Modeling State of Health for a Li-On Battery

Karthik Nataraj (kartnat@stanford.edu) Hampus Carlens (hcarlens@stanford.edu) Julian Cooper (jelc@stanford.edu)

Motivation & our dataset

Lithium ion batteries are of great and increasing importance in today's society due to their high energy density. Li-on batteries' performance capability can be characterized by their State of Health (SOH). SOH is a measure of usable capacity over rated capacity. Accurately predicting how many cycles a battery can perform before reaching 80% of rated capacity (Remaining Useful Life (RUL)) is important for reliability. While much work has already been done in this space to accurately predict RUL, recent studies ([1], [3]) have called for further research into white box models that predict the SOH curve to enable deployment for online systems. In this project, we develop a scalable (and explainable) **white box model to predict the SOH curve until 80% threshold** (as opposed to a point estimate of RUL) given the first 100 cycles. Since our predicted SOH curve implies an RUL prediction, we also compare our implied RUL prediction error to that of previous studies.

The **dataset we have selected contains approximately 96,700 cycles (approx. 780 cycles per battery for 124 batteries)**. For each cycle the authors captured voltage, applied current and temperature sampled at 2.5 second intervals. This is the largest publicly available dataset for 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)[4].

1. Neural Network

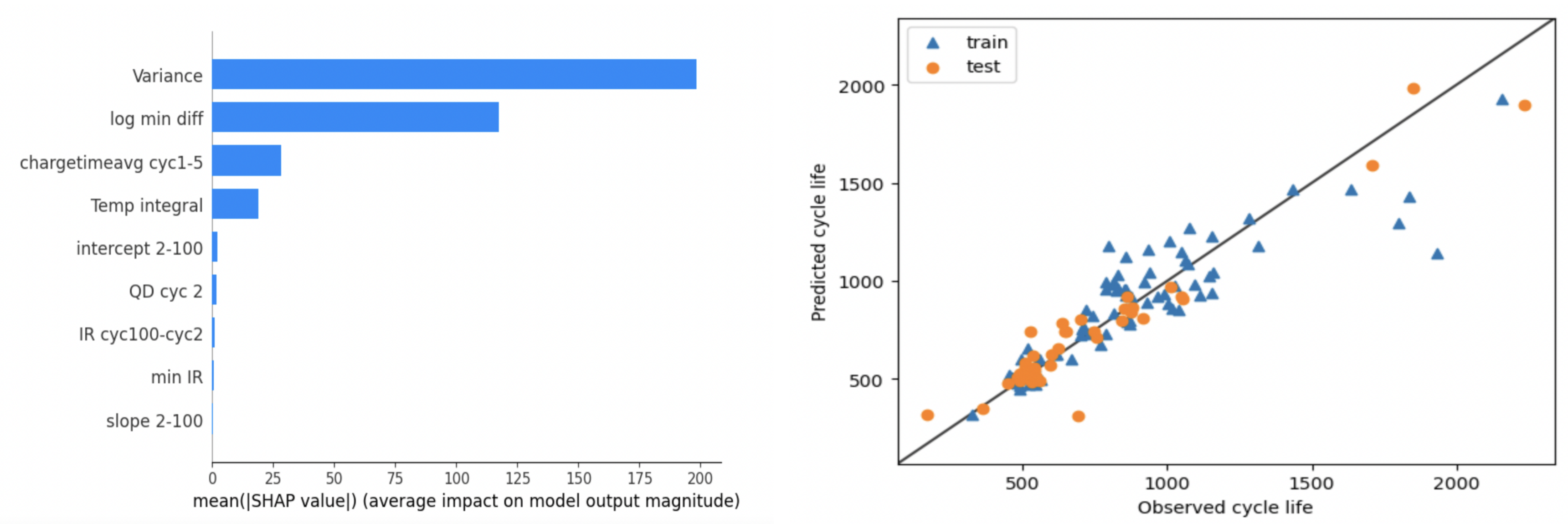
Our Neural Network model predicts RUL of a battery given data features derived only from its first 100 cycles. The purpose was to improve results from [4] and perform feature importance analyses to inform subsequent time series and Bayesian Inference models.

(a) **Architecture**: 2-layer feed-forward, densely connected neural network. First layer had 1076 neurons and the second 96. We then used the learning rate of .001 with the Adam optimizer, ReLU activation, and full-batch gradient descent (since the training dataset is already small). Training and validation: 81 batteries; Test: 43

(b) **Feature engineering**: Due to slow training convergence and poor performance despite l1/l2 regularization, condensed original derived feature set of [4] consisting of 20 features, to a total of 3, using Shapley values ((a) of Figure 2): variance of difference in charge curves, average temperature and current

Results. Obtained training RUL MAPE $\approx 8.6\%$, test RUL MAPE $\approx 10.8\%$, better than the best model of [4], which obtained test MAPE 13%.

MAPE RUL 10.8%



(a) Feature importance for non time-series variables (b) Predicted vs. actual cycle life for test set batteries

Figure 1. Descriptive plots for neural network approach to predicting

Discussion and takeaways

- Even with an small training dataset it was possible to improve performance of a regularized linear model with shallow neural networks, using a very small sample of very good features.
- Eliminating noisy features proved more helpful in improving performance than hyperparameter tuning. Our dataset was too small to support significant hyperparameter calibration, even with cross-validation techniques.
- **Next steps**: Going forward we'd focus on decay curve predictions using LSTM networks, given the partial success of ARIMA models (see next column)

2. Auto-regression (ARIMA)

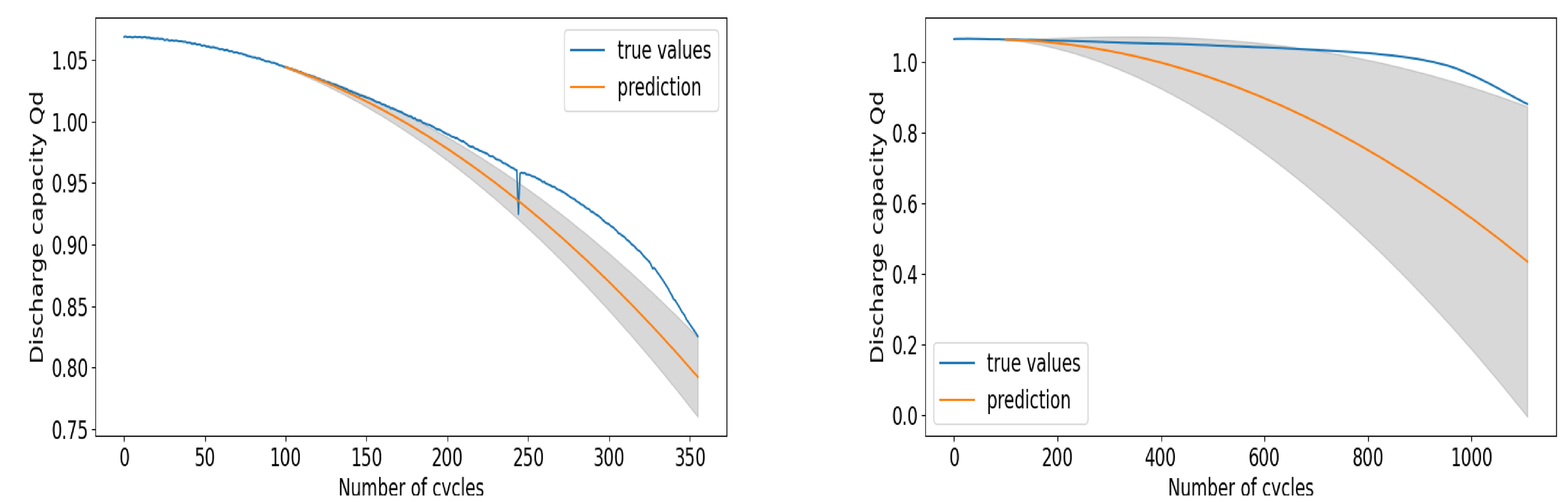
The purpose of using time series models was to predict the entire capacity degradation curve, not only the RUL as with the neural net. The ARIMA model considered only one measurement per cycle, the discharge capacity Q_D and from a machine learning point of view there was a separate model for each battery.

(a) **Choice of hyperparameters** was done through Box-Jenkins method which is based on autocorrelation and partial autocorrelation. The results were then sense checked using an automatic parameter selection algorithm called auto-ARIMA. We found that p=2, d=2, q=2 worked well, and in particular that our results were very sensitive to d being at least 2.

(b) **Exogenous variables**; ARIMA models can be used with or without exogenous variables. We tried adding exogenous variables to the ARIMA model, however, since the battery life cycles were of different lengths it was hard to consistently apply exogenous variables in training and testing. We ended up not using exogenous variables.

Results Fit evaluation using ARIMA(2,2,2) on out-of-sample test batteries given information from the first 100 cycles only.

MSE SOH 0.049
MAPE RUL 26.5%



(a) Good prediction example on test battery (b) Poor prediction example on test battery

Figure 2. Predicted discharge capacity over cycle life until 0.8 threshold

Discussion and takeaways

- The ARIMA model results had a very large variance, as seen in Figure (2).
- It was almost impossible to estimate any 1000+ cycle battery life using just 100 cycles. However, battery cycle lives of 300-700 were possible to estimate with reasonable accuracy.
- The largest limitation of the model was the inability to learn across batteries.
- **Next steps**: Future work could focus on implementing exogenous variable to impose data priors on the algorithm. This could include an RUL estimation from the neural network combined with a generic decay curve.

3. Bayesian inference model

We explored Bayesian Inference as a way to impose more known physics on the problem. Main idea: since we know the discharge capacity curve of each battery must be a decay curve, why don't we specify such a functional form and only ask our model to learn the shape and translation.

(a) **Functional form**: We selected an inverse sigmoid $\hat{y} = \gamma - 1/1 + \exp(-\alpha(x - \beta))$ because it described our data well in the region of interest (y = 0.8 to 1.2) and its parameters (shape, translation, y-asymptote) were highly interpretable.

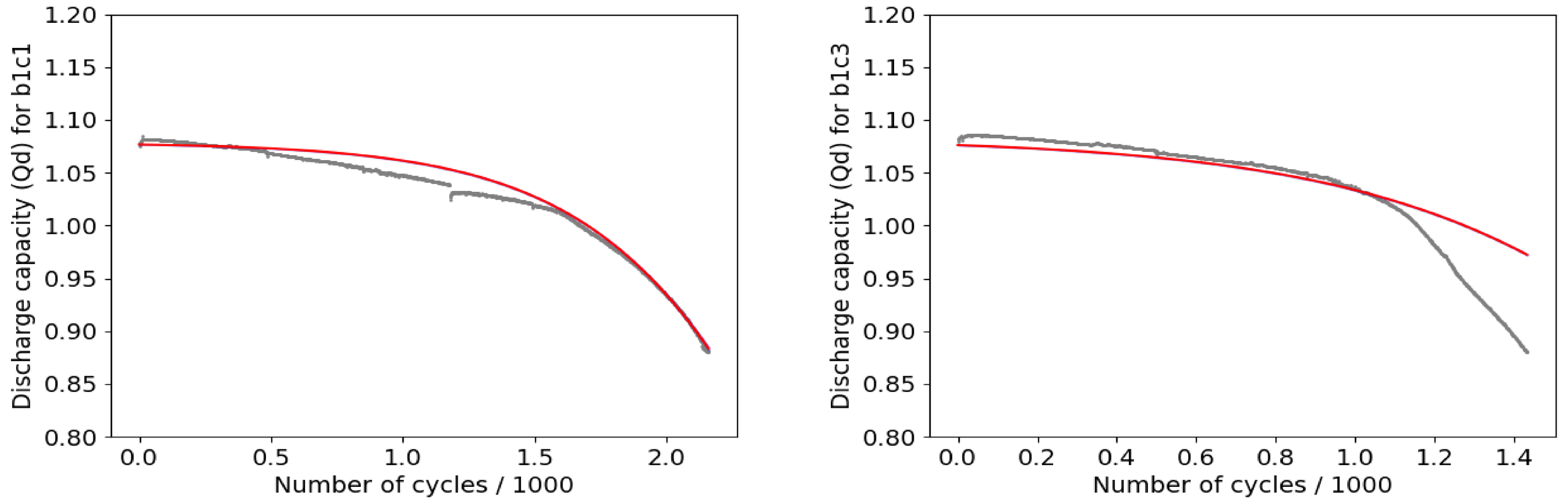
(b) **Parameterization**: Relating parameters α, β, γ to features x_i identified by Neural Network:

$\alpha = a_0 + a_1x_1 + a_2x_2 + a_3x_3 + a_4x_4 + a_5x_5$	linear model for rate of decay
$\beta = b_0 + b_1x_1 + b_2x_2 + b_3x_3 + b_4x_5 + b_5x_5$	linear model for horizontal translation
$\gamma = x_1$	asymptote set to nominal discharge capacity

(c) **Prior specification**: We assume that our labels y are generated from a gaussian with mean $\hat{y}$ (our predictions) and variance $\sigma^2 \sim$ Gamma(2, 1). Informative intercept priors $a_0 \sim N(3, 1)$, $b_0 \sim N(2, 1)$ (empirically estimated) and standard normals for $a_i, b_i \sim N(0, 1)$ $i = 1, 2, ..., d$.

Results Fit evaluation on out-of-sample test batteries given information from the first 100 cycles only.

MSE SOH 0.015
MAPE RUL 33.5%



(a) Good prediction example on test battery (b) Poor prediction example on test battery

Figure 3. Predicted discharge capacity over normalized cycle life

Discussion and takeaways

- Our model predicts well for batteries with mid-long cycle lives (Figure 3a), but often overshoots for short cycle life batteries (e.g. Figure 3b).
- This is partly due us having so few short cycle life training observations (variance), but also may be indicative of inflexibility in the model parameterization (bias).
- **Next steps**: We use a linear model to parameterize α and β, but a deeper understanding of the physics involved might lead us to specify something different (e.g. interaction terms).

[1] Sungwoo Jo, Sunkyu Jung, and Taemoon Roh. Battery state-of-health estimation using machine learning and preprocessing with relative state-of-charge. *Energies*, 14(21), 2021. ISSN 1996-1073. doi:10.3390/en14217206. URL `https://www.mdpi.com/1996-1073/14/21/7206`.

[2] Darius Roman, Saurabh Saxena, Valentin Robu, Michael Pecht, and David Flynn. Machine learning pipeline for battery state-of-health estimation. *Nature Machine Intelligence*, 3(5):447–456, 2021.

[3] Darius V. Roman, Ross W. Dickie, David Flynn, and Valentin Robu. A review of the role of prognostics in predicting the remaining useful life of assets. In Marko Čepin and Radim Briš, editors, *Safety and Reliability. Theory and Applications*, pages 897–904. CRC Press, 2017. ISBN 9781138629370. URL `http://esrel2017.org/`. 27th European Safety and Reliability Conference 2017, ESREL 2017 ; Conference date: 18-06-2017 Through 22-06-2017.

[4] Kristen A Severson, Peter M Attia, Norman Jin, Nicholas Perkins, Benben Jiang, Zi Yang, Michael H Chen, Muratahan Aykol, Patrick K Herring, Dimitrios Fraggedakis, et al. Data-driven prediction of battery cycle life before capacity degradation. *Nature Energy*, 4(5):383–391, 2019.