Supplementary Information for

Prediction of the Knee Point in Battery Aging Curve Based on Machine Learning

Xijun Tan[1,2], Nathan Zeng[3], Yixiang Deng[4,5,6], Shuguo Sun[1,2], Bo Rui[1,2], Jun Xu[1,21]

* Corresponding authors: junxu@udel.edu

Affiliations:

[1]Department of Mechanical Engineering, University of Delaware, Newark, DE 19711, USA

[2]Energy Mechanics and Sustainability Laboratory (EMS), University of Delaware, Newark, DE 19711, USA

[3]Dover-Sherborn High School, Dover, MA 02030

[4]Department of Computer and Information Sciences, University of Delaware, Newark, DE 19711, USA

[5]Department of Biomedical Engineering University of Delaware, Newark, DE 19711, USA

[6]Data Science Institute, University of Delaware, Newark, DE 19711, USA

[1]Corresponding to: Prof. Jun Xu (junxu@udel.edu)

Weight coefficient:

For the weight coefficient in multiple-feature models, the initial start was equal for each features input (50% for two-features model and 20% for five-features model). Then, $\pm$1% of each weight coefficient was adjusted in each model and sum of each weight coefficient kept 1. Since the train data and test data was randomly selected, we ran 50 times for each weight coefficient condition and compared the main MAPE (main absolute percentage error). The best weight coefficient for each model is shown in Table 1.

Train test Results:

In some cases, the training test showed a higher MAPE than the primary test, likely due to differences in battery chemistry. The training data included two different battery chemistries, whereas the primary test set contained only one. This combination of chemistries in the training data may introduce conflicts that increase the MAPE in the training test relative to the primary test. When the EMSLab dataset, representing a different chemistry system, was removed from the training process, the training test MAPE no longer exceeded that of the primary test.

Supplementary Figures

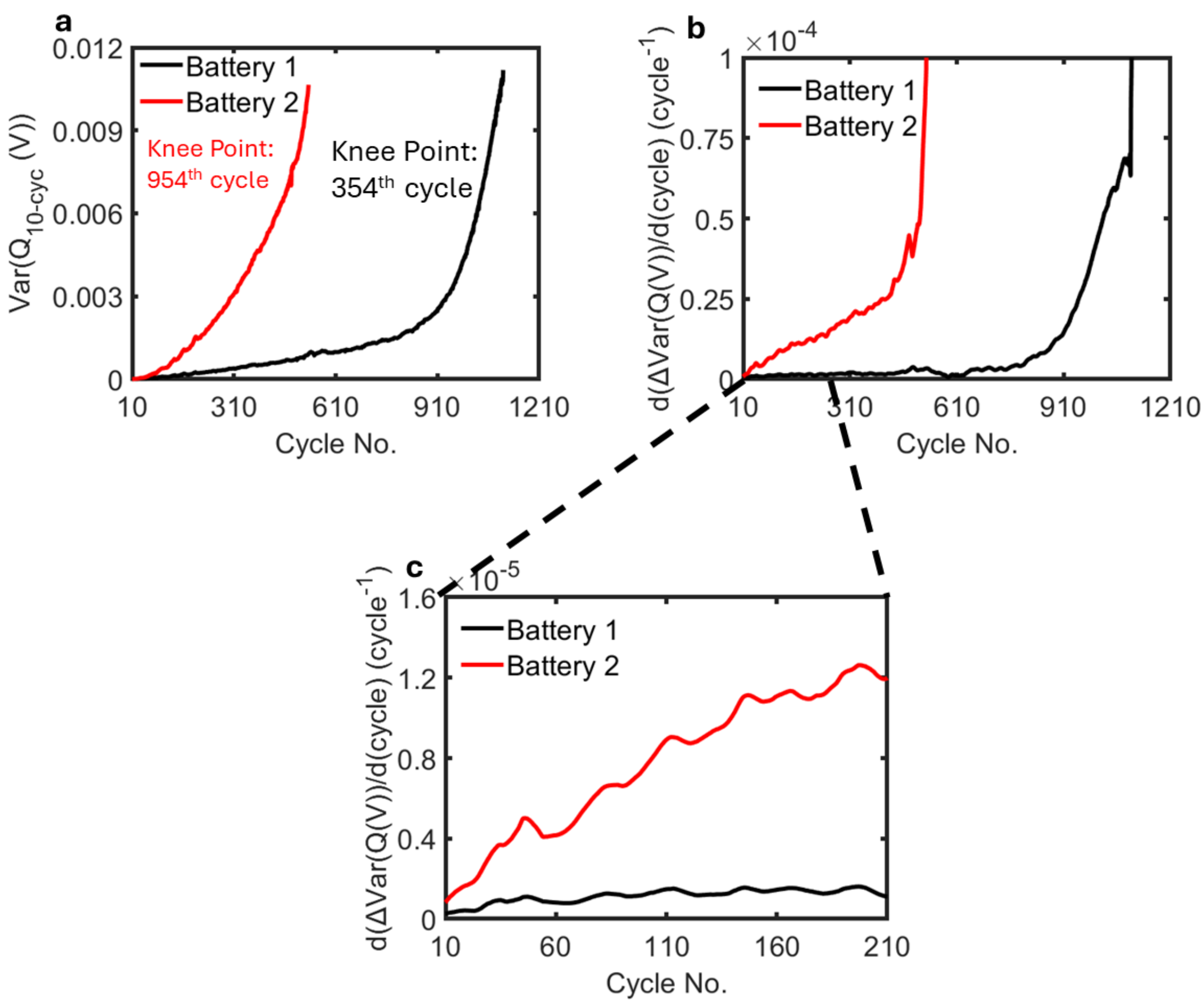

Fig S1: Feature extraction for $\mathrm{d}^2(\mathrm{var}(\Delta \mathrm{Q}_{\mathrm{cycle}-10}(\mathrm{V})))$

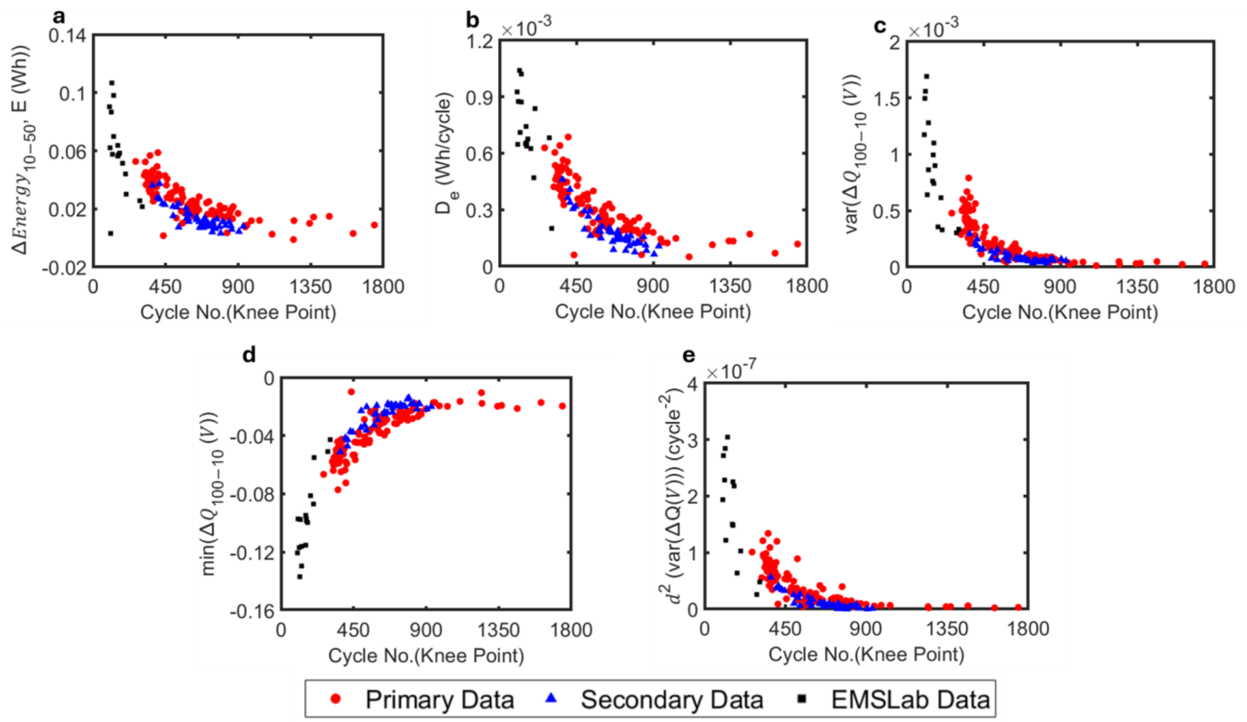

Fig. S2: Relationship between knee point cycle number and features ((a) $\Delta\mathrm{Energy}_{10-100}$, (b) $\mathrm{D_E}$, (c) $\mathrm{var}\left(\Delta\mathrm{Q}_{100-10}(\mathrm{V})\right)$, (d) min $(\Delta\mathrm{Q}_{100-10}(\mathrm{V}))$, and (e) $\mathrm{d}^2(\mathrm{var}(\Delta\mathrm{Q}_{\mathrm{cycle}-10}(\mathrm{V}))))$ in first 100 cycles

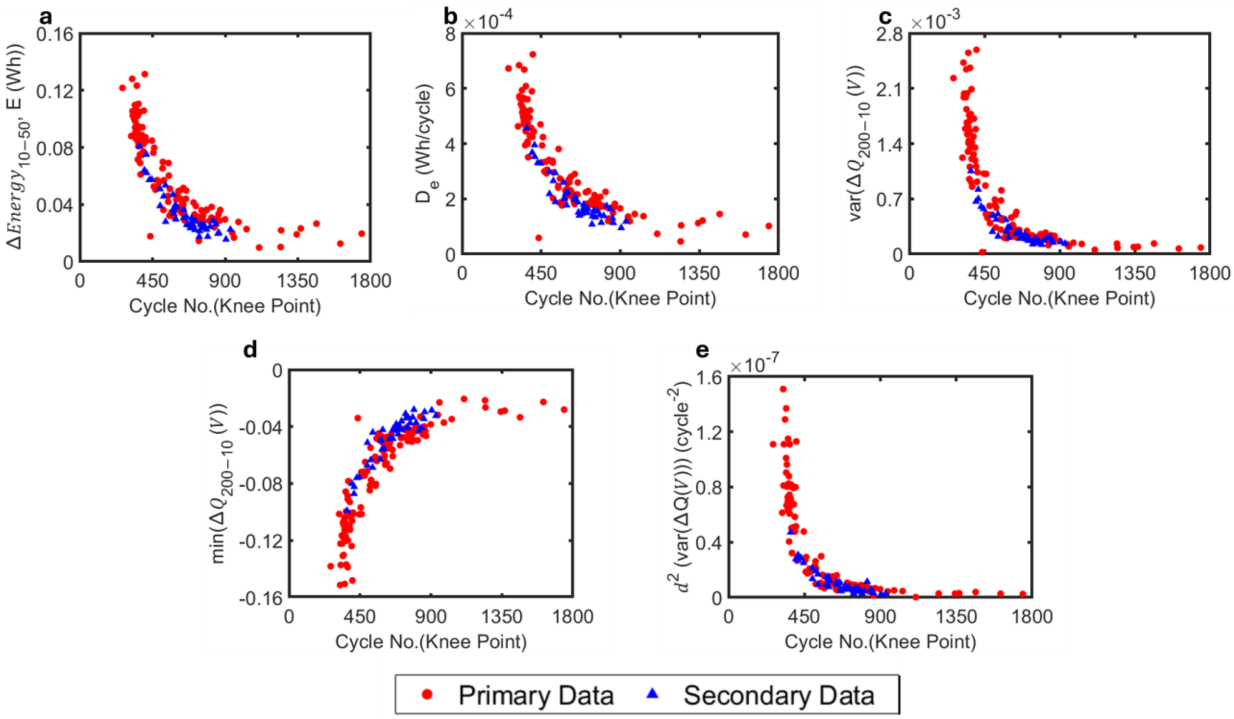

Fig. S3: Relationship between knee point cycle number and features ((a) $\Delta \mathrm{Energy}_{10-100}$, (b) $\mathrm{D_E}$, (c) $\mathrm{var}\left(\Delta \mathrm{Q}_{100-10}(\mathrm{V})\right)$, (d) min $(\Delta \mathrm{Q}_{100-10}(\mathrm{V}))$, and (e) $\mathrm{d}^2(\mathrm{var}(\Delta \mathrm{Q}_{\mathrm{cycle}-10}(\mathrm{V})))$) in first 200 cycles

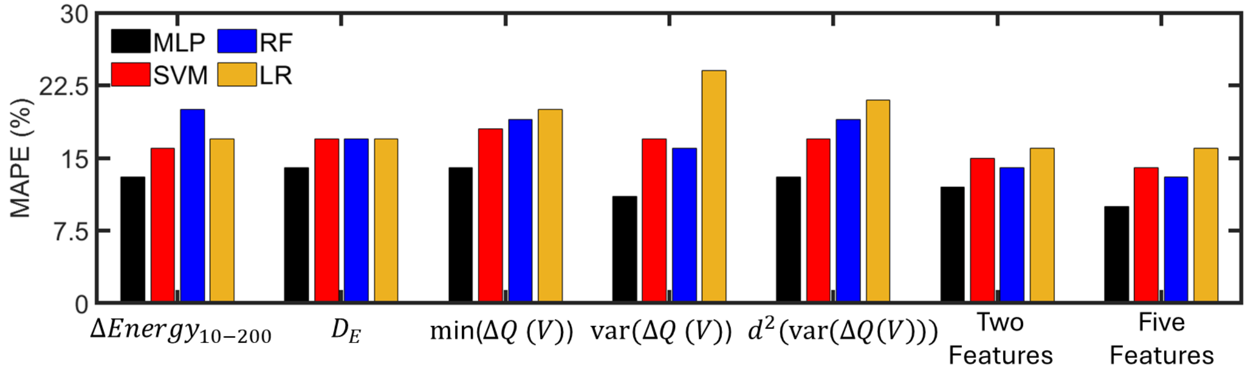

Fig S4: Predicted results by using different algorithms including multilayer perceptron (MLP), support vector machine (SVM), random forest (RF), and linear regression (LR).

Supplementary Tables

Table S1. Network parameters for MLP pretraining

Network parameters	Value
Input layer (neurons)	1(30)
Hidden layers (neurons)	4(30)
Output layer (neurons)	1(1)
Activation function	relu
Epochs	300
Batch Size	10
Optimizer	Adam
Learning rate	0.001
Shuffle frequency	1 time each epoch