

# CS 229 Project Milestone

## Project Category: Physical Sciences

### Modelling State-of-Health for a Li-ion Battery

Karthik Nataraj (kartnat), Hampus Carlens (hcarlens) and Julian Cooper (jelc)

October 28, 2022

## 1 Motivation

Lithium ion batteries are of great and increasing importance in today's society due to their high energy density. Li-ion batteries' performance capability can be characterized by their so called state of health (SOH). SOH is a measure of usable capacity over rated capacity. Accurately predicting how many cycles (i.e charge - discharge) a battery can perform, at any given time, before reaching it's end of useful life is difficult, however important for reliability etc. Prediction is hard in part due to poor understanding of how the measurable parameters (voltage, current etc) effect the SOH. Today three main methods are used: electrochemical, equivalent circuit, and data-driven models.

We aim to use a data-driven model to predict remaining useful life (RUL) and future SOH degradation. This method requires little knowledge about the underlying factors. There have been previous studies successfully achieving data-driven SOH prediction with high accuracy. For example some recognized studies are (1), (2) and (3). These articles show predictions of SOH and RUL with a range of different mean average percentage errors, all less than 10%. The lowest percentage errors are obtained using deep neural networks such as in (2) and (3), with MAPE's  $< 2\%$  being reported. However, the authors of (3) and (4) point towards the importance of future work in trying to scale the models, to make them suitable for front-end embedded systems. In (1) they opt for a simple linear model over neural networks. They achieved a significantly lower accuracy and resolution of the predictions, mainly being able to predict RUL using the first hundred cycles. We specifically aim to query the interactions learnt by the neural network model (SHAP, LIME, or ideally following methods in (5) etc.) to infer complex features we might use to build a simpler, explainable white-box model with reasonable accuracy. The model should be able to predict SOH using fewer cycles than (1) and with higher accuracy.

## 2 Data exploration

The dataset we have selected contains approximately 96,700 cycles (approx. 780 cycles per battery for 124 batteries). For each cycle we capture voltage, current applied and temperature sampled at 2.5 second intervals. This is largest publicly available dataset for nominally identical commercial lithium-ion batteries cycled under controlled conditions, and is the same source used by two recent papers that motivated our project: "Machine learning pipeline for battery state-of-health estimation" (Nature, 2021)(2) and "Data-driven prediction of battery cycle life before capacity degradation" (Nature, 2019)(1).

- **Discharge capacity vs cycle number.** Our goal is to predict the discharge curves below given information from the first 50-100 cycles. A few things to note. First, the batteries in our population are either rated at 1.1 Ah or 1.05 Ah nominal capacity. By nominal here we mean manufacturer rated initial capacity. Second, in practice, our initial discharge capacity rarely perfectly matches the nominal capacity rating and so instead of two discrete starting points at 1.1 and 1.05, our starting points range continuously in that range. Third, our data is meant to reflect cycling each battery from its initial capacity to 80% of its nominal. This explains the two distinct end points we observe in the right hand chart at 0.88 Ah and 0.84 Ah, the 80% thresholds of 1.1 Ah and 1.05 Ah nominal capacities respectively. Last, in plotting these curves we identified some obvious outliers. Batteries in the left hand plot with capacities either above 1.2 or below 0.8 are erroneous measurements that we remove from the data before training.

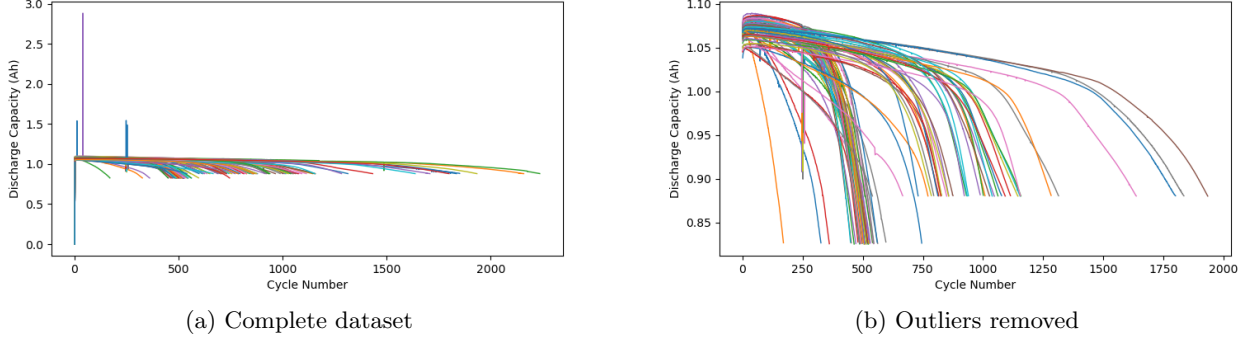


Figure 1: Discharge capacity by cycle number for 124 batteries

- Distribution of our target variable.** Cycle life is the number of cycle it takes for a given battery to reach 80% of its nominal capacity. This is our target variable. When plotting the complete dataset we identified a skewed normal distribution. Ideally we want to roughly maintain this distribution for our validation and test data. To achieve this we re-used logic from "Data-driven prediction of battery cycle life before capacity degradation" (1) to split our train, validation and test data.

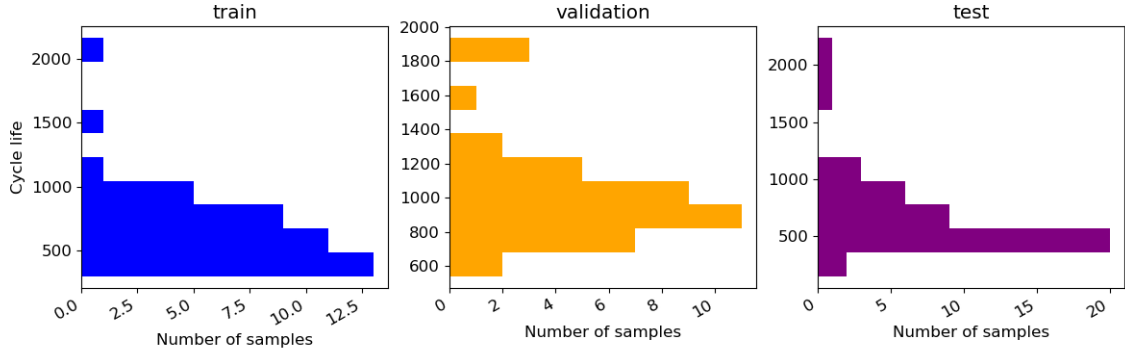


Figure 2: Distribution of target for train, validation and test data

- Applied current and charge cycle.** The charts below illustrate current, charge capacity, and discharge capacity for all cycles of an example battery data point (ref: b1c0). We can immediately confirm that the vast majority of cycles follow similar charge profiles. In particular, applying positive current (charging) for first 500-700 time steps, then switching to an applied negative current (discharging) until depleted at the end of the cycle. The charge capacity ( $Q_c$ ) and discharge capacity ( $Q_d$ ) charge plots also show this transition, with  $Q_c$  increasing up until the changeover to negative applied current, and  $Q_d$  increasing only after the changeover.

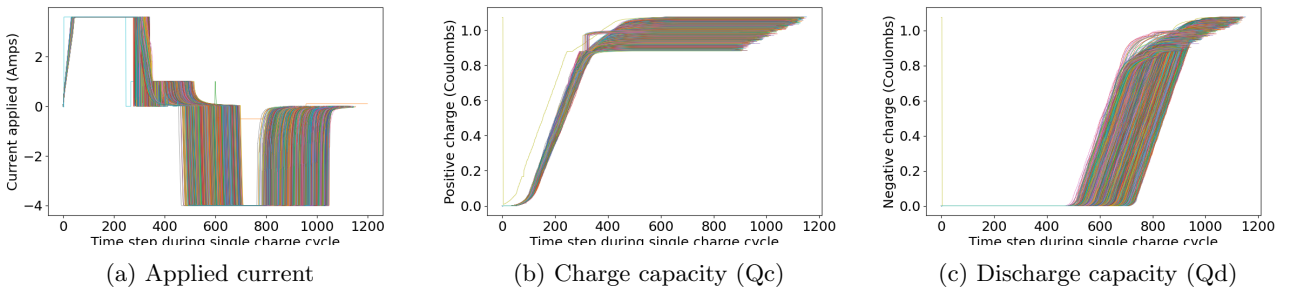


Figure 3: Charge and discharge cycles for an example battery (ref: b1c0)

We also investigated how voltage and temperature vary over the cycles of the same example battery. It is interesting to note that while temperature has a roughly gaussian distribution throughout any given cycle, the voltage measure has almost no variance during charge but significant variance across cycles during discharge.

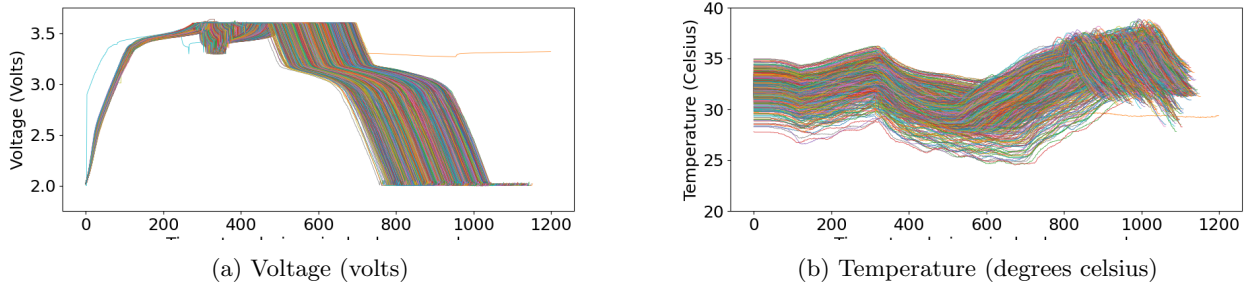


Figure 4: Voltage and temperature over cycles for an example batter (ref: b1c0)

- **Correlation with cycles @ 5% fade.** Finally, we wanted to directly plot some measure of capacity fade (in this case number of cycles to reach 5% decrease from nominal) during the initial cycles against cycle life to see if we can recover the correlated behaviour we expect between early trajectory and end point of the discharge capacity curve. Encouragingly, we see the two are highly correlated, achieving a pearson correlation coefficient of approximately 0.94.

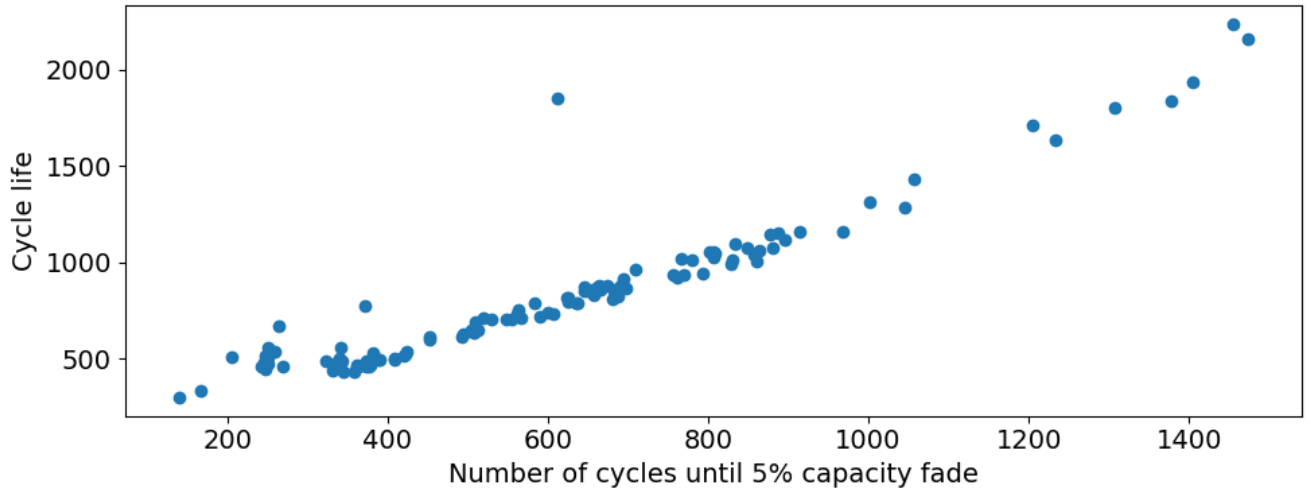


Figure 5: Scatter plot demonstrating correlation between cycle life and 5% capacity fade

## 3 Methodology

### 3.1 Model fit evaluation

Our plan is to use 80% of the data for training, leaving 10% for validation and 10% for out-of-sample test. For out-of-sample test we will initialize with the first 50-100 cycles, and then ask the model to predict State-of-Health (SOH) for the remaining cycle until end-of-life threshold).

To evaluate out-of-sample performance, we will compute two error metrics:

1. **State-of-Health Curve (SOH):** Compute Mean Average Percentage Error for predicted SOH values, one value per cycle for each battery used for validation.
2. **Remaining-Useful-Life (RUL):** Compute Mean Average Percentage Error for predicted RUL values, one value per battery used for validation.

While these error metrics are related ( $RUL = \text{first cycle for which SOH dips below end-of-life threshold}$ ), they do measure different things. For example, we might correctly predict the number of cycles before end-of-life (let's say 10,000), but guess that the path is linear instead of parabolic or logistic. This might mean that at 50 cycles our prediction of available capacity (SOH) is much worse than the true value despite good RUL accuracy.

### 3.2 Prediction and inference

Our goal is to build two models, one for prediction and the other for inference:

1. **Prediction:** Black-box model (e.g., LSTM Neural Network) that is intended to be a useful tool for battery manufacturers to test (predict RUL and SOH profile given 100 cycles) their batteries meet desired capacity specifications before leaving the factory.
2. **Inference:** By querying the results of our black box model (e.g., partial dependence plots) we hope to learn the marginal effects of our feature variables. Given this and our understanding of the battery physics, we would like to construct a alternative white box model (e.g., Auto-regression) from which we can perform inference

Given what we learn from our initial black box model, we hope to develop a white box model which comparable performance metrics. Such a model is often more desirable for engineering applications since its relationships and behavior are transparent and simple to sense check against known physics.

## 4 Experiments

### 4.1 Black-box model (completed first experiments)

Interestingly in (1) only a regularized elastic net was used to predict remaining useful life of a set of batteries, using information gleaned from the first 100 cycles. Our first step was to use a similar feature set but using a black-box neural network instead, to see if we obtain a lower MAPE. We tried a preliminary experiment using 81 batteries for training and 43 for testing. Our training input consisted of the internal resistance and average temperature measurements over the 100 initial cycles. Additionally we used the following very predictive feature: log variance of the difference between the discharge capacity curves as a function of voltage, between the 10th and 100th initial cycles. In this initial model since we were only looking to replicate the initial results of this paper, we used a shallow neural network with a thousand node hidden layer, ReLU activation, and standard mean squared error loss. We achieved a test MAPE of about 12%, which was right in line with what the paper achieved (about 15% on a primary test set, 11 % on a secondary).

Studying the test set predictions there are a couple of test set batteries that exhibit anomalous behavior (non-standard capacity - voltage curves), though we have yet to confirm whether they should be taken out of the test set. Also (1) had additional models using many more features based on intra-cycle information, which we have yet to leverage. Lastly, we may be able to increase the complexity of our network (number of layers and/or nodes/layer) as well as run for a greater number of epochs to further refine our test MSE with this approach. Spending a little time to see if we can really improve results for the RUL problem can give us more insight whether it would be sensible to try and predict the whole SoH decay curve over the remaining cycle life.

### 4.2 White-box model (next steps)

We want to build a highly performant autoregressive model, with feature engineering based on assessment of our black box model. Coming soon!

## References

- [1] K. A. Severson, P. M. Attia, N. Jin, N. Perkins, B. Jiang, Z. Yang, M. H. Chen, M. Aykol, P. K. Herring, D. Fraggdakis *et al.*, “Data-driven prediction of battery cycle life before capacity degradation,” *Nature Energy*, vol. 4, no. 5, pp. 383–391, 2019.
- [2] D. Roman, S. Saxena, V. Robu, M. Pecht, and D. Flynn, “Machine learning pipeline for battery state-of-health estimation,” *Nature Machine Intelligence*, vol. 3, no. 5, pp. 447–456, 2021.
- [3] S. Jo, S. Jung, and T. Roh, “Battery state-of-health estimation using machine learning and preprocessing with relative state-of-charge,” *Energies*, vol. 14, no. 21, 2021. [Online]. Available: <https://www.mdpi.com/1996-1073/14/21/7206>
- [4] D. Roman, R. Dickie, D. Flynn, and V. Robu, “A review of the role of prognostics in predicting the remaining useful life of assets,” in *Safety and Reliability. Theory and Applications*, M. Čepin and R. Briš, Eds. CRC Press, 2017, pp. 897–904, 27th European Safety and Reliability Conference 2017, ESREL 2017 ; Conference date: 18-06-2017 Through 22-06-2017. [Online]. Available: <http://esrel2017.org/>
- [5] M. Tsang, D. Cheng, and Y. Liu, “Detecting statistical interactions from neural network weights,” *ArXiv*, vol. abs/1705.04977, 2018.