Li-ion Battery Remaining Lifetime Prediction

Ruhui Shen, Yang Li
Contact: shenruhui@outlook.com

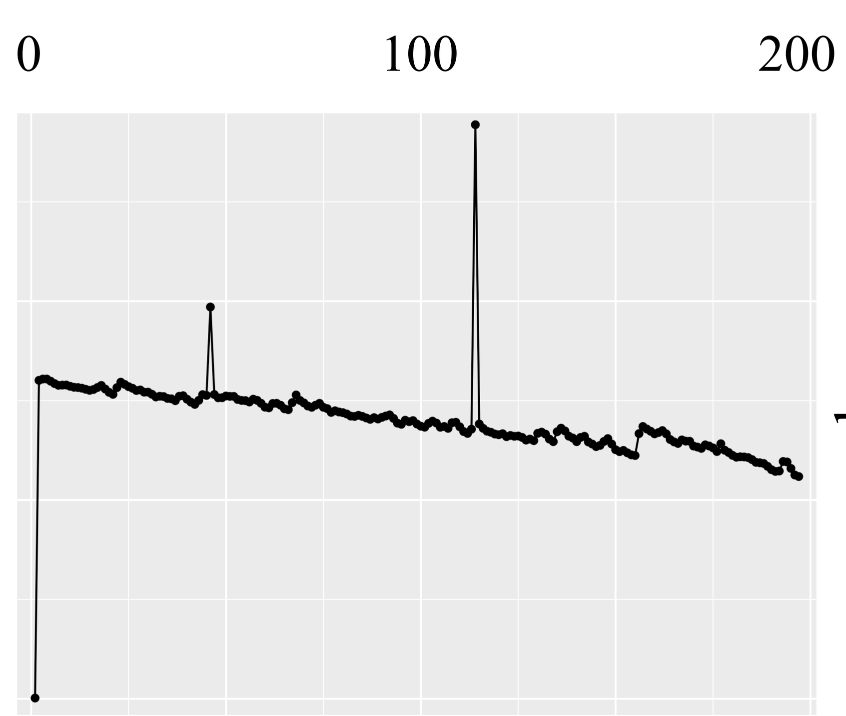
Abstract

Quantitative cycle life prediction using earlier cycles of battery is an interesting and important task in lithium-ion battery systems. Different selections of features including voltage, location, capacity and current extracted from semi-supervised NASA dataset with different regression methods are used to acquire an accurate early-stage fade prediction. Significant improvement to the baseline method was acquired in the 1A and 4A group.

Problem Definition

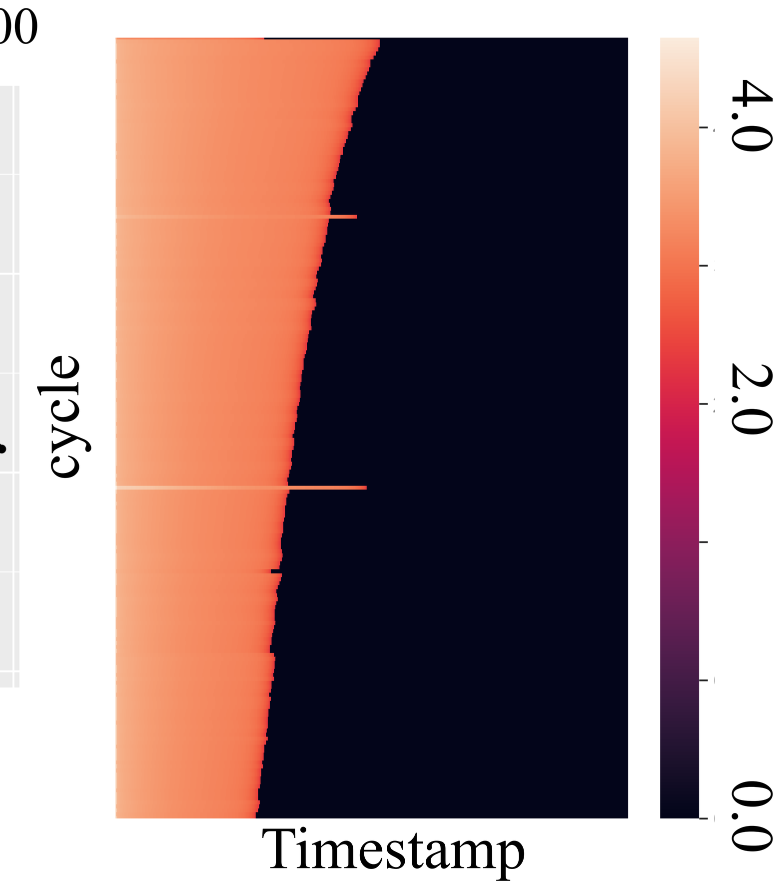
<u>Input</u>:
- **Capacity** during the first t cycles $Q_0, \dots, Q_t$ of a battery.
- Other per-cycle measurements: **voltage**, current, etc.

<u>Goal</u>: Predict fade at $t + kth$ cycle: y: $y = 1 - \frac{Q_{t+k}}{Q_0}$.



Capacity vs. Cycle

Voltage sequences in all cycles



Timestamp

<u>Research Question</u>:
- **Which features should we use to predict fade?**
- **How to predict fade accurately using few cycles and limited training data.**

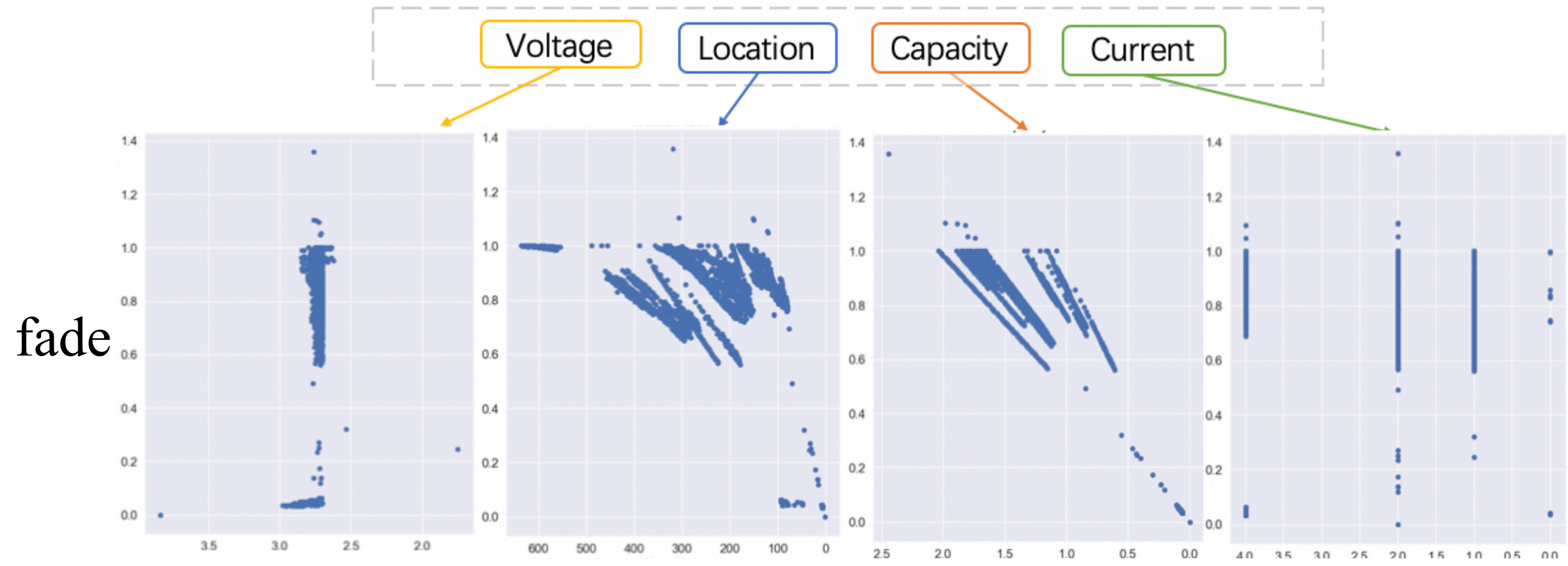
Dataset

- Extracted from the NASA Battery Data Set.
- 24 Batteries charged/discharged at different temperatures and different currents (1A, 2A, 4A)

Method

(1) Feature selection:

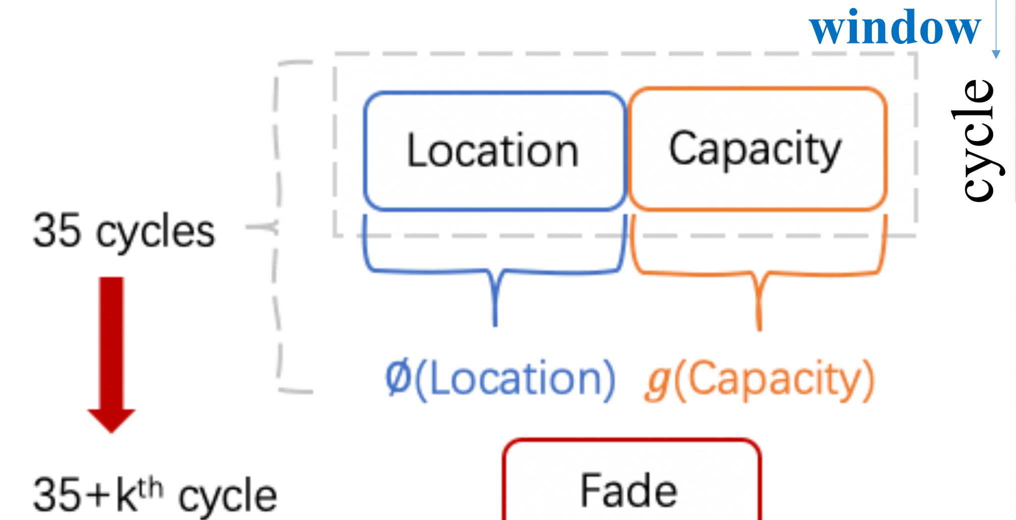
Voltage : The final value of voltage sequence (a sequence of voltage measured in each discharge) before the cut-off voltage.
Location : The length of voltage sequence before the cut-off voltage.
Capacity : The capacity in each discharge cycle.
Current : The experiment current.



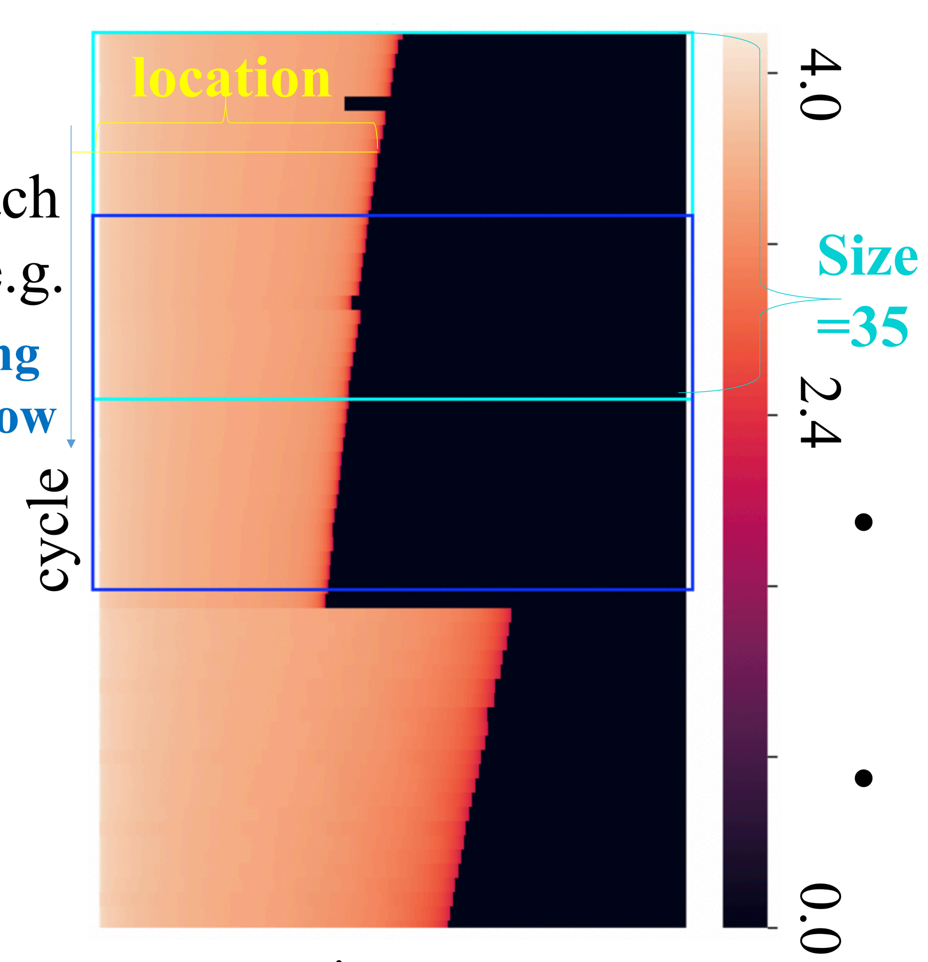
- Voltage is not informative for fade prediction.
- Location and Capacity are negatively correlated with fade.
- Current mainly has 3 unique values: 1A(11 samples), 2A(10 samples) and 4A(6 samples).

(2) Fade prediction:

<u>Data pre-processing</u>:
- Group samples by currents
- Data augmentation: sample t consecutive cycles in each battery using a sliding window (e.g. t=35, step size=1)



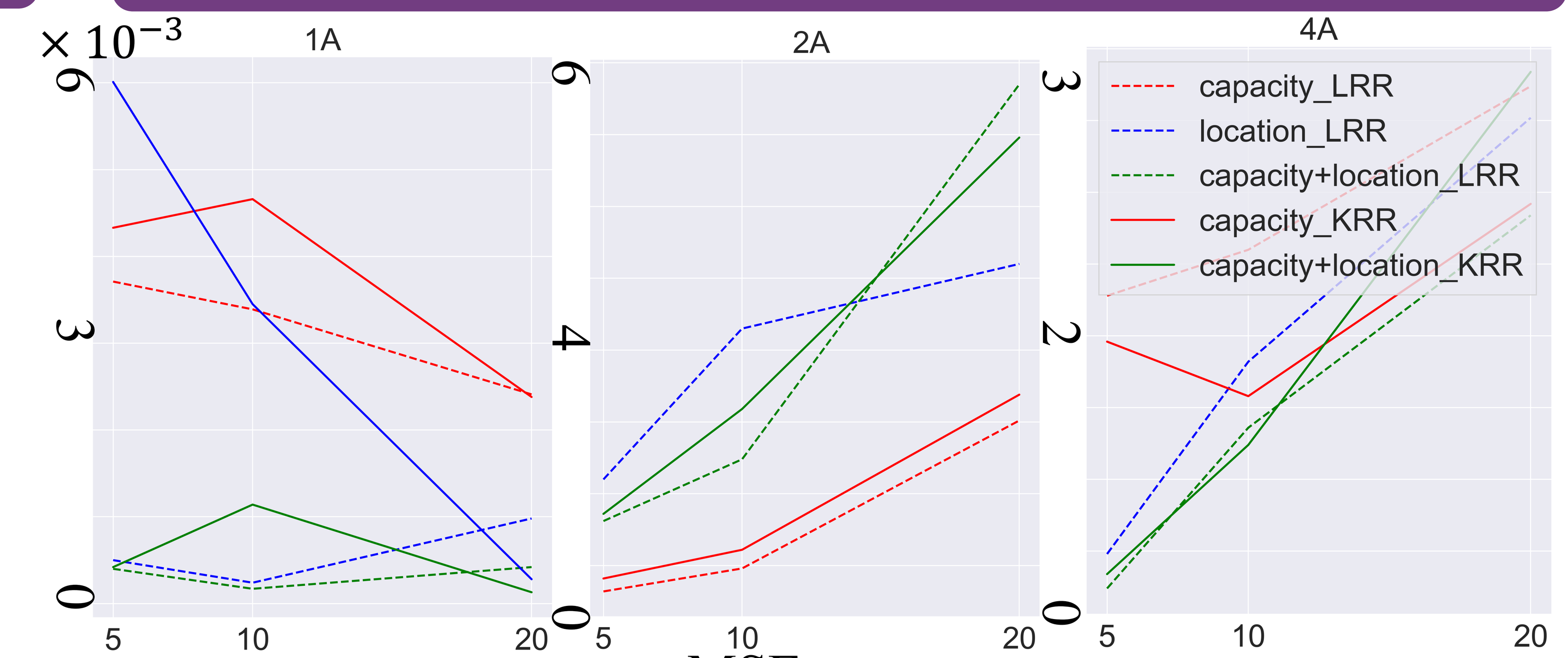
Voltage sequences in all cycles



Timestamp

<u>Regression methods</u>: linear regression (LR), linear ridge regression (LRR), kernel ridge regression (KRR). CNN-based methods perform poorly due to the small sample size.

Results

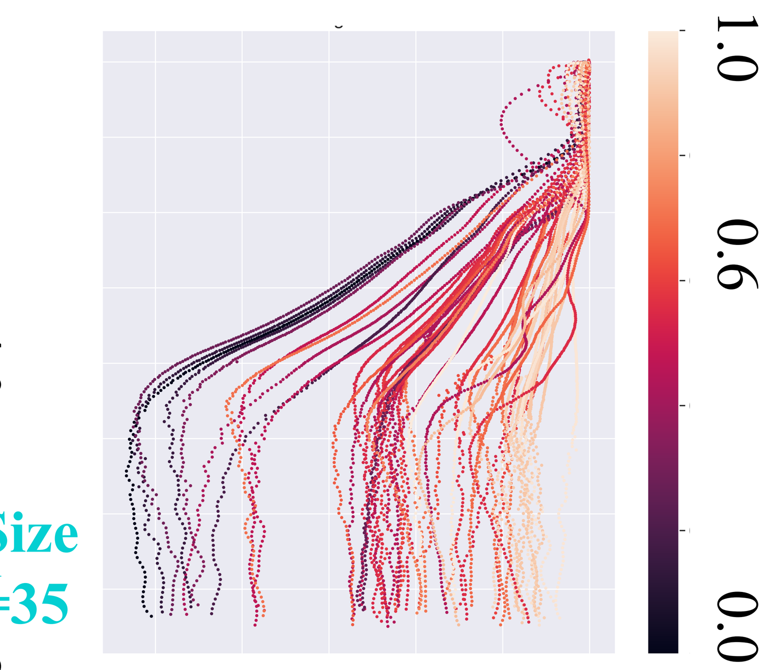


MSE vs. step

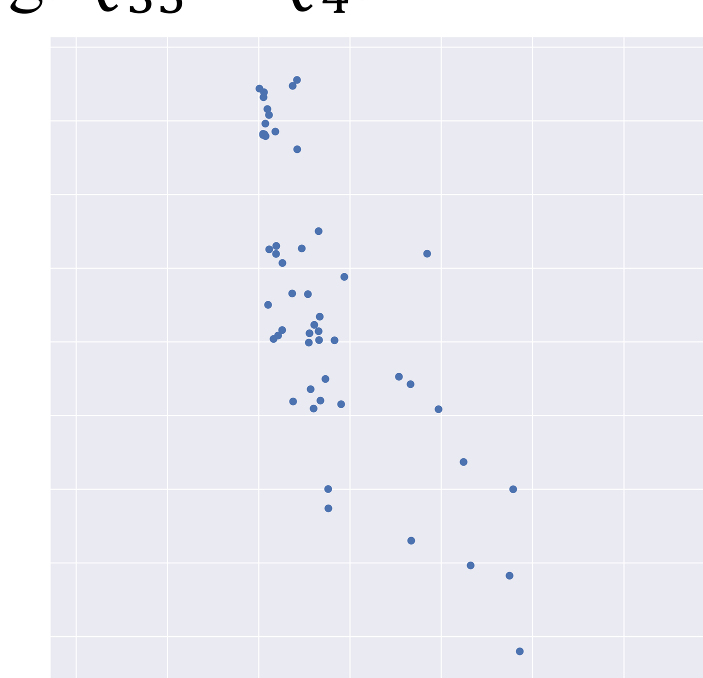
- t=35 and k(step)=5,10,20.
- # of samples in each group after data augmentation: 179 (1A), 1535 (2A), 644 (4A).
- Capacity+ location (green line) feature performs better than capacity(redline baseline) in most case.
- LRR (dotted line) performs better than KRR (solid line) in most case.

Future Works

- Utilize other features such as difference of capacity sequences(Current measured × Time in each discharge cycle.) e.g. $Q_{35} - Q_4$.



Voltage vs. $Q_{35} - Q_4$



Fade vs. Var($Q_{35} - Q_4$)

- Consider other regression models: Gaussian process, RNN- based sequence-to-sequence prediction.
- Utilize multiple heterogenous datasets to improve prediction accuracy.

The author would like to acknowledge the help and guidance provided by Prof. Yang Li, Prof. Xuan Zhang and Zihao Zhou.
Reference: [1]. Severson, K.A., Attia, P.M., Jin, N. et al. Data-driven prediction of battery cycle life before capacity degradation. Nat Energy 4, 383–391 (2019). https://doi.org/10.1038/s41560-019-0356-8. [2]. Peter Attia, Marc Deetjen, Jeremy Witmer. Accelerating battery development via early prediction of cell lifetime (2019).