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 26, 2022

Format
Your milestone should be at most 3 pages, excluding references. Similar to the proposal, it should include - Motivation: What problem are you tackling, and what's the setting you're considering? Method: What machine learning techniques have you tried and why?

- Preliminary experiments: Describe the experiments that you've run, the outcomes, and any error analysis that you've done. You sould have tried at least one baseline.

- Next steps: Given your preliminary results, what are the next steps that you're considering?

1   Motivation

Lithium ion batteries are of great and increasing importance in today's society due to their high energy density. We are very dependent on their performance. Li-on 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.

The predicting is hard in part due to poor understanding of how the measurable parameters (voltage, current etc) effect the SOH. Today three main methods; electrochemical models, equivalent circuit models and data-driven models, are used.

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 over 90% accuracy. The highest accuracies are obtained using deep neural networks such as in (2) and (3). There, accuracies > 97% have been reported. However, the authors of (3) and (4) point toward the importance of future work in trying to scale the models, to make them suitable for front-end embedded systems. In (1) they attempt to not use deep learning and use a linear model instead. They achieved a significantly lower accuracy and resolution of the predictions, mainly being able to predict RUL using a lot of cycles. We specifically aim to query the interactions learned by the neural network model (SHAP, LIME, 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   Method

We can classify our approaches into two categories, black-box and white-box methods: WORK IN PROGRESS ...

2.1   Black-box model

These include sequential deep learning techniques such as RNN's and LSTM networks, which we can use to predict the decay curves of capacity versus cycle number based on a feature set including various voltage, current, and temperature information.

2.2 White-box model

Various autoregressive models, feature engineering based on assessment of black box model with target being the SoH or "capacity" (number of hours for which a battery can provide a current equal to the discharge rate at battery's rated voltage).

Since we ultimately want to use early cycle information to predict the exact decay curve, it might be the case that the sequential neural networks will not have enough information–in this case we may revert to regular dense, fully connected layers, which have the added benefit of easily incorporating other features.

3 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 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.
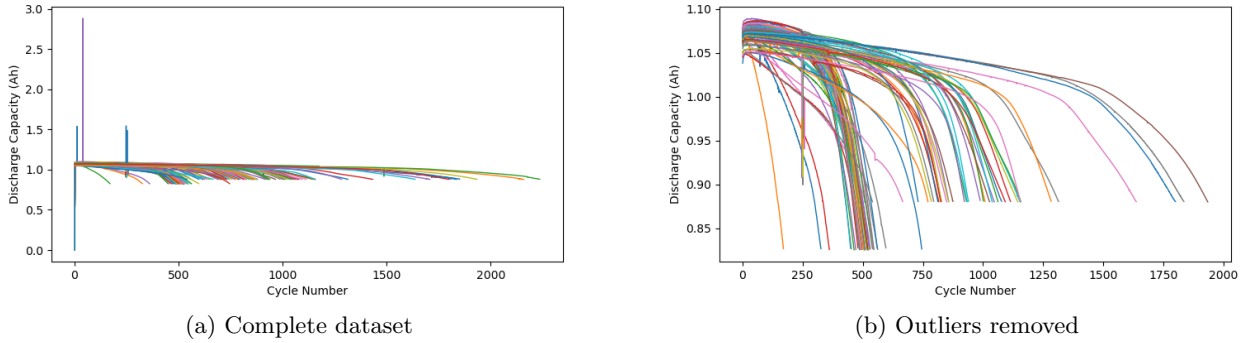

(a) Complete dataset   (b) Outliers removed

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" (Nature, 2019)(1) to split our train, validation and test data.
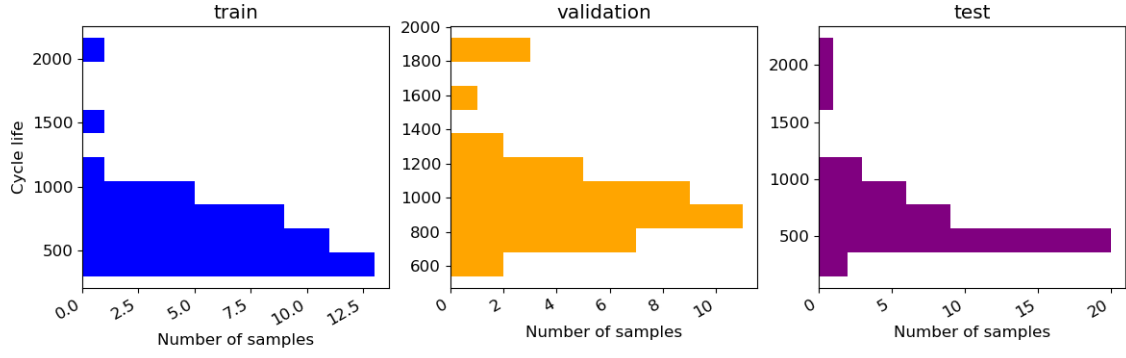
Figure 2: Distribution of target for train, validation and test data

- **Applied current and charge cycle**. The charts below illustrate current and positive / negative charge for all cycles of an example battery data point (ref: c1b0). 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 positive (Qc) and negative (Qd) charge plots also show this transition, with Qc increasing up until the changeover to negative applied current, and Qd increasing only after the changeover.

(a) Applied current

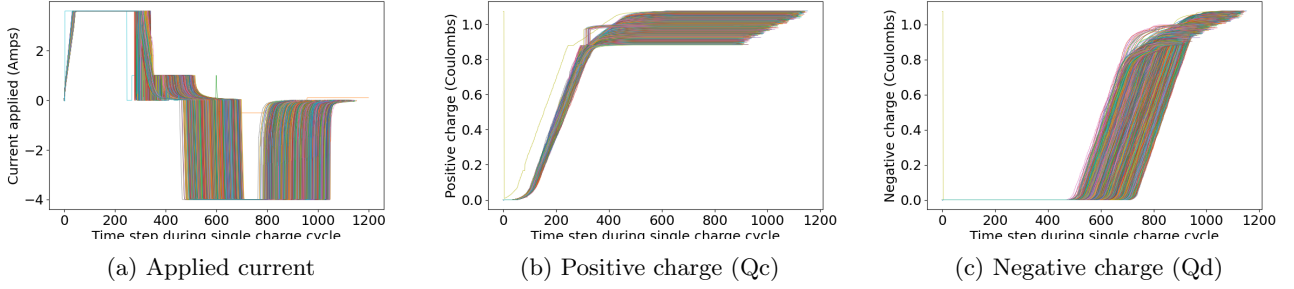(b) Positive charge (Qc)

(c) Negative charge (Qd)

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.
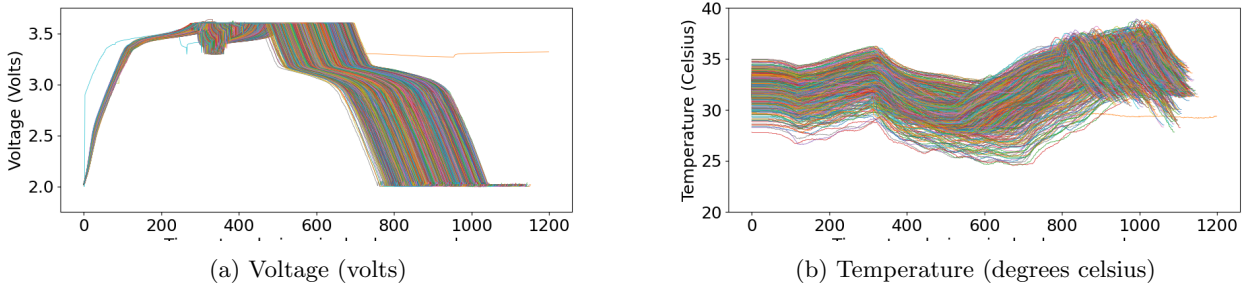

(a) Voltage (volts)

(b) Temperature (degrees celsius)

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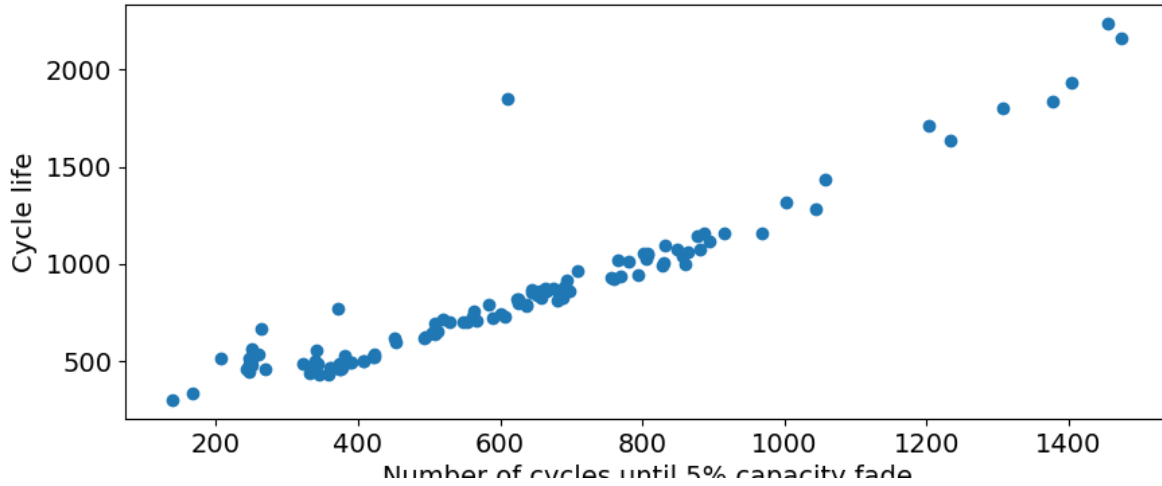
Figure 5: Scatter plot demonstrating correlation between cycle life and 5% capacity fade

4 Experiments

4.1 Model fit evaluation

Our plan is to use 70% of the data for training, leaving 15% for cross-validation and 15% for out-of-sample validation. For out-of-sample validation we will initialize with the first 10 cycles, and then ask the model to predict State-of-Health (SOH) for the next 90 cycles (or until we hit end-of-life threshold).

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

1. **State-of-Health (SOH)**: Compute Mean Squared Error for predicted SOH values, one value per cycle for each battery used for validation.

2. **Remaining-Useful-Life (RUL)**: Compute Mean Squared Error for predicted RUL values, one value per battery used for validation.

While these error metrics are related (RUL = 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 100), 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.

4.2 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 10 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.

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. Fraggedakis *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/