

# CS 229 Project Proposal

## Project Category: Physical Sciences

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

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

October 7th, 2022

## 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.

## 2 Method

### 2.1 Machine learning techniques

[Karthik to insert text here]

### 2.2 About the data source

The data is the same as used in the article "Machine learning pipeline for battery state-of-health estimation" (Nature April 5th 2021). It is comprised of thousands of charge-discharge cycles for over 120 li-on batteries. Every cycle contains measurements of voltage, current etc sampled at 2.5 second intervals. Link to data can be found [here](#).

## 3 Experiments

### 3.1 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.

### 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 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.