Article

A Deep Learning Method for Health State Prediction of Lithium Ion Batteries Based on LUT-Memory and Quantization

Mohamed H. Al-Meer *

Posted Date: 3 January 2024

doi: 10.20944/preprints202311.1246.v2

Keywords: Lithium-Ion Batteries; SoH; SoC; RUL; Batteries; Deep Learning; LUT

Article

A Deep Learning Method for Health State Prediction of Lithium Ion Battery Based on LUT-Memory and Quantization

Mohamed H. Al-Meer

almeer@qu.edu.qa

Abstract: The precise determination of the State-of-Health (SOH) of lithium-ion batteries is critical in the domain of battery management systems. The proposed model in this research paper emulates any deep learning or machine learning model by utilizing a Look Up Table (LUT) memory to store all activation inputs and their corresponding outputs. The operation that follows the completion of training is referred to as the LUT memory preparation procedure. The lookup operation performed on this method simply substitutes for the inference process. This is achieved by discretizing the input data and features before binarizing them. The term for the aforementioned operation is the LUT inference method. The procedure was evaluated in this study using two distinct neural network architectures: a bidirectional long short-term memory (LSTM) architecture and a standard fully connected neural network (FCNN). It is anticipated that considerably greater efficiency and velocity will be achieved during the inference procedure when the pre-trained deep neural network architecture is inferred directly. The principal aim of this research is to construct a lookup table that effectively establishes correlations between the SOH of lithium-ion batteries and ensures a degree of imprecision that is tolerable. According to the results obtained from the NASA PCoE lithium-ion battery dataset, the proposed methodology exhibits performance that is largely comparable to that of the initial machine learning models. Utilizing the error assessment metrics RMSE, MAE, and (MAPE), the accuracy of SOH prediction has been quantitatively evaluated. The indicators mentioned above demonstrate a significant degree of accuracy when predicting SOH.

Keywords: Lithium-Ion Batteries; SOH; SoC; RUL; Batteries; Deep Learning; LUT

1. Introduction

The use of lithium-ion batteries has rapidly increased due to their low-cost, high-energy densities, low self-discharge rate, and long lifetime compared to other batteries [1–4]. Hence, lithium-ion batteries have gained significant prominence across diverse domains, including but not limited to mobile computing devices, aerospace applications, electric cars, and energy storage systems [5,6]. Despite the noteworthy advantages of lithium-ion batteries, a significant drawback is the occurrence of capacity fading upon repeated utilization. In addition, it is imperative to diligently observe and precisely assess the capacity, since an inaccurate evaluation of capacity might result in irreversible harm to the battery through excessive charging or discharging [7]. The assessment of battery capacity fade relies heavily on, what is called, the state of health or SOH, which serves as a pivotal indication. Hence, it is vital to precisely determine the SO) of lithium-ion batteries in order to ensure their safety and dependability [8].

Numerous research endeavors have been undertaken to get a precise assessment of the SOH. In general, the studies can be categorized into three distinct groups: model-based methods [9–14], data-driven methods [15,16], and hybrid methods [17,18].

The evaluation of battery health state, as determined by the SOH metric, was conducted based on the varying beginning capacities of each battery. There exist several approaches for determining the SOH of a battery. However, the predominant methods revolve around assessing the impedance and its useful capacity. The utilization of the battery's impedance as a measuring technique is deemed unsuitable for online applications due to the necessity of specialized instruments, for example the electrochemical-impedance-spectroscopy. Hence, most studies employed the SOH metric of a battery, which was determined by its usable capacity.

The subsequent sections of this article are structured in the following manner. Section 1 outlines the many contemporary techniques employed for forecasting the state of health (SOH) of batteries. Section 2 presents the employed approach and its corresponding context. Section 3 enumerates other relevant studies. Section 4 provides a comprehensive summary of the dataset utilized. Section 5 demonstrates the utilization of FCNN and LSTM training models. Section 6 of the paper focuses on the assessment of performance and the specific measurements employed. Section 7 presents the findings and outcomes of the study besides discussions. The last section is the conclusion.

Numerous studies have been undertaken to precisely assess the state of health (SOH). The methods can be categorized into three distinct groups: direct measurement (experimental) methods, model-based methods, and data-driven methods.

Starting with the experimental methods, the Cycle counting determines the age of a battery by analyzing its chronological charge and discharge cycles. The most straightforward approach to determine the SOH is by tallying the total number of battery cycles. During each cycle, the degradation of batteries is significantly influenced by characteristics such as charge and discharge depth temperatures, as well as the C rating. A coefficient can be derived to establish a connection between those characteristics and the full discharge (100% depth of discharge) [57].

Charge Counting: The charge counting method, also known as ampere-hour counting, is regarded as a precise and straightforward approach for measuring battery capacity. The ratio of the actual battery capacity (Qact) to the nominal capacity (Qnom) yields an accurate SOH. The measurement of the charge transferred throughout a complete charging or discharging process at a low C-rate and controlled temperature (usually 25°C) provides a precise assessment of the remaining capacity [57–59].

$$SOH = \left(\frac{Qactl}{Qnom}\right) \tag{1}$$

The internal resistance approach is based on the observation that the internal resistance (IR) of a battery increases as it ages. This property makes it a good parameter for estimating the SOH of the battery. To compute the IR, one can take a small step in the current and apply Ohm's law, as shown in equation 2. Here, ΔUt represents the voltage pulse, ΔU_t represents the current pulse [57,58].

$$R_0 = \frac{\Delta U_t}{\Delta U_L} \tag{2}$$

The Electrochemical Impedance Spectroscopy (EIS) method involves analyzing the impedance properties of a battery by subjecting it to a sinusoidal current voltage signal across a broad range of frequencies. Consequently, the response of current or voltage can be examined. A Nyquist plot can facilitate comprehension of the estimation. In addition, Warburg elements, in addition to inductors, can be used in an ECM to enhance the capture of intricate dynamics. At low frequencies, the impedance is essentially ohmic, and the significance of capacitive effects increases. The ohmic resistance exhibits a direct relationship with battery ageing, therefore making it a suitable indicator for SOH assessment [57]. Despite its accuracy, this system is deemed challenging to apply due to its high cost and complexity.

The Incremental Capacity Analysis Method (ICA) examines the rate of change of the charge Q in relation to the voltage V. The morphology of the IC curve undergoes alterations as it ages, rendering it a potential predictor of the SOH. As the battery degrades, the incline of the voltage curve during the constant-current charging phase intensifies. The primary drawbacks of the ICA technique are the requirement for a relatively low C-rate compared to typical charge rates and the computationally intensive numerical differentiation carried out by the BMS.

The differential voltage analysis method (DVA) [56,64] shares similarities with the independent component analysis (ICA) method. The rate of change of voltage, denoted as V, with respect to the charge is calculated during the constant current (CC) charging phase. The peaks in question correspond to specific chemical reactions within the cell that persist independent of battery ageing.

The gap between these peaks indicates a segment of the battery's capacity that can be utilized for estimating the SOH.

The inside examination of a battery cell, without causing any harm, can be accomplished by employing ultrasonic inspection or X-ray techniques. Both tasks can be executed manually during maintenance in order to obtain an approximation of the SOH. The idea relies on the sensitivity of ultrasonic wave propagation in liquid-filled porous media to several factors such as electrode tortuosity, porosity, thickness, density, elastic modulus, fluid density, and ion concentration [60,61]. A more thorough comprehension of the ageing process in different types of batteries has been achieved using techniques such as X-ray computed tomography [62] and others [63]. Characterizing the internal information with accuracy can be a difficult undertaking; the fundamental ageing mechanisms must be deduced from the examination of exterior signals. Several research integrate ultrasonic examination and machine learning to accurately evaluate the SOH.

The battery SOH estimation can be achieved by a number of model-based approaches, which entail building a model of a battery and including the internal deterioration process inside the model. Knowledge of the internal battery reaction mechanisms, the accurate formulation of the mathematical equations driving these reactions, and the development of effective simulation models are all essential for model-based approaches. There may be difficulties in putting these ideas into practice. Lithium-ion battery life was predicted using a weighted ampere-hour throughput model in a study by the author in [19]. The researcher determined how severely the batteries had been damaged. To determine a battery's SOH, the author of [20] applied an analog circuit model and an improved Kalman filter. To determine a battery's SOH. In a study presented by in [21], authors relied on nominal filtering methods. They discovered that battery capacity is proportional to the square of the internal impedance. The researchers in [22] presented a method for estimating a battery's SOH, or state of health. This method takes battery discharge rates as input variables into a state-space model. The SOH of a battery was predicted using a single-particle model developed by a study [23] that is grounded on the physics of electrolytes. The degeneration of the internal mechanisms and the batteries were the focus of this investigation. From a chemical perspective. The construction of a precise aging model for lithium-ion batteries is difficult due to the complex structure of the chemical interactions occurring within the battery, although model-based approaches are useful in forecasting the SOH of batteries. Furthermore, various environmental variables, such as operational temperature, anode and cathode materials, and related parameters, have a major impact on the performance of lithium-ion batteries. Therefore, there are difficulties in establishing a reliable aging model for lithium-ion batteries.

Improvements in hardware have allowed computers to do increasingly complex mathematical operations in recent years. Concurrently, there has been an uptick in the number of databases that can be mined for information on a battery SOH thanks to the growing use of data-driven methodologies for this purpose. This paved the way for the broad adoption of data-driven approaches. Even without a thorough understanding of the battery's internal structure and aging mechanisms from an electrochemical standpoint, a data-driven method allows for reliable prediction of the battery's SOH. Therefore, these methods can be implemented with little to no familiarity with a battery's electrochemical properties or the environmental context. The success of data-driven approaches relies greatly on the accuracy and usefulness of the information that is collected. Acquiring attributes that are highly linked with the degradation process is crucial to the success of a data-driven model. In a related study found in [24], authors suggested using sparse Bayesian predictive modeling with a sample entropy metric to boost the reliability of voltage sequence predictions. The SOH of a battery can be estimated and analyzed with the help of a hidden Markov model, which was introduced by the authors in a study [25]. In different research [26], authors used Gaussian process regression to estimate health status in their investigation. This method combines covariance and mean functions into a single estimate. Another work [27] presented a data-driven prognosis method that makes use of deep neural networks to foretell the SOH and RUL of lithium-ion batteries. In research, the authors [28] developed a method for adaptive SOH estimate based on an online alternating current (AC) complex impedance and a Fully Connected neural network (FCNN). The use of recurrent neural networks to foresee battery performance decline has been demonstrated by a work in [29]. In this work [30], the authors proposed using a deep convolutional neural network (CNN) trained on recorded current and voltage to estimate battery capacity.

In this work, we develop an efficient processing method to replace the inference operation through the trained network with a pre calculated LUT memory. The accuracy and efficiency of SOH estimation by using this substitute proved to be acceptable. This method is completely non-using the different machine learning models trained, but solely relies on the quantization operation and chosen bits per feature.

2. Proposed Methodology

The objective of machine learning is to identify and construct an appropriate model based on the provided training samples in order to establish the relationship between the training data and the target output, denoted as y. In this context, x represents the charge voltage of charging curves, N is the of training samples number, and y denotes the SOH of a Li-ion battery. The nonlinear mapping f(·) can be defined as follows:

$$y_i = f(x_i) \tag{3}$$

Prior to proceeding, it is important to establish a function capable of discretizing the characteristics of the neural network model inputs into a limited range of values. The desired quantization process involves converting real numbers represented in floating point format into a narrower range of lower precision values. One commonly used option for a quantization function is explained in the next equation.

$$Q(r) = Int\left(\frac{R}{S}\right) - Z \tag{4}$$

The quantization operator, denoted as Q, operates on a real-valued input (activation or weight) represented by the variable r. The scaling factor, denoted as S, is also a real-valued variable. Lastly, the integer zero point is represented by the variable Z. In addition, the Int function converts a real value to an integer value by performing a rounding operation, such as rounding to the closest integer or truncating the decimal part. Essentially, this function is establishing a correspondence between real values r and integer values. The technique of quantization described here is commonly referred to as uniform quantization, as it produces quantized values (also known as quantization levels) that are evenly distributed.

Dequantization is the process by which real values, denoted as r^, can be recovered from quantized values, denoted as Q(r). It should be noted that the true values that have been recovered, denoted as r^, will not perfectly match the original values due to the rounding operation.

$$r^{\wedge} = S(Q(r) + Z) \tag{5}$$

The selection of the scaling factor S is a crucial aspect in uniform quantization. This scaling factor essentially divides a specified range of real values, denoted as r, into a specific number of partitions.

$$S = \frac{\beta - \alpha}{2^b - 1} \tag{6}$$

The clipping range, denoted by [α, β], is a bounded range used to restrict the real values. The quantization bit width is represented by b. In order to determine the scaling factor, it is necessary to first establish the clipping range. One simple option is to use the signal's min/max for the clipping range, that is, $\alpha = r_{min}$ and $\beta = r_{max}$. This method employs an asymmetrical quantization strategy.

Symmetric quantization is a widely used technique for quantizing weights and biases in neural networks [65] and also makes the implementation more straightforward. Also, quantization exhibits a noticeable error called the quantization error. The following formula will be used to determine the power of the quantization noise:

$$E[v^2] = \frac{q^2}{12} \tag{7}$$

From this, one may extract the formula for calculating the Signal-to-Quantization-Noise Ratio, also known as SQNR.

$$SQNR = 20log_{10}\,(2^Q) \mathrel{\hat{\approx}} 6.02\,.\,Q\ dB \tag{8}$$

where Q is the number of quantization bits. This study will incorporate quantization to convert continuous numerical values belonging to features into discrete digital representations in the form of binary numbers. The range of bits transformed will span from 2 to 8 bits. Table 1 shows the quantization bits distribution per feature and corresponding SQNR level and total memory size needed. When the number of bits rises, the memory size needed gets larger, but the SQNR gets more improved.

In this study, we will utilize a neural network model to acquire knowledge about nonlinear mapping. Both the Fully Connected Neural Network (FCNN) and Long Short-Term Memory (LSTM) have been utilized as learners for nonlinear mapping. The decision was clearly made after doing thorough testing, and it is possible to use any machine learning approach or other types of neural network models. The proposed method emulates the correlation between inputs and outputs.

Table 1. Quantization bits distribution per feature and corresponding SQNR level and total memory size needed.

Bits /	Values	Bits Total	SQNR	Memory
Feature	given	(Address)	dB	Size
2	4	14	12.04	16K
3	8	21	18.06	2M
4	16	28	24.08	256M
5	32	35	30.10	32G
6	64	42	36.12	4T
7	128	49	42.14	----
8	256	56	48.16	----

In this study, we will utilize a neural network model to acquire knowledge about nonlinear mapping. Both the Fully Connected Neural Network (FCNN) and Long Short-Term Memory (LSTM) have been utilized as learners for nonlinear mapping. The decision was clearly made after doing thorough testing, and it is possible to use any machine learning approach or other types of neural network models. The proposed method emulates the correlation between inputs and outputs.

2.1. LUT-Memory Creation and Usage

The chosen machine learning architecture is firstly prepared by selecting the seven features input, which are prepared in the dataset as 7X1 format. It is chosen to have seven inputs according to this constraint. The input and output variables of the estimation model need to have a certain correlation. This study takes the next measurements of the battery (capacity, ambient temperature, date-time, measured volts, measured current, measured temperature, load voltage, and load current), as the input variable of the model. The models chosen to be tested are FCNN and LSTM models. Figure 1 is the architecture of the LUT estimation system, which consists of LUT generation and inference modules. The LUT generation module consists of Integer address generation, bits grouping, Digital to Analog Conversion (DAC), features normalization, model training, then finally LUT filling. The LUT inference module consists of features preparation and normalization, binarization, binary address bits binding, inferring the LUT memory, and finally analysis results calculation. The entire process can be described as follows:

2.2. LUT Generation:

- **Step 1**. The address bit combination starts from 0 up to 2^{7n}. The binary address generated depends highly on the number of bits assigned to each of the seven features. The ones that will be tested are 2, 3, 4, 5, 6, 7, and 8 bits. Table 3 shows the details.
- **Step 2**. Then the generated address bits are grouped into seven feature groups, while each feature owns its own number of bits, generating a feature binary address bit.
- **Step 3**. The address bit value for each feature is normalized as bits value / 2^n, where n is the number of bits selected for the feature.
- **Step 4**. The seven normalized feature values are presented to the trained deep neural network.
- **Step 5**. The value inferred from the model is stored in the LUT memory at the given address.
- **Step 6**. Then the next address is selected, and the whole operation is repeated (going to step 1).

The execution of the inference procedure involves accessing memory after converting the continuous values of the features into binary address bits. This process can be described as follows:

2.3. LUT Usage

- Step 1. Starting from the seven feature values (capacity, ambient temperature, date-time, measured volts, measured current, measured temperature, load voltage, and load current)
- Step 2. Each of the seven feature values will be normalized (0, 1).
- Step 3. Then those will be quantized based on the next configurations: 2 bits, 3 bits, 4 bits, 5 bits, 6 bits, and 8 bits, depending on the adaptation.
- Step 4. Quantization produces the binary bits for each feature.
- Step 5. Combining all bits into one address, as shown in Figure 3,

In general, the methodology utilized in this study is founded on the application of pre inference process. It is evident that the process will require a total of 2n7 iterations to go through all combinations of features. The aim is to populate LUT memory with the range of values that are obtained through the process of approximation and quantization. These values are derived from the inference process of the selected model. The process of converting the inference operation results in a significant boost in speed, mostly attributed to the simplicity of the conversion and the use of less complex hardware for implementation. Significantly, this transformation is not contingent upon the existence of the trained model.

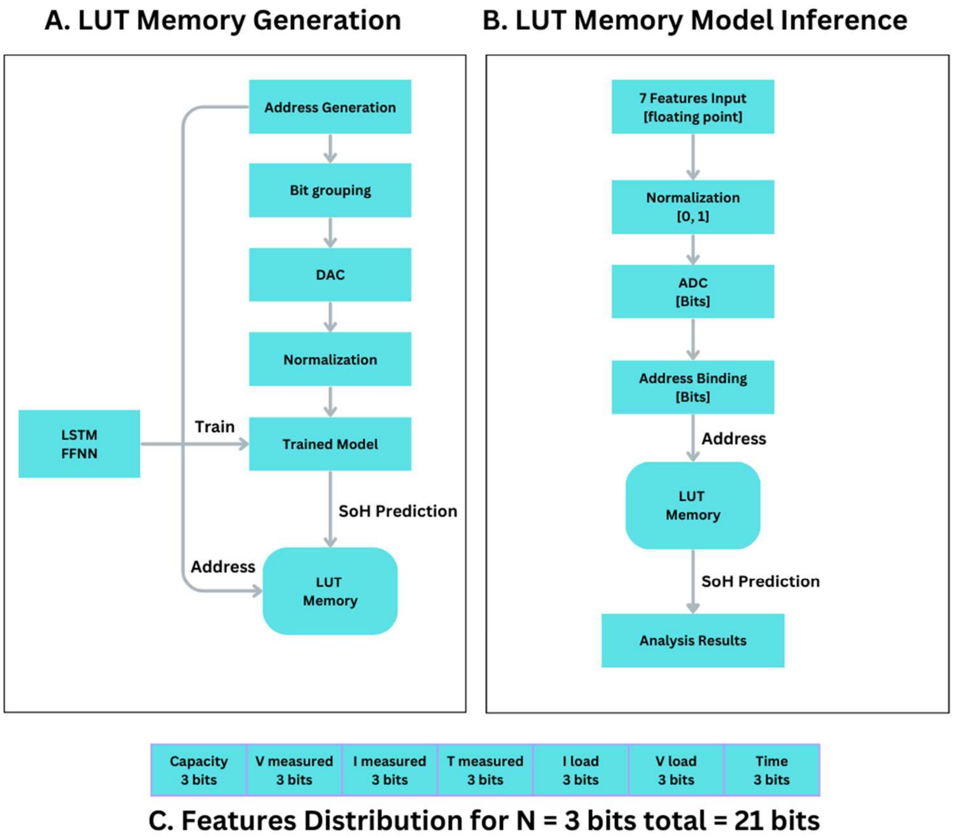


Figure 1. The new algorithm flowchart and features-bit distribution. A. LUT Memory Generation, and B. LUT Memory Inference.

3. Related Works on Quantization in DNN

The 2-bit quantization of activations, a uniform quantization [35–38] convert full-precision quantities to uniform levels of quantization (0; 1/3; 2/3; 1). It was initially suggested in research conducted by the authors of [39] that the weights be quantified using only the binary values -1 and +1. Then, trained ternary quantization (TTQ) [40] was implemented to symbolize the weights using the ternary values -1, 0, and +1. Scale parameters are employed in research found in [41] to binarize the activations; however, they remain full-precision values when solving the activations in these networks. It is noteworthy to notice that a majority of uniform quantization techniques employ linear quantizers, namely utilizing the *round()* function, in order to uniformly quantize floating-point numbers.

In numerous instances (e.g., quantization of activations), non-uniform quantization has been implemented to align the quantization of weights and activations with their respective distributions. It is suggested in a work implemented to fit a combination of Gaussian prior models with cluster centroids serving as quantization levels, as opposed to weights [42]. In accordance with this concept, another work implemented layer-wise clustering in order to convert weights into cluster centroids [43]. Logarithmic quantizers were implemented in the two works of [44] and [45] to symbolize the weights and activations with powers-of-two values.

Despite the fact that Deep Neural Networks (DNNs) are extensively employed in real-time applications, they are computationally intensive tasks, consisting primarily of linear computation operators, which place a tremendous strain on the constrained hardware resources. Considerable effort and research have been devoted to developing a cost-effective and efficient DNN inference algorithm. Model compression [46,47], operator optimization for sophisticated computation [48–51], tensor compilers for operator generation [52,53], and customized DNN accelerators [54–56] are a few

doi:10.20944/preprints202311.1246.v2

of these techniques. In order to accommodate various deployment conditions, these methods necessitate the repetitive redesign or reimplementation of computation operators, accelerators, or model structures.

In contrast to the aforementioned approaches, this article investigates a novel possibility for substituting computation operators in DNNs in order to reduce inference costs and the laborious process of operator development. LUT, a novel system that enables DNN inference via table search, is proposed in order to address this inquiry. The system computes the LUT in accordance with the learned typical features after traversing every permutation for each of the seven SOH features.

4. Dataset Description

The information utilized in this work for the purpose of estimating the lithium-ion batteries SOH was acquired from the NASA Prognostics Center [31]. The four batteries used, indicated as #5, #6, #7, and #18, have been extensively utilized for the purpose of SOH estimate [32–34]. The dataset utilized in this study comprised operational profiles and impedance measurements of 18,650 lithium-ion batteries measurement trials throughout the processes of charging and discharging at ambient temperature (around 24C). The batteries underwent a charging process where a 1.50 A current was maintained in a continuous flow until the voltage reached 4.2 V. Subsequently, the charging proceeded at 4.2 V of constant voltage until the current decreased to a value below 20 mA. Battery numbers 5, 6, 7, and 18 were discharged at a constant current of 2 A until their voltages hit 2.7 V for B0005, 2.5 V for B0006, 2.2 V for B0007, and 2.5 V for B0018 batteries. This process took place until the batteries' individual voltages reached the desired levels. Eventually, the batteries gradually start to decrease in capacity as the number of charging and discharging continues to grow. It is taken that, once the battery dropped 30% from the rated capacity, its end-of-life (EOL) is reached, nominally from 1.4Ah to 2.0 Ah.

The aforementioned collected datasets provide the capability to predict the batteries SOH. The battery's aging process is correlated with the number of cycles, as seen in Figure 2. Additionally, Table 2 illustrates the features value boundaries observed throughout an aging cycle for battery #5.

Table 2. Boundaries of the measured features for B0005 Battery.

	Capacity	V_m	I_m	T_m	I_{Load}	V_{Load}	Time (s)
Min	1.28745	2.44567	-2.02909	23.2148	-1.9984	0.0	0
Max	1.85648	4.22293	0.00749	41.4502	1.9984	4.238	3690234

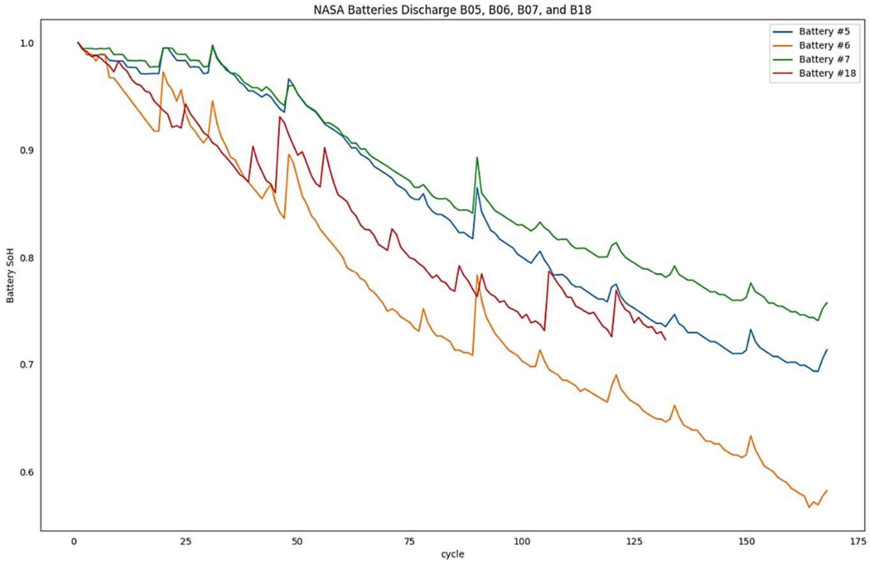


Figure 2. Degradation of lithium-ion batteries as a function of cycle count.

The primary indication of battery deterioration is the decline in capacity, which is mostly associated with the SOH. The definition of SOH is determined by its capacity, which may be calculated easily using the next formula. In this formula, Cusable and Crated denote the actual and notional capabilities, correspondingly. This study we utilize lithium-ion battery time series data for the purpose of predicting the SOH. The research also delves into the essential process of data preprocessing in order to ensure accurate and reliable predictions.

$$SoH = \frac{C_{Usable}}{C_{Rated}} \tag{9}$$

To illustrate more, the term "C_{usable}" refers to the usable capacity of a device, representing the maximum amount of capacity that may be released when it is entirely discharged. On the other hand, "C_{rated}" denotes the rated capacity, which is the capacity value supplied by the manufacturer. The available capacity diminishes as time progresses.

Data normalization is a widely utilized technique in depth modeling methods, as it is seen suitable for enhancing both the convergence of the model and the accuracy of the prediction. Normalization will be conducted using the minimum-maximum approach, which involves scaling the data within the range of 0 to 1. The aforementioned relationship is mathematically represented by the following equation. In the given context, the symbol "x" is used to denote the processed data, while "x" represents the original data. Additionally, "x_{max}" and "x_{min}" are used to indicate the maximum and minimum values of the original data, respectively.

$$x_n = \frac{x - x_{min}}{x_{max} - x_{min}} \tag{10}$$

5. Background and Preliminaries

In this study, the Fully Connected Neural Network FCNN and the LSTM models are proposed to undergo attesting for the new method. The model structure of FCNN and LSTM is presented in Table 3, which primarily consists of 5 layers for the FCNN and 9 layers for the LSTM model. The quantization concept besides its generated noise (error) will be presented as well. The internal structure of both models is described in detail as follows:

5.1. Fully Connected Deep Neural Network

The FCNN is a type of artificial neural network characterized by its acyclic graph structure. It is considered to be the most basic and straightforward form of neural network. The FCNN is a crucial component in machine learning, serving as the foundation for various architectures. It has multiple layers of neurons, each incorporating a nonlinear activation function.

A neural network comprises three distinct layers: the input layer, the hidden layer, and the output layer. Hidden layers are a set of intermediary layers positioned between the input and output layers in a neural network. These hidden layers are established by adjusting the parameters of the network. In the case of standard rectangular data, it is commonly observed that the utilization of 2 to 5 hidden layers is typically adequate. The quantity of nodes included in each layer is contingent upon the quantity of characteristics inside the dataset; however, there exists no rigid guideline governing this relationship. The computational load of the model is influenced by the quantity of hidden layers and nodes; hence the objective is to identify the most parsimonious model that exhibits satisfactory performance. The output layer generates the intended output or forecast, and its activation is contingent upon the specific modeling methodology employed. In regression problems, it is common for the output layer to consist of a single node that is responsible for generating continuous numeric predictions. Conversely, in binary classification issues, the output layer normally comprises a single node that is utilized to estimate the probability of success. In the case of multinomial output, the output layer of a neural network consists of a number of nodes that corresponds to the total number of classes being predicted. Overall, layers and nodes play a crucial role in determining the complexity and performance of neural network models.

5.2. Long Short-Term Memory (LSTM) Deep Neural Network

The Long Short-Term Memory (LSTM) is a distinct variant of the recurrent neural network (RNN) architecture, which facilitates the utilization of outputs from previous time steps as inputs for the purpose of processing sequential data. One notable distinction between LSTM (Long Short-Term Memory) and simple RNN (Recurrent Neural Network) lies in the conditioning of the weight on the self-loop. Unlike simple RNN, LSTM incorporates contextual information to dynamically adjust the weight on the self-loop. The model effectively captures and preserves enduring relationships among sequential input data, such as time series, text, and speech signals. LSTM models employ memory cells and gates to effectively control the information flow, enabling the selective retention or removal of information as required. LSTM networks consist of three distinct types of gates, namely the input, forget, and output gates. The input gate controls the data stream towards the memory cell, and the other gate, the forget gate, controls the extraction of data from the memory cell. Additionally, the output gate is responsible for controlling the transmission of information from the LSTM unit to the output. The construction of these gates involves the utilization of sigmoid functions, and their training is accomplished through the process of backpropagation. The gates within a Long Short-Term Memory (LSTM) model exhibit dynamic behavior by modulating their openness or closure in response to the current input besides the previous hidden state. This adaptive mechanism enables the model to effectively choose whether to retain or discard information. The cells of a Long Short-Term Memory (LSTM) network are interconnected in a recurrent manner. The accumulation of input values computed by a standard neuron unit can occur within the state, provided that the input gate permits it. The state unit is equipped with a self-loop that is regulated by the forget gate, while the output gate has the ability to inhibit the output of the LSTM cell. Figure 3 shows a general diagram for the FCNN model and LSTM models, while Table 3 lists the structural component of the FCNN and LSTM models used in this study.

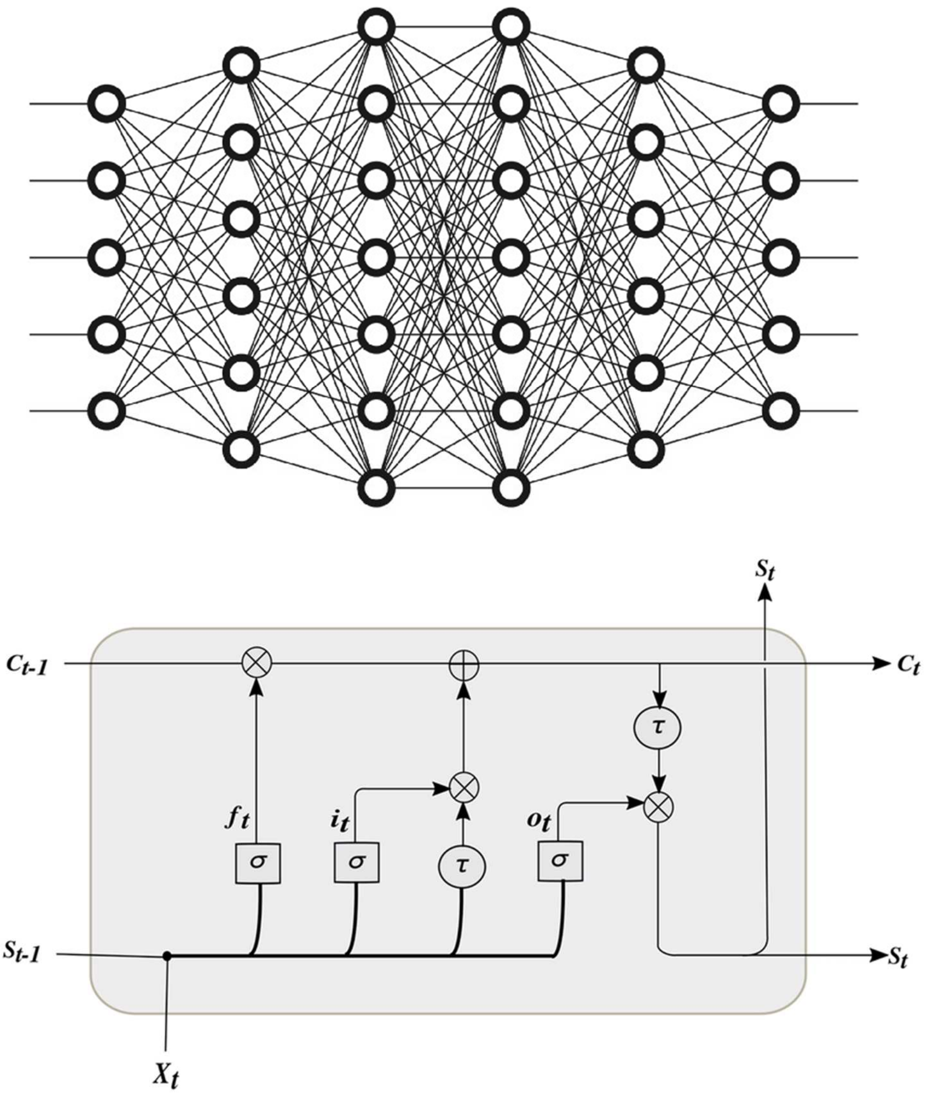


Figure 3. (Upper) Deep FCNN and (Lower) LSTM State of a layer

Table 3. Training Model Specification.

	Layers	**Output Shape**	**Parameters No.**
Model 1 **FCNN**	Dense	(node, 8)	217
	Dense	(node, 8)	
	Dense	(node, 8)	
	Dense	(node, 8)	
	Dense	(node, 1)	
	Layers	**Output Shape**	**Parameters No.**
Model 2 **LSTM**	LSTM 1	(N, 7, 200)	1.124 M
	Dropout 1	(N, 7, 200)	
	LSTM 2	(7, 200)	
	Dropout 2	(N, 7, 200)	

LSTM 3	(N, 7, 200)
Dropout 3	(N, 7, 200)
LSTM 4	(N, 200)
Dropout 4	(N, 200)
Dense	(N, 1)

6. Performance Evaluation and Metrics

The following is a description of the experimental setup used in this work. Hardware specifications include a 64-bit operating system, x64-based processor, Intel(R) Core (TM) i5-9400T CPU @ 1.80GHz running at 1.80 GHz, and 8 GB RAM. Kaggle Notebooks is used to enables and explore and run machine learning code with a cloud computational environment based on Jupyter that enables reproducible and collaborative analysis. Python 3.7 was the main programming language used.

Four datasets are utilized to predict the SOH of lithium-ion batteries, one for training, the DS0005, and the other three for validity of prediction use, DS0006, DS0007, and DS0018. In details, the datasets are partitioned into one subset as a training besides validation, and three as testing sets. The purpose of the training set is to facilitate the training process of the model, while the validation set is used to fine-tune the model's parameters. Lastly, the testing set is employed to evaluate the performance of the model. In order to ensure that the model's prediction is accurate, a number of suitable hyperparameters must be chosen. We use grid search and cross-validation to obtain optimal parameters for model performance.

6.1. Performance Evaluation Indicators

This work utilizes three error evaluation metrics, namely Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE). These metrics are employed to quantitatively analyze the accuracy of the proposed State of Health (SOH) prediction model and its Quantized Approximators (QA). The definitions of these metrics are provided as follows:

$$RMSE = \sqrt{\frac{1}{N}\left[\sum_{i=1}^{N}(y_i - y_i\hat{}\,)^{\,2}\right]} \tag{11}$$

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|y_i - \hat{y}_i\,| \tag{12}$$

$$MAPE = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{y_i - y^{\hat{}}{}_i}{y_i}\right| X\ 100\% \tag{13}$$

In this context, y_i represents the actual SOH value, while $y^{\hat{}}{}_i$ represents the expected (estimated) SOH value. In relation to metrics such as RMSE, MAE, and MAPE, the predictive accuracy increases as these indicators approach zero.

6.2. Models Training

In preparation for the model training phase, we split the training dataset into training and testing data. We tested the SOH dataset with two different ML models, including one with a FCNN and one with an LSTM. The tabulated model architectures make up Table 4. In the dataset used for training, which was divided into training and validation data, the training data to validation data ratio was found to be 2:1. The simulation was implemented with Kaggle Environment based on Jupyter and Python 3.7. All machine learning models were performed five times, and the average value of the results was used.

Table 4. Training Parameters for B0005 Battery Dataset.

Model	Batch size	Epochs	Time(s)	Loss
FCNN	25	50	200	0.0243
LSTM	25	50	7453	3.1478E-05

Table 4 shows the batch size, the epochs, the time, and the loss functions emerged from the training process. It is clearly realized that because of the large size of the LSTM model, it's time for convergence is so large when compared with the simple FCNN model.

7. Evaluation Results and Discussion

The efficiency of the suggested methodology in extracting findings was demonstrated by adapting FCNN and LSTM learning models for training and inference. We compare the original prediction from the trained model in both forms, the FCNN and the LSTM, with the LUT prediction for different quantization bits and show the results. To compare and show the difference between the true SOH and its estimated value utilizing the two neural network learning models, SOH estimation executed for the FCNN and LSTM learning models, as shown in Table 5. Training uses B0005 and prediction uses B0006, B0007, and B0018 batteries. In Table 5, batteries tested without quantization using actual model validation. Figure 4 shows that SOH estimation follows the same pattern as real SOH with a little deviation due to training on a different battery than tested.

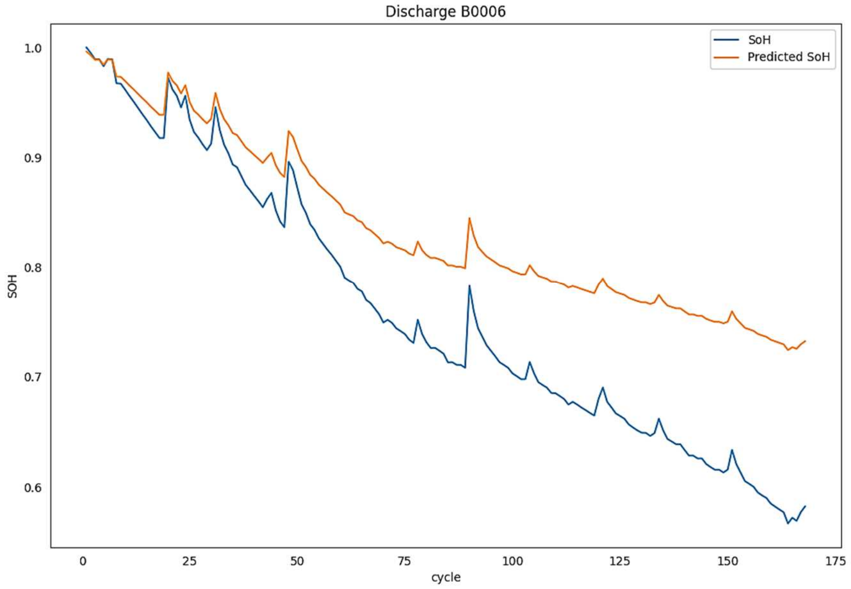


Figure 4. Real B0006 battery SOH and its estimated SOH. The Model trained on B0005 Battery.

Table 5. Estimation errors results for FCNN training model when trained on B0005 Battery and tested on the others using actual model validation (testing) without quantization.

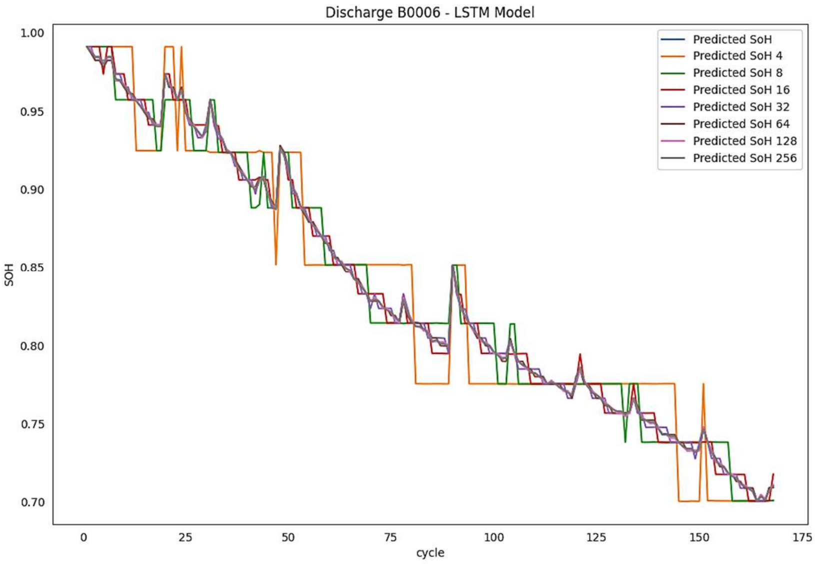
Battery	Model	RMSE	MAE	MAPE
B0006	FCNN	0.080010	0.068220	0.100970
	LSTM	0.076270	0.067620	0.098770
B0007	FCNN	0.019510	0.018019	0.021460
	LSTM	0.029282	0.024710	0.030434
B0018	FCNN	0.015680	0.013610	0.016890

LSTM	0.018021	0.016371	0.020547

Another setup of SOH predictions for the different batteries at different quantization levels was executed to illustrate the accuracy of the suggested quantization on the estimated SOH. Table 6 displays relevant results. RMSE, MAE, and MAPE are used as error evaluation measures. The table shows SOH predictions from deep learning models for batteries B0006, B0007, and B0018 with different quantization bits assigned. Figure 5 shows SOH prediction without and with quantization for all bits, as b = 2, 3, 4, 5, 6, 7, and 8 bits when tested on both models, the FCNN and the LSTM.

To supplement the preceding results, a visual comparison between the quantized and real inferences was prepared. Figure 6 illustrates the comparison between the original estimated SOH and its quantized counterparts for 2 bits and 5-bit.

Initially, the evaluation of the training models will be conducted without incorporating quantization operations. This approach aims to provide an understanding of the outcomes in their unaltered state, prior to exploring the impact of quantization. The tight correspondence between the SOH and the observed SOH for battery B0006 is shown in Figure 4. This phenomenon is expected to be applicable to all other batteries, despite the lack of empirical evidence, as has been observed. According to the findings presented in Table 3, it is evident that the two training models exhibit significant differences in terms of elapsed training time, although having identical batch sizes and epochs. The rationale behind this observation stems from the architectural differences between LSTM and FCNN models. Specifically, the LSTM model possesses a significantly larger building structure, consisting of over 1.14 million parameters that need to be mathematically computed. In contrast, the FCNN model exhibits a much smaller parameter count, totaling only 217.



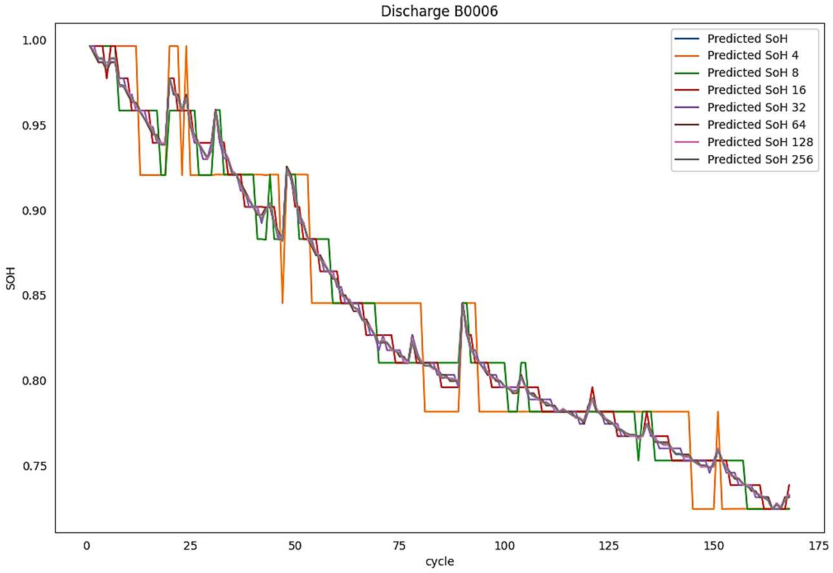


Figure 5. SOH predictions for B0006 Battery Dataset trained on B0005 Battery dataset showing original SOH (A. above) and all quantized bits for FFN and LSTM models (B. bottom).

Table 5 presents the outcomes of the estimation error analysis for the FCNN training model. The model was trained on the B0005 battery and subsequently evaluated on different batteries using actual model validation, without employing quantization. Divergent error estimations are detected across the two training models across different battery types. In the B0006 battery, the root mean square error (RMSE) of the FCNN model exhibited a superior performance compared to the long short-term memory (LSTM) model by a margin of 0.003, equivalent to a relative improvement of 4.6%. In the case of the B0007 battery, the root mean square error (RMSE) exhibited a difference favoring the LSTM model by 0.0097, which corresponds to a 50% improvement. In the case of the B0018 battery, both the FCNN and LSTM models exhibited similar performance, with a slight advantage of 0.0023 (15%) in favor of the FCNN. The mean absolute error (MAE) metric revealed a significant disparity in performance for the B0007 battery, with an observed value of 0.0067 (37%). In contrast, the remaining batteries had similar scores in this regard. In terms of the MAPE error metric, the B0006 battery exhibited the smallest disparity with a score of 2.1%, whilst the FCNN model demonstrated the biggest disparity with a score of 42%.

Next, we will commence the examination of the impact of quantization on the SOH forecasts for various batteries across different quantization levels and analyze its influence on the accuracy of these predictions. Based on the findings presented in Table 6, it is evident that the accuracy of SOH prediction utilizing quantization approaches diminishes as the number of quantization bits increases, in comparison to the original non-quantized model's SOH prediction. To provide additional support for this assertion, Figure 7 exhibits the three-error metrics in relation to the number of quantization bits for the B0006 batteries when subjected to training using FFN and LSTM models. The results demonstrate that even when employing a limited number of quantization bits, the magnitude of mistake remains quite small. Notably, the errors tend to become inconsequential after utilizing a quantization of n = 5 bits. Table 6 provides compelling evidence for the identification of distinct patterns. The findings indicate that the variations in battery types and the training models, FFN or LSTM, are inconsequential, with the primary distinguishing factor being the quantization bits.

Table 6 provides information on the B06 battery test, which is based on the dataset of the FCNN model and further approximated using the LUT. The following findings can be observed: The root mean square error (RMSE), mean absolute error (MAE), and mean absolute percentage error (MAPE) measurements begin at 0.0195370, 0.0159236, and 0.0190499, respectively, when b = 2 bits. They conclude at 0.0003125, 0.0002565, and 0.0003088, respectively, when b = 8 bits. The ratios will be 65 X, 61 X, and 61 X, respectively. The LSTM model demonstrates consistent findings with the RMSE, MAE, and MAPE error measurements. At b = 2 bits, the initial values for these measurements are 0.0216045, 0.0185078, and 0.0225291, respectively. At b = 8 bits, the final values for these measurements are 0.0003309, 0.0002835, and 0.0003446, respectively. The ratios will be 65 X, 66 X, and 65 X, respectively. An identical computation can be performed for the remaining batteries.

Figure 5 presents the predictions of the SOH without quantization, as well as with quantization for varying bit values, specifically 2, 3, 4, 5, 6, 7, and 8 bits. However, Figure 6 illustrates that the deviation from the initial prediction (before quantization) becomes insignificant when utilizing 5 or more bits. This implies that a minimum of 5 bits is required in order to get a distinct correspondence with the prediction.

Table 6. Prediction results after comparing between SOH of Non-Quantized and the Quantized versions based on different models for batteries B0006, B0007, and B0018.

Battery	Model	Quantization Bits	RMSE	MAE	MAPE (%)
B0006	FCNN	2	0.0195370	0.0159236	0.0190499
		3	0.0098006	0.0080317	0.0096645
		4	0.0046815	0.0037988	0.0045664
		5	0.0024301	0.0020093	0.0024294
		6	0.0012535	0.0010379	0.0012461
		7	0.0006150	0.0005068	0.0006144
		8	0.0003125	0.0002565	0.0003088
	LSTM	2	0.0216045	0.0185078	0.0225291
		3	0.0104658	0.0088477	0.0107360
		4	0.0050010	0.0042487	0.0051737
		5	0.0025885	0.0022293	0.0027206
		6	0.0013394	0.0011620	0.0014114
		7	0.0006609	0.0005692	0.0006974
		8	0.0003309	0.0002835	0.0003446
B0007	FCNN	2	0.0187614	0.0162685	0.0191451
		3	0.0101181	0.0088282	0.0103004
		4	0.0050026	0.0043651	0.0051114
		5	0.0024498	0.0021127	0.0024730
		6	0.0012030	0.0010481	0.0012269
		7	0.0006394	0.0005566	0.0006533
		8	0.0003060	0.0002578	0.0003013
	LSTM	2	0.0209633	0.0181984	0.0219105
		3	0.0113147	0.0099692	0.0119157
		4	0.0056382	0.0049296	0.0059140
		5	0.0027386	0.0023843	0.0028542
		6	0.0013495	0.0011826	0.0014153
		7	0.0007212	0.0006320	0.0007581
		8	0.0003432	0.0002912	0.0003475
B00018	FCNN	2	0.0205289	0.0159912	0.0189426

	3	0.0096451	0.0077552	0.0092191
	4	0.0050730	0.0040254	0.0047780
	5	0.0022966	0.0017585	0.0020886
	6	0.0011492	0.0008754	0.0010336
	7	0.0006432	0.0005005	0.0005950
	8	0.0002954	0.0002268	0.0002719
LSTM	2	0.0218554	0.0189299	0.0233109
	3	0.0109069	0.0094792	0.0116619
	4	0.0057440	0.0049472	0.0060704
	5	0.0026591	0.0022228	0.0027317
	6	0.0013255	0.0011411	0.0014012
	7	0.0007168	0.0006208	0.0007612
	8	0.0003431	0.0002941	0.0003649

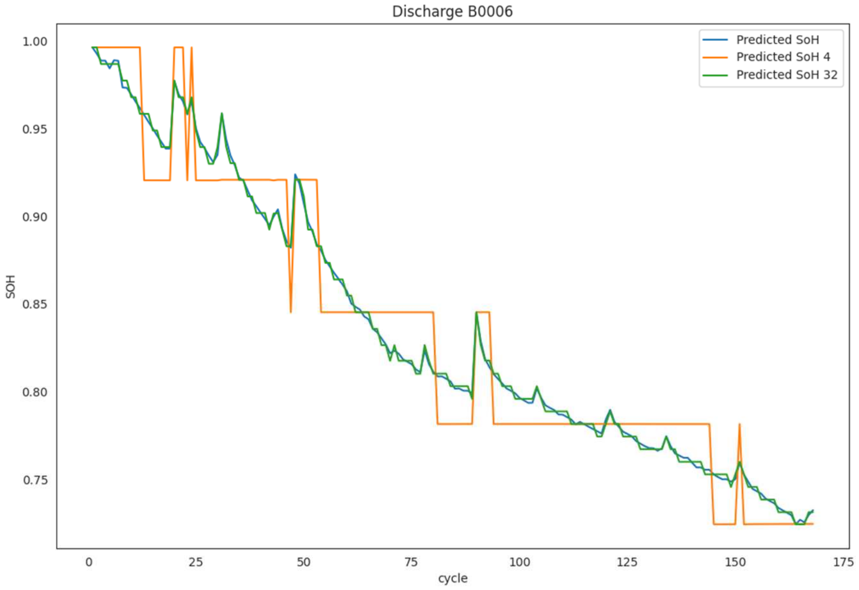
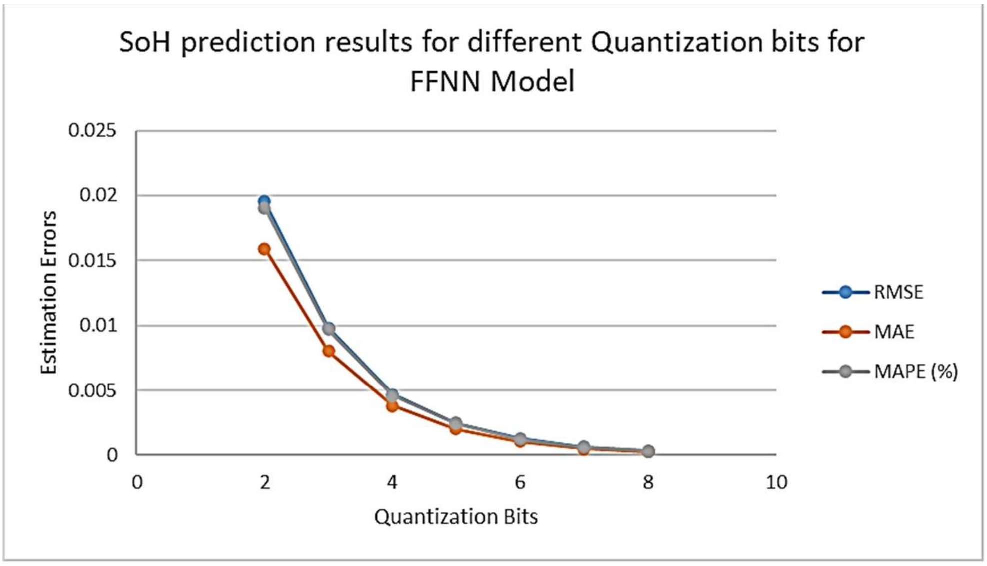


Figure 6. SOH predictions for B0006 Battery trained on B0005 Battery dataset showing original SOH and its 2-bit and 5-bit quantized versions.



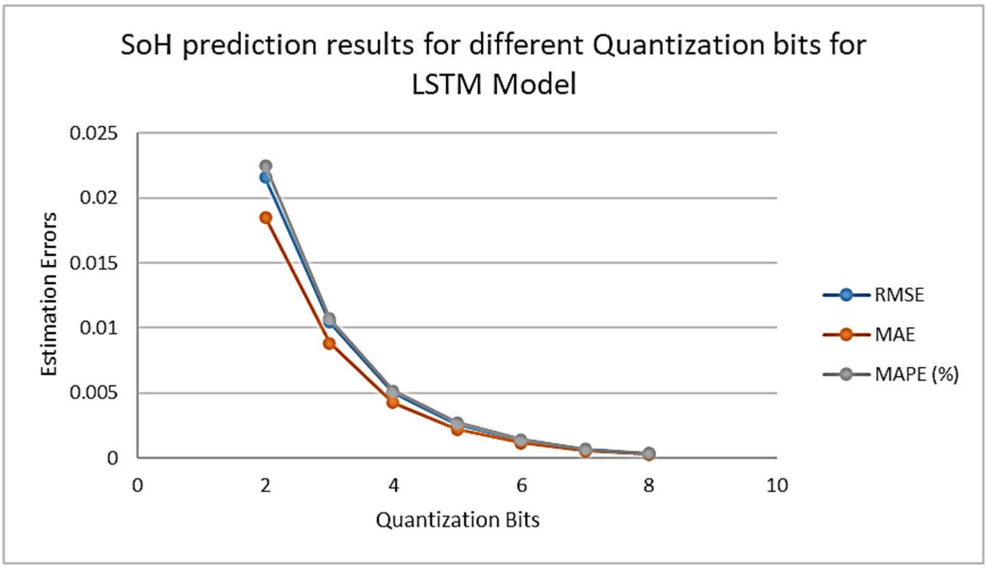


Figure 7. Effect of Quantization on SOH prediction error for B006 battery trained with FFN and LSTM models respectively.

8. Conclusions

This article proposes a replacement and an approximation for the neural network model for battery health estimation. By quantizing all the features used in battery SOH prediction problem, the traditional FCNN and LSTM network, a novel LUT memory network replacement model is constructed. Furthermore, the training data and testing data are collected from NASA to train and test the combined model, respectively. In addition, the BP algorithm for the neural networks, traditional FCNN, and LSTM are employed to estimate SOH to compare with the proposed method. The experiment result shows that the proposed method is superior when compared to its algorithms and model that it mimics. It was found that accuracies were slightly affected by this adoption, when compared to the actual inferred SOH estimation without quantization. The prediction accuracy error is improved up 65X when compared for b =2 bits to a b = 8 bits.3. An average b=3 bits was found to be a good quantization level which gave RMSE, MAE, and MAPE error accuracies of 0.0098006, 0.0080317, and 0.0096645, respectively. The novel finding was that b=5 bits will mostly quantize the seven SOH battery prediction issue features, totaling 35 bits. However, even with b=3 bits, the results were satisfactory, totaling 21 bits. This discovery made LUT operation reasonable with 2 MB LUT stored in a memory device giving virtually as close results as a full inference model.

References

1. Whittingham, M.S. Electrical Energy Storage and Intercalation Chemistry. Science 1976, 192, 1126–1127.
2. Stan AI, Świerczyński M, Stroe DI, Teodorescu R, Andreasen SJ. Lithium ion battery chemistries from renewable energy storage to automotive and back-up power applications—An overview. In2014 International Conference on Optimization of Electrical and Electronic Equipment (OPTIM) 2014 May 22 (pp. 713-720). IEEE.
3. Nishi, Y. Lithium-Ion Secondary Batteries; Past 10 Years and the Future. J. Power Sources 2001, 100, 101–106
4. Huang SC, Tseng KH, Liang JW, Chang CL, Pecht MG. An online SOC and SOH estimation model for lithium-ion batteries. Energies. 2017 Apr 10;10(4):512.
5. Goodenough JB, Kim Y. Challenges for rechargeable Li batteries. Chemistry of materials. 2010 Feb 9;22(3):587-603.
6. Nitta, N.;Wu, F.; Lee, J.T.; Yushin, G. Li-Ion Battery Materials: Present and Future. Mater. Today 2015, 18, 252–264.
7. Dai H, Jiang B, Hu X, Lin X, Wei X, Pecht M. Advanced battery management strategies for a sustainable energy future: Multilayer design concepts and research trends. Renewable and Sustainable Energy Reviews. 2021 Mar 1;138:110480.
8. Lawder MT, Suthar B, Northrop PW, De S, Hoff CM, Leitermann O, Crow ML, Santhanagopalan S, Subramanian VR. Battery energy storage system (BESS) and battery management system (BMS) for grid-scale applications. Proceedings of the IEEE. 2014 May 7;102(6):1014-30.
9. Lai X, Gao W, Zheng Y, Ouyang M, Li J, Han X, Zhou L. A comparative study of global optimization methods for parameter identification of different equivalent circuit models for Li-ion batteries. Electrochimica Acta. 2019 Feb 1;295:1057-66.
10. Wang Y, Gao G, Li X, Chen Z. A fractional-order model-based state estimation approach for lithium-ion battery and ultra-capacitor hybrid power source system considering load trajectory. Journal of Power Sources. 2020 Feb 15;449:227543.
11. Cheng G, Wang X, He Y. Remaining useful life and state of health prediction for lithium batteries based on empirical mode decomposition and a long and short memory neural network. Energy. 2021 Oct 1;232:121022.
12. Rechkemmer SK, Zang X, Zhang W, Sawodny O. Empirical Li-ion aging model derived from single particle model. Journal of Energy Storage. 2019 Feb 1;21:773-86.
13. Li K, Wang Y, Chen Z. A comparative study of battery state-of-health estimation based on empirical mode decomposition and neural network. Journal of Energy Storage. 2022 Oct 1;54:105333.
14. Geng Z, Wang S, Lacey MJ, Brandell D, Thiringer T. Bridging physics-based and equivalent circuit models for lithium-ion batteries. Electrochimica Acta. 2021 Mar 10;372:137829.
15. Xu, N.; Xie, Y.; Liu, Q.; Yue, F.; Zhao, D. A Data-Driven Approach to State of Health Estimation and Prediction for a Lithium-Ion Battery Pack of Electric Buses Based on Real-World Data. Sensors 2022, 22, 5762.
16. Alipour, M.; Tavallaey, S. Improved Battery Cycle Life Prediction Using a Hybrid Data-Driven Model Incorporating Linear Support Vector Regression and Gaussian. ChemPhysChem 2022, 23, e202100829.
17. Li, X.; Wang, Z. Prognostic health condition for lithium battery using the partial incremental capacity and Gaussian process regression. J. Power Sources 2019, 421, 56–67.
18. Li, Y.; Abdel-Monem, M. A quick on-line state of health estimation method for Li-ion battery with incremental capacity curves processed by Gaussian filter. J. Power Sources 2018, 373, 40–53.
19. Onori, S.; Spagnol, P.; Marano, V.; Guezennec, Y.; Rizzoni, G. A New Life Estimation Method for Lithium-Ion Batteries in Plug-in Hybrid Electric Vehicles Applications. Int. J. Power Electron. 2012, 4, 302–319.
20. Plett, G.L. Extended Kalman Filtering for Battery Management Systems of LiPB-Based HEV Battery Packs: Part 3. State and Parameter Estimation. J. Power Sources 2004, 134, 277–292.
21. Goebel, K.; Saha, B.; Saxena, A.; Celaya, J.R.; Christophersen, J.P. Prognostics in Battery Health Management. IEEE Instrum. Meas. Mag 2008, 11, 33–40.
22. Wang, D.; Yang, F.; Zhao, Y.; Tsui, K.L. Battery Remaining Useful Life Prediction at Different Discharge Rates. Microelectron. Reliab. 2017, 78, 212–219.
23. Li, J.; Landers, R.G.; Park, J. A Comprehensive Single-Particle-Degradation Model for Battery State-of-Health Prediction. J. Power Sources 2020, 456, 227950.
24. Hu, X.; Jiang, J.; Cao, D.; Egardt, B. Battery Health Prognosis for Electric Vehicles Using Sample Entropy and Sparse Bayesian Predictive Modeling. IEEE Trans. Ind. Electron. 2015, 63, 2645–2656.
25. Piao, C.; Li, Z.; Lu, S.; Jin, Z.; Cho, C. Analysis of Real-Time Estimation Method Based on Hidden Markov Models for Battery System States of Health. J. Power Electron. 2016, 16, 217–226
26. Liu, D.; Pang, J.; Zhou, J.; Peng, Y.; Pecht, M. Prognostics for State of Health Estimation of Lithium-Ion Batteries Based on Combination Gaussian Process Functional Regression. Microelectron. Reliab. 2013, 53, 832–839.

27. Khumprom, P.; Yodo, N. A Data-Driven Predictive Prognostic Model for Lithium-Ion Batteries Based on a Deep Learning Algorithm. Energies 2019, 12, 660.
28. Xia, Z.; Qahouq, J.A.A. Adaptive and Fast State of Health Estimation Method for Lithium-Ion Batteries Using Online Complex Impedance and Artificial Neural Network. In Proceedings of the 2019 IEEE Applied Power Electronics Conference and Exposition (APEC), Anaheim, CA, USA, 17–21 March 2019; pp. 3361–3365.
29. Eddahech, A.; Briat, O.; Bertrand, N.; Delétage, J.Y.; Vinassa, J.M. Behavior and State-of-Health Monitoring of Li-Ion Batteries Using Impedance Spectroscopy and Recurrent Neural Networks. Int. J. Electr. Power Energy Syst. 2012, 42, 487–494.
30. Shen, S.; Sadoughi, M.; Chen, X.; Hong, M.; Hu, C. A Deep Learning Method for Online Capacity Estimation of Lithium-Ion Batteries. J. Energy Storage 2019, 25, 100817.
31. Saha B, Goebel K. Battery data set, NASA ames prognostics data repository. NASA Ames, Moffett Field, CA, USA.
32. Ren, L.; Zhao, L.; Hong, S.; Zhao, S.; Wang, H.; Zhang, L. Remaining Useful Life Prediction for Lithium-Ion Battery: A Deep Learning Approach. IEEE Access 2018, 6, 50587–50598
33. Khumprom, P.; Yodo, N. A Data-Driven Predictive Prognostic Model for Lithium-Ion Batteries Based on a Deep Learning Algorithm. Energies 2019, 12, 660.
34. Choi, Y.; Ryu, S.; Park, K.; Kim, H. Machine Learning-Based Lithium-Ion Battery Capacity Estimation Exploiting Multi-Channel Charging Profiles. IEEE Access 2019, 7, 75143–75152.
35. Gong R, Liu X, Jiang S, Li T, Hu P, Lin J, Yu F, Yan J. Differentiable soft quantization: Bridging full-precision and low-bit neural networks. InProceedings of the IEEE/CVF international conference on computer vision 2019 (pp. 4852-4861).
36. Choi J, Wang Z, Venkataramani S, Chuang PI, Srinivasan V, Gopalakrishnan K. Pact: Parameterized clipping activation for quantized neural networks. arXiv preprint arXiv:1805.06085. 2018 May 16.
37. Esser SK, McKinstry JL, Bablani D, Appuswamy R, Modha DS. Learned step size quantization. arXiv preprint arXiv:1902.08153. 2019 Feb 21.
38. Yang Z, Wang Y, Han K, Xu C, Xu C, Tao D, Xu C. Searching for low-bit weights in quantized neural networks. Advances in neural information processing systems. 2020;33:4091-102.
39. Courbariaux M, Bengio Y, David JP. Binaryconnect: Training deep neural networks with binary weights during propagations. Advances in neural information processing systems. 2015;28.
40. Zhu C, Han S, Mao H, Dally WJ. Trained ternary quantization. arXiv preprint arXiv:1612.01064. 2016 Dec 4.
41. Rastegari M, Ordonez V, Redmon J, Farhadi A. Xnor-net: Imagenet classification using binary convolutional neural networks. InEuropean conference on computer vision 2016 Sep 17 (pp. 525-542). Cham: Springer International Publishing.
42. Ullrich K, Meeds E, Welling M. Soft weight-sharing for neural network compression. arXiv preprint arXiv:1702.04008. 2017 Feb 13.
43. Xu Y, Wang Y, Zhou A, Lin W, Xiong H. Deep neural network compression with single and multiple level quantization. InProceedings of the AAAI conference on artificial intelligence 2018 Apr 29 (Vol. 32, No. 1).
44. Zhou A, Yao A, Guo Y, Xu L, Chen Y. Incremental network quantization: Towards lossless cnns with low-precision weights. arXiv preprint arXiv:1702.03044. 2017 Feb 10.
45. Miyashita D, Lee EH, Murmann B. Convolutional neural networks using logarithmic data representation. arXiv preprint arXiv:1603.01025. 2016 Mar 3.
46. Blalock D, Gonzalez Ortiz JJ, Frankle J, Guttag J. What is the state of neural network pruning?. Proceedings of machine learning and systems. 2020 Mar 15;2:129-46.
47. Gou J, Yu B, Maybank SJ, Tao D. Knowledge distillation: A survey. International Journal of Computer Vision. 2021 Jun;129:1789-819.
48. Google. 2019. TensorFlow: An end-to-end open source machine learning platform. https://www.tensorflow.org/
49. MACE. 2020. https://github.com/XiaoMi/mace
50. Microsoft. 2019. ONNX Runtime. https://github.com/microsoft/
51. Wang M, Ding S, Cao T, Liu Y, Xu F. Asymo: scalable and efficient deep-learning inference on asymmetric mobile cpus. InProceedings of the 27th Annual International Conference on Mobile Computing and Networking 2021 Sep 9 (pp. 215-228).
52. Wang M, Ding S, Cao T, Liu Y, Xu F. Asymo: scalable and efficient deep-learning inference on asymmetric mobile cpus. InProceedings of the 27th Annual International Conference on Mobile Computing and Networking 2021 Sep 9 (pp. 215-228).
53. Liang R, Cao T, Wen J, Wang M, Wang Y, Zou J, Liu Y. Romou: Rapidly generate high-performance tensor kernels for mobile gpus. InProceedings of the 28th Annual International Conference on Mobile Computing And Networking 2022 Oct 14 (pp. 487-500).

54. Jiao Y, Han L, Long X. Hanguang 800 npu–the ultimate ai inference solution for data centers. In2020 IEEE Hot Chips 32 Symposium (HCS) 2020 Aug 1 (pp. 1-29). IEEE Computer Society.
55. Jouppi NP, Yoon DH, Ashcraft M, Gottscho M, Jablin TB, Kurian G, Laudon J, Li S, Ma P, Ma X, Norrie T. Ten lessons from three generations shaped google's tpuv4i: Industrial product. In2021 ACM/IEEE 48th Annual International Symposium on Computer Architecture (ISCA) 2021 Jun 14 (pp. 1-14). IEEE.
56. Wechsler O, Behar M, Daga B. Spring hill (nnp-i 1000) intel's data center inference chip. In2019 IEEE Hot Chips 31 Symposium (HCS) 2019 Aug 1 (pp. 1-12). IEEE Computer Society.
57. Xiong R, Li L, Tian J. Towards a smarter battery management system: A critical review on battery state of health monitoring methods. Journal of Power Sources. 2018 Nov 30;405:18-29.
58. Hannan MA, Lipu MH, Hussain A, Mohamed A. A review of lithium-ion battery state of charge estimation and management system in electric vehicle applications: Challenges and recommendations. Renewable and Sustainable Energy Reviews. 2017 Oct 1;78:834-54.
59. Waag W, Fleischer C, Sauer DU. Critical review of the methods for monitoring of lithium-ion batteries in electric and hybrid vehicles. Journal of Power Sources. 2014 Jul 15;258:321-39.
60. Gold, L.; Bach, T.; Virsik, W.; Schmitt, A.; Müller, J.; Staab, T.E.M.; Sextl, G. Probing Lithium-Ion Batteries' State-of-Charge Using Ultrasonic Transmission—Concept and Laboratory Testing. J. Power Sources 2017, 343, 536–544
61. Robinson JB, Owen RE, Kok MD, Maier M, Majasan J, Braglia M, Stocker R, Amietszajew T, Roberts AJ, Bhagat R, Billsson D. Identifying defects in Li-ion cells using ultrasound acoustic measurements. Journal of The Electrochemical Society. 2020 Sep 1;167(12):120530.
62. R-Smith NA, Leitner M, Alic I, Toth D, Kasper M, Romio M, Surace Y, Jahn M, Kienberger F, Ebner A, Gramse G. Assessment of lithium ion battery ageing by combined impedance spectroscopy, functional microscopy and finite element modelling. Journal of Power Sources. 2021 Nov 15;512:230459.
63. Liu X, Zhang L, Yu H, Wang J, Li J, Yang K, Zhao Y, Wang H, Wu B, Brandon NP, Yang S. Bridging multiscale characterization technologies and digital modeling to evaluate lithium battery full lifecycle. Advanced energy materials. 2022 Sep;12(33):2200889.
64. Han X, Ouyang M, Lu L, Li J, Zheng Y, Li Z. A comparative study of commercial lithium ion battery cycle life in electrical vehicle: Aging mechanism identification. Journal of power sources. 2014 Apr 1;251:38-54.
65. Wu H, Judd P, Zhang X, Isaev M, Micikevicius P. Integer quantization for deep learning inference: Principles and empirical evaluation. arXiv preprint arXiv:2004.09602. 2020 Apr 20.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.