses a predetermined threshold, determined through a series of empirical tests, we identify it as a significant contributing factor to the observed SOH drop. The outcome of this phase is a list containing the cycle number at which the drop is detected and the corresponding percentage of deviation between the typical and untypical SOH values. The cycle number is then employed to extract the data related to current, voltage, and temperature from the preceding ten cycles leading up to the detection point, while the percentage of deviation aids in categorizing the severity of the SOH drop.

Algorithm 1 The getSequences function

function $getSequences$(typical_cell, untypical_cell)
  $DSOH_{\text{typical}} \leftarrow$ derivation(typical_cell[SOH])
  $DSOH_{\text{untypical}} \leftarrow$ derivation(untypical_cell[SOH])
  $SOH1 \leftarrow$ typical_cell[SOH]
  $SOH2 \leftarrow$ untypical_cell[SOH]
  $index \leftarrow$ compare($DSOH_{\text{typical}}, DSOH_{\text{untypical}}, SOH1, SOH2$)
  return $index$
end function
function $compare(df1, df2, soh1, soh2)$
  $size \leftarrow \min(\text{len}(df1), \text{len}(df2))$
  $differences \leftarrow$ abs(np.array($df1[0:size]$) − np.array($df2[0:size]$))
  $index \leftarrow []$
  for i in range(len($differences$)) **do**
    if differences[i] ≥ 0.001 **then**
      $x \leftarrow 100 - \left(\frac{\text{soh2}[i+1]\times 100}{\text{soh1}[i+1]}\right)$
      index.append($(i+1, x)$)
    end if
  end for
  return index
end function

- **Events identification (Algorithm 2):** In the following step, we identify events from the sequences obtained from the previous step. An event is represented as a tuple containing the data (d), charge policy (chp), and cycle number (cn). The data includes the current, voltage, and temperature values during both the charge and discharge phases. Based on the preceding result, we construct three DataFrames, corresponding to each category:
  - Df_low: DataFrame of sequences that lead to a low degradation in the SOH values.
  - Df_medium: DataFrame of sequences that lead to a medium degradation in the SOH values.
  - Df_high: DataFrame of sequences that lead to high degradation in the SOH values

Algorithm 2 The getEvents function

function $getEvents(Sequence, untypical_{cell})$
  $df \leftarrow$
    DataFrame(events_extract($Sequence_DF, untypical_{cell}$))
  $SOH_categories \leftarrow [1, 9, 14, \infty]$
  $df['result'] \leftarrow$ Discretize($df['result']$, bins $= SOH_categories$, $labels = ['low', 'medium', 'high']$)
  ▷ Discretize the SOH column
  $Df_{low} \leftarrow df[df['result'] == 'low']$
  $Df_{medium} \leftarrow df[df['result'] == 'medium']$
  $Df_{high} \leftarrow df[df['result'] == 'high']$

  return $Df_{low}, Df_{medium}, Df_{high}$
end function
function $events_extract(sequence, data)$
  $Events \leftarrow []$
  $result \leftarrow []$
  for $sequence_i$ in $sequence$ **do**
    $Events$.add($sequence_i.d, sequence_i.chp, sequence_i.cn$)
    $result.add(sequence_i.SOH_{decrease})$
  end for

  return $Events, result$
end function

- **Transformation:** To identify patterns and recurring occurrences, we undertake a transformation process that shifts the original format of the data, which consists of an array of L values for each feature, into a string representation. This string representation allows us to determine whether two events are equivalent or not, an essential for identifying recurring patterns. The conversion of data into strings streamlines the comparison of events and facilitates the detection of similarities that may not be readily apparent in their numerical representations. To accomplish this transformation, we begin by comparing the shape of each feature, utilizing the correlation coefficient, and we consider the charge policy. This approach ensures that events sharing both the same feature shape and charge policy are assigned the same label or name. We store the true values associated with each feature in a list, allowing us to return to them when necessary. Based on the results of the previous steps, we construct three DataFrames, one for each category (low, medium, high). However, in these DataFrames, we utilize data in their transformed string representation.

- **The PrefixSpan algorithm:** Next, we use the PrefixSpan algorithm [39] to extract frequent patterns from each DataFrame. It is a powerful method for extracting frequent sequential patterns from a sequence database. It efficiently identifies the most common sequential patterns that occur across multiple sequences, making it suitable for handling large-scale data.

The output from the preceding step consists of three DataFrames, each dedicated to a specific category, encapsulating the most prevalent sequences of events contributing to SOH degradation. These categorized DataFrames serve as critical inputs, collectively employed with our model, to execute a predictive maintenance strategy for SOH in LIBs.

6. Results and discussion

This section presents the results of the different experiments carried out as part of this work.

6.1. SOH estimation and explainability

The models were trained and tested using a machine with the following specifications: an Intel Core i5 12th generation CPU, an NVIDIA GeForce RTX 3070 GPU with 6 GB of VRAM, 32 GB of RAM, and a 512 GB SSD. This machine provided the necessary computing power to train and test our models efficiently and effectively. The high-end NVIDIA GPU played a crucial role in accelerating the training process, enabling faster iterations and more rapid experimentation. A large amount of RAM was useful for handling and processing large datasets, and the high-speed SSD enabled quick access to data during training and testing. In summary, the machine used for training and testing the models was equipped with top-of-the-line hardware, enabling us to perform complex computations.

6.1.1. Error metrics

To evaluate the performance of our models, we used several metrics commonly used in DL. These metrics include:

- Mean Squared Error (MSE): As defined in Eq. (6), it quantifies the average squared difference between the predicted and actual values, providing an indicator of the overall accuracy of the model's predictions (n represents the total number of samples).

$$MSE = \frac{1}{n}\sum(predicted\ value - observed\ value)^2 \tag{6}$$

- Mean Absolute Error (MAE): As defined in Eq. (7), MAE measures the average absolute difference between the predicted and actual values. This metric offers insight into the magnitude of errors in the model's predictions.

$$MAE = \frac{\sum|predicted\ value - observed\ value|}{n} \tag{7}$$

- Root Mean Squared Error (RMSE): As defined in Eq. (8) is the square root of the MSE, and it is commonly used to measure the average difference between the predicted and actual values in the same units as the original data.

$$RMSE = \sqrt{\frac{\sum(predicted\ value - observed\ value)^2}{n}} \tag{8}$$

- Mean Absolute Percentage Error (MAPE): As defined in Eq. (9), MAPE is expressed as a percentage and represents the average absolute difference between the actual values and the predicted values as a percentage of the actual values. It is a metric commonly used to assess the accuracy of forecasting models.

$$MAPE = \frac{\sum|predicted\ value - observed\ value|}{observed\ value} * 100 \tag{9}$$

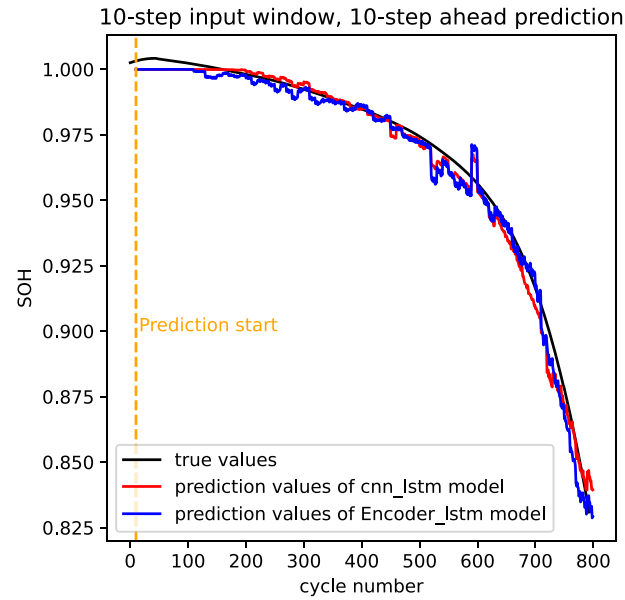


Fig. 8. SOH estimation using 10-step input window for 10-step output window.

Table 3
E-LSTM and CNN–LSTM performance results, with 10 step inputs window and 10 step ahead estimation.

Metric	E-LSTM model		CNN–LSTM model	
	Mean	Std	Mean	Std
MAE ($\times 10^{-2}$)	0.86	0.06	0.90	0.08
MSE ($\times 10^{-3}$)	0.17	0.02	0.22	0.03
RMSE ($\times 10^{-2}$)	1.30	0.44	1.40	0.54
MAPE	0.91	0.06	0.96	0.09

Table 4
E-LSTM and CNN–LSTM performance results, with 25-step inputs window and 25 step ahead estimation.

Metric	E-LSTM model		CNN–LSTM model	
	Mean	Std	Mean	Std
MAE ($\times 10^{-2}$)	0.83	0.07	1.11	0.23
MSE ($\times 10^{-3}$)	0.16	0.02	0.35	0.12
RMSE ($\times 10^{-2}$)	1.26	0.44	1.87	1.07
MAPE	0.89	0.07	1.25	0.24

These four criteria are used to measure the accuracy of both E-LSTM and CNN–LSTM models used in the proposed SOH estimation approach. Results of this evaluation are presented in Table 3 where the reported values represent the mean performance across the five folds, with the accompanying Standard deviations (Std) providing insight into the variability of the results. Fig. 8 shows the SOH estimation values for both models compared to the SOH true values of a given cell. We also tested the models performances using a 25-step input window and 25-step ahead estimation and the results of this test are presented in Table 4.

6.1.2. Discussion

The comparative analysis of models revealed that both models performed well with low errors, but the E-LSTM model outperformed the CNN–LSTM model, attributed to the encoder of the E-LSTM model, which altered the input even when the error was small, which helped prevent overfitting.

To ensure a fair comparison with other related works, we adjusted the step window to 25 inputs and 50 outputs for both models and compared our models with those developed in articles [12,13,20,23]. The results of this comparative analysis are presented in Table 5. It

Table 5
Performance comparison of the proposed E-LSTM and CNN–LSTM models with state-of-the-art over the MIT battery dataset.

Metric	E-LSTM	CNN–LSTM	AE-LSTM [12]	BiXLSTM [13]	XLSTM [13]	LSTM [20]	LSTM [23]	CNN [23]
MAE ($*10^{-2}$)	0.89 ± 0.05	1.19 ± 0.18	2.43	1.10	1.07	9.79	10.54	12.12
MSE ($*10^{-3}$)	0.19 ± 0.02	0.35 ± 0.09	0.76	0.28	0.25	11.51	11.41	16.90
RMSE ($*10^{-2}$)	1.37 ± 0.44	1.87 ± 0.94	2.77	1.70	1.60	10.72	11.90	13.01
MAPE	0.95 ± 0.05	1.27 ± 0.20	2.66	N/A	N/A	10.01	10.80	12.44

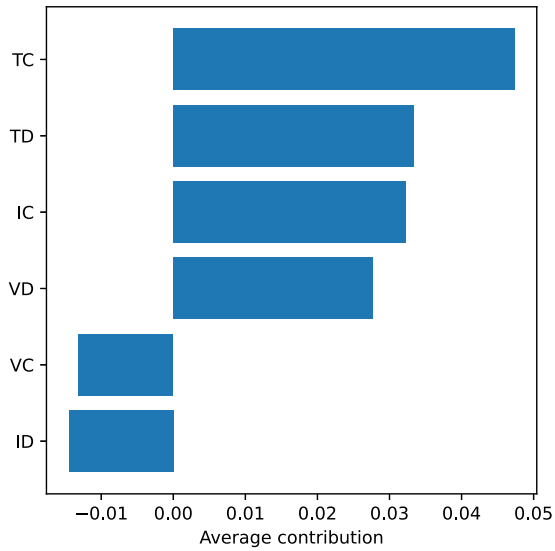


Fig. 9. The average contribution of each feature in the model's prediction using E-LSTM.

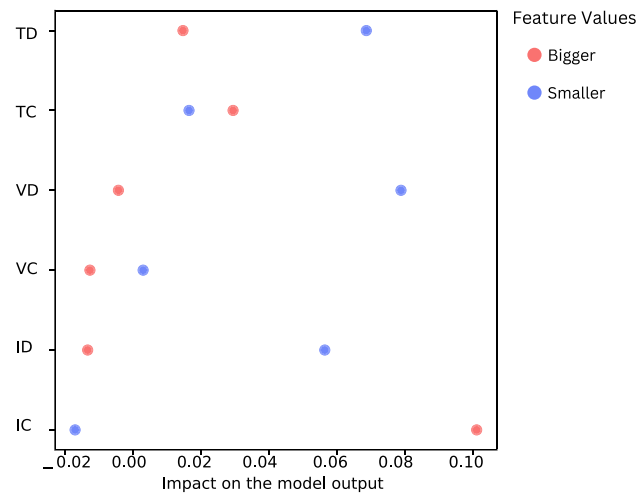


Fig. 10. Impact of feature values on SOH estimation using the E-LSTM.

is evident that E-LSTM model yield promising results compared to the models presented in those articles, these results could be attributed to the use of the k-fold cross-validation technique that helped reduce the overfitting and enhance the model's ability to generalize to new data.

We applied the developed function detailed in Section 4.4 initially to the E-LSTM model, yielding the results shown in Fig. 9 which represent the average contribution for each feature in the model prediction sorted in a descendant way. This analysis was performed for 1000 random inputs taken from the test set. Upon analysis, it is evident that the charge temperature (TC) has a significant impact on the prediction, and that the model relies on the feature to make the prediction. Additionally, we assess the impact of bigger and smaller feature values on the prediction. This relationship is illustrated in Fig. 10. For instance, if the values of TD increase, the SOH will decrease, and vice versa. Similarly, if the IC increases, the SOH will increase, and the opposite is correct. This insight is based on our model's behavior and can provide valuable information regarding the impact of feature values on SOH estimation. The results reflect that in our dataset, the cycling process of cells is almost identical, differing only in the charging phase since the discharging phase is the same for all cells. This is why the features that most affect our models come from the charging phase. Regarding temperature, since the charge phase varies and the discharge phase occurs directly after the charge phase without sufficient resting time for the cell to return to normal temperature, the discharging temperature also has a significant impact. When visualizing the shape of each feature across all cells, you can observe similar current and voltage shapes. These only differ in their maximum and minimum values and the width, which represents the charge and discharge duration. This is not the case for temperature.

6.2. SOH abnormal decrease detection

Through the analysis, we successfully detected and plotted events that could potentially indicate a drop in SOH, as depicted in Figs. 11 and 13. When detecting events, we represent them in a simplified manner that allows us to gain additional insights into the events, aiding in the interpretation of the underlying factors influencing SOH changes. This representation also enables us to track changes in the values of each feature over time, providing a more comprehensive view of the events as they unfold. As illustrated in Fig. 12, each event is associated with data regarding current, voltage, and temperature. This association offers experts a straightforward way to analyze these data points and identify any abnormal behaviors that may lead to SOH degradation. By examining these features over time, experts can better understand the conditions and factors contributing to changes in SOH, facilitating more informed decisions and interventions.

Our function has demonstrated the capability to detect events that lead to SOH drop. This is achieved through the identification of frequent items by comparing a large number of untypical cells with a typical one.

However, the results are not optimal in detecting all events, as in Fig. 13. This limitation occurs because certain events are infrequent and represent rare occurrences in the dataset, making them challenging to detect accurately. Despite this, our approach provides valuable insights into the most common and recurring events that influence the SOH.

7. Conclusion

In this study, we presented an approach for SOH estimation using two models: the E-LSTM model, which combines an encoder with LSTM architecture, and the CNN–LSTM model. Both models were trained and tested on the MIT battery dataset. Our experiments demonstrated that both models yielded improved results compared to state-of-the-art methods, with the E-LSTM model outperforming the CNN–LSTM model.

Additionally, we proposed a method for identifying events that contribute to SOH degradation using the E-LSTM model in conjunction with a frequent pattern mining-based technique. This method leverages derivatives and correlation coefficients to effectively detect and visualize the factors responsible for SOH degradation. By plotting and

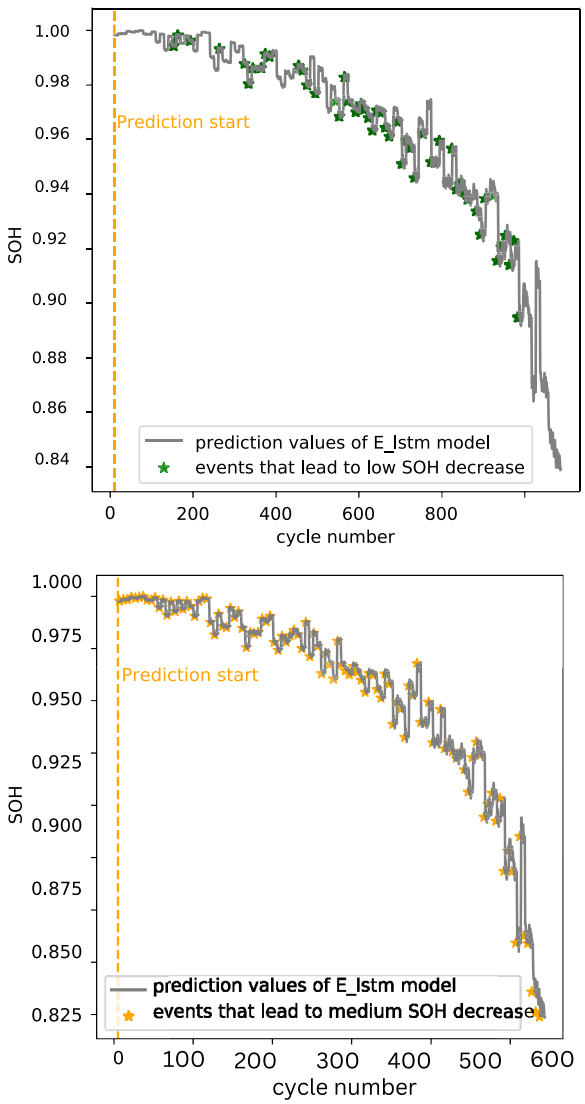


Fig. 11. SOH drop down events-(low-medium) for the same cell.

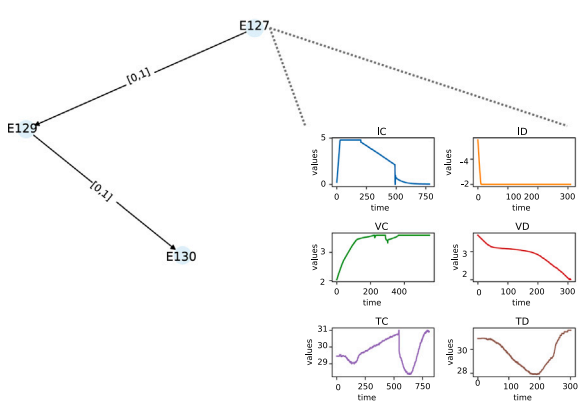


Fig. 12. Simple representation of the events and the data related to them.

providing a simple representation of these events, we gained insights into the underlying causes of SOH decline.

To enhance the explainability of our models, we developed a specialized function to calculate SHapley Additive exPlanations (SHAP) values. This approach, based on Strumbelj et al. [38] and using Monte

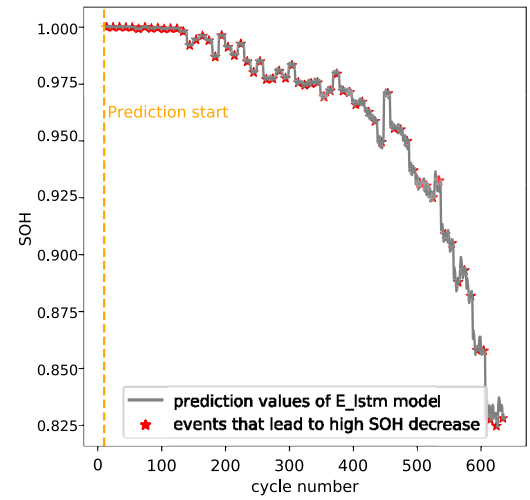


Fig. 13. SOH drop down events-high.

Carlo sampling, allowed us to approximate the SHAP value for each feature. By quantifying the contribution of each feature to the prediction, we gained valuable insights into the importance of various features in the context of battery performance data, aiding in decision-making and further model improvement.

Several interesting directions for future work emerged from this study. Firstly, we plan to incorporate real-world data into our experiments to move beyond online datasets, enhancing the robustness of our models. Secondly, we aim to improve our models to predict the complete SOH curve for a given cell using the growing window technique. These enhancements will increase the practicality and effectiveness of our SOH estimation models in real-world applications.

CRediT authorship contribution statement

Slimane Arbaoui: Writing – original draft, Visualization, Validation, Conceptualization. **Ahmed Samet:** Writing – review & editing, Supervision. **Ali Ayadi:** Writing – review & editing, Supervision. **Tedjani Mesbahi:** Writing – review & editing, Supervision. **Romuald Boné:** Writing – review & editing, Supervision.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Code availability

The code for implementing the approach described in this paper is accessible at https://github.com/arslimane/Energitic-project-1.git.

Data availability

Data will be made available on request.

Acknowledgments

This research received partial funding from the French National Research Agency (ANR) under the project “ANR-22-CE92-0007-02”. Additionally, support was provided by the European Union through the Horizon Europe program and the innovation program under “GAP-101103667”.

References

[1] Fit for 55. 2024, https://www.consilium.europa.eu/en/policies/green-deal/fit-for-55-the-eu-plan-for-a-green-transition/. [Accessed 07 June 2024].

[2] Zero emission vehicles: first 'Fit for 55' deal will end the sale of new CO2 emitting cars in Europe by 2035. 2024, https://ec.europa.eu/commission/presscorner/detail/en/ip_22_6462. [Accessed 07 June 2024].

[3] Why is the lithium-Ion battery the best battery?. 2024, https://energyx.com/blog/why-is-the-lithium-ion-battery-the-best-battery/. [Accessed 07 June 2024].

[4] Lee Junghwan, Sun Huanli, Liu Yongshan, Li Xue. A machine learning framework for remaining useful lifetime prediction of li-ion batteries using diverse neural networks. Energy AI 2024;15:100319.

[5] Ren Zhong, Du Changqing. A review of machine learning state-of-charge and state-of-health estimation algorithms for lithium-ion batteries. Energy Rep 2023;9:2993–3021.

[6] dos Reis Gonçalo, Strange Calum, Yadav Mohit, Li Shawn. Lithium-ion battery data and where to find it. Energy AI 2021;5:100081.

[7] Meng Huixing, Li Yan-Fu. A review on prognostics and health management (PHM) methods of lithium-ion batteries. Renew Sustain Energy Rev 2019;116:109405.

[8] Wen Jianping, Chen Xing, Li Xianghe, Li Yikun. SOH prediction of lithium battery based on IC curve feature and BP neural network. Energy 2022;261:125234.

[9] Gu Xinyu, See KW, Li Penghua, Shan Kangheng, Wang Yunpeng, Zhao Liang, Lim Kai Chin, Zhang Neng. A novel state-of-health estimation for the lithium-ion battery using a convolutional neural network and transformer model. Energy 2023;262:125501.

[10] Fan Yuqian, Li Yi, Zhao Jifei, Wang Linbing, Yan Chong, Wu Xiaoying, Zhang Pingchuan, Wang Jianping, Gao Guohong, Wei Liangliang. Online state-of-health estimation for fast-charging lithium-ion batteries based on a transformer–long short-term memory neural network. Batteries 2023;9(11).

[11] Feng Hailin, Shi Guoling. SOH and RUL prediction of Li-ion batteries based on improved Gaussian process regression. J Power Electron 2021;21:1845–54.

[12] Audin Paul, Jorge Inès, Mesbahi Tedjani, Samet Ahmed, Bertr François De, Beuvronand De, Boné Romuald. Auto-encoder LSTM for Li-ion SOH prediction : a comparative study on various benchmark datasets. In: 2021 20th IEEE international conference on machine learning and applications. ICMLA, vol. 8, 2021, p. 1526–36.

[13] Jorge Inès, Mesbahi Tedjani, Samet Ahmed, Boné Romuald. Time series feature extraction for lithium-ion batteries state-of-health prediction. J Energy Storage March 2023;59:106436.

[14] Liu Kang, Kang Longyun, Xie Di. Online state of health estimation of lithium-ion batteries based on charging process and long short-term memory recurrent neural network. Batteries 2023;9.

[15] Saxena Saurabh, Ward Logan, Kubal Joseph, Lu Wenquan, Babinec Susan, Paulson Noah. A convolutional neural network model for battery capacity fade curve prediction using early life data. J Power Sources 2022;542:231736.

[16] Wang Fujin, Zhao Zhibin, Ren Jiaxin, Zhai Zhi, Wang Shibin, Chen Xuefeng. A transferable lithium-ion battery remaining useful life prediction method from cycle-consistency of degradation trend. J Power Sources 2022;521:230975.

[17] Ardeshiri Reza Rouhi, Liu Ming, Ma Chengbin. Multivariate stacked bidirectional long short term memory for lithium-ion battery health management. Reliab Eng Syst Saf 2022;224:108481.

[18] Toughzaoui Yassine, Toosi Safieh Bamati, Chaoui Hicham, Louahlia Hasna, Petrone Raffaele, Masson Stéphane Le, Gualous Hamid. State of health estimation and remaining useful life assessment of lithium-ion batteries: A comparative study. J Energy Storage 2022;51:104520.

[19] Gong Dongliang, Gao Ying, Kou Yalin, Wang Yurang. State of health estimation for lithium-ion battery based on energy features. Energy 2022;257:124812.

[20] Nguyen Van Chi, Quang Duy Ta. Estimation of SoH and internal resistances of Lithium ion battery based on LSTM network. Int J Electrochem Sci 2023;18(6):100166.

[21] Chen Liping, Xie Siqiang, Lopes António M, Li Huafeng, Bao Xinyuan, Zhang Chaolong, Li Penghua. A new SOH estimation method for Lithium-ion batteries based on model-data-fusion. Energy 2024;286:129597.

[22] Chen Minzhi, Ma Guijun, Liu Weibo, Zeng Nianyin, Luo Xin. An overview of data-driven battery health estimation technology for battery management system. Neurocomputing 2023;532:152–69.

[23] Wang Fujin, Zhai Zhi, Liu Bingchen, Zheng Shiyu, Zhao Zhibin, Chen Xuefeng. Open access dataset, code library and benchmarking deep learning approaches for state-of-health estimation of lithium-ion batteries. J Energy Storage 2024;77:109884.

[24] Qi Guangheng, Ma Ning, Wang Kai. Predicting the remaining useful life of supercapacitors under different operating conditions. Energies 2024;17(11).

[25] Du Mengnan, Liu Ninghao, Hu Xia. Techniques for interpretable machine learning. Commun ACM 2019;63(1):68–77.

[26] Wang Fujin, Zhao Zhibin, Zhai Zhi, Shang Zuogang, Yan Ruqiang, Chen Xuefeng. Explainability-driven model improvement for SOH estimation of lithium-ion battery. Reliab Eng Syst Saf 2023;232:109046.

[27] Kim Sung Wook, Oh Ki-Yong, Lee Seungchul. Novel informed deep learning-based prognostics framework for on-board health monitoring of lithium-ion batteries. Appl Energy 2022;315:119011.

[28] Shi Junchuan, Rivera Alexis, Wu Dazhong. Battery health management using physics-informed machine learning: Online degradation modeling and remaining useful life prediction. Mech Syst Signal Process 2022;179:109347.

[29] LeCun Yann, Bengio Yoshua, Hinton Geoffrey. Deep learning. Nature 2015;521(7553):436–44.

[30] Yamashita Rikiya, Nishio Mizuho, Do Richard Kinh Gian, Togashi Kaori. Convolutional neural networks: an overview and application in radiology. Insights Imaging 2018;9:611–29.

[31] Hinton GE, Salakhutdinov RR. Reducing the dimensionality of data with neural networks. Science 2006;313:504–7.

[32] What are autoencoders? Introduction to autoencoders in deep learning. 2024, https://www.simplilearn.com/tutorials/deep-learning-tutorial/what-are-autoencoders-in-deep-learning. [Accessed 07 June 2024].

[33] Understanding variational autoencoders (VAEs). 2024, https://towardsdatascience.com/understanding-variational-autoencoders-vaes-f70510919f73. [Accessed 07 June 2024].

[34] Vincent Pascal, Larochelle Hugo, Bengio Yoshua, Pierre-Antoine. Extracting and composing robust features with denoising autoencoders. In: Proceedings of the 25 th international conference on machine learning, Helsinki, Finland. vol. 8, 2008, p. 1096–103.

[35] What happens in sparse autoencoder. 2024, https://medium.com/@syoya/what-happens-in-sparse-autencoder-b9a5a69da5c6. [Accessed 07 June 2024].

[36] Severson Kristen A, Attia Peter M, Jin Norman, Perkins Nicholas, Jiang Benben, Yang Zi, Chen Michael H, Aykol Muratahan, Herring Patrick K, Fraggedakis Dimitrios, Bazant Martin Z, Harris Stephen J, Chueh William C, Braatz Richard D. Data-driven prediction of battery cycle life before capacity degradation. Nat Energy 2019;4:383–91.

[37] A gentle introduction to k-fold cross-validation. 2024, https://machinelearningmastery.com/k-fold-cross-validation/. [Accessed 07 June 2024].

[38] Štrumbelj Erik, Kononenko Igor. Explaining prediction models and individual predictions with feature contributions. Knowl Inf Syst 2014;41(3):647–65.

[39] Fournier-Viger Philippe, Lin Jerry Chun-Wei, Kiran Rage Uday, Thomas Rincy. A survey of sequential pattern mining. Data Sci Pattern Recognit 2017;1:54–77.