Machine Learning Methods for Early Prediction of End-of-Life for Lithium-ion Batteries

James Chapman
CIS 732 Machine Learning – Term Project
Kansas State University
jachapman@ksu.edu

I. Introduction

The global market for lithium-ion batteries is expected to increase by a factor of 5 to 10 in the next decade [1]. The demand is driven by a combination of environmental concerns and technological advances that make lithium-ion batteries the most reliable source of electrical power for many applications around the world including consumer electronics, electric vehicles (EVs), and renewable energy storage. Lithium-ion batteries have high power and high energy density while providing great reusability and recyclability. It is crucial that we continually refine battery technology in ways such as reducing costs, enhancing efficiency and safety, and improving reliability. Better battery systems have far-reaching benefits that affect many people by providing affordable and abundant energy.

End-of-life is defined as the point when a battery can only retain 80% of its initial capacity [2]. Lithium-ion batteries have high variability of end-of-life "Fig. 1.", and accurate end-of-life prediction is essential for the reliable operation of lithium-ion batteries [3]. Battery failure can have devastating consequences when used in important operations such as grid-scale support and large industrial systems. Batteries that are close to end-of-life are more likely to approach thermal runaway [4]. Malfunctions can be avoided by identifying premature failures and scheduling battery management and replacement. End-of-life is influenced by a number of factors related to the design and manufacturing process.

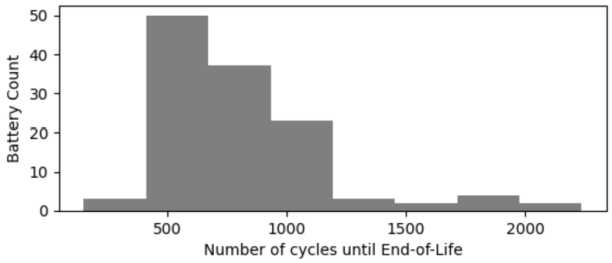

Fig. 1. High Variability in Battery Life, Severson et al. Data Set - End-of-life

Battery degradation is a complex electrochemical process that reduces the charge capacity via multiple concurrent mechanisms and highly irregular behavior. The principal mechanisms of degradation include lithium plating, solid-electrolyte interface growth, particle fragmentation, and positive electrode structure changes and decomposition [5]. Modeling and simulation of the physical decay of batteries requires a deep understanding of electrochemistry and material science, including the characteristics of the individual batteries.

Data-driven modeling is an alternative to modeling the physical decay of battery degradation [6]. The use of artificial intelligence and machine learning methods can be applied without the need for a deep understanding of each battery. Data-driven modeling is mechanism-agnostic, and the complicated electrochemical behaviors and conditions are captured in experimental data [2]. Machine learning can detect even delicate fluctuations in data and recognize patterns that affect the end-of-life of each battery. These nonlinear dynamics are normally very difficult to express mathematically.

Machine learning is increasingly being used in the fields of chemistry, engineering, and material science [7], yet data-driven modeling comes with its challenges. Lithium-ion battery life data is expensive because they have relatively long lives, making experiments last several months. Because of this, there is a sparsity of data, and machine learning often requires a large amount of data for models to accurately learn.

II. Battery Data Set

In 2019, Severson et al. [2] at MIT, published the largest open-source database of lithium-ion battery lifecycle data. The project was a combined effort of Stanford University, Toyota Research Institution, and Berkeley Livermore National Lab. The data set consists of 124 LFP-graphite cells that were charged and discharged until they reached end-of-life (80% of their initial capacities). The cells were tested under a variety of conditions, and 72 different charging protocols were used. The cells were tested in an environmental chamber so that the relative temperature changes could be observed. Current (A), temperature (T), voltage (V), and capacity (Ah) were recorded

throughout the life of the batteries "Fig. 2." Summary data and pre-calculated features were included in the raw data. Severson et al. provided additional pre-processing calculation and code including training/test split for reproducibility (see Data & Code section).

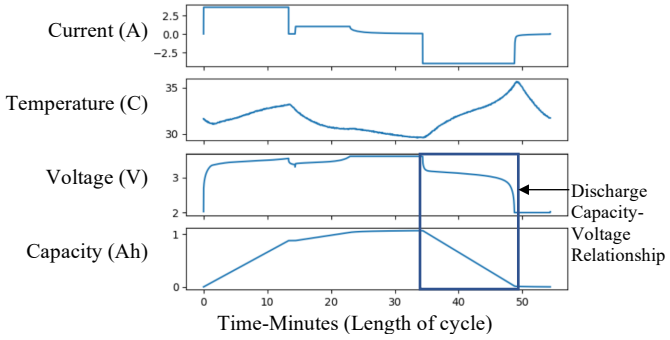

Fig. 2. Timeseries Data ,
Sample - cycle 1 of battery 'b1c0'

Attia et al. [8] produced a follow-up paper that acted as a more comprehensive benchmark for the data set. They focused strictly on the voltage capacity relationship and improved the interpretability of the machine learning models.

Tabular Summary Data

Severson et al. originally considered 20 features from the data and 3 models were presented to predict the end-of-life based on data from the first 100 cycles. The "variance model " used only one feature: the logarithm of the variance of the difference between the 100th & 10th cycle discharge voltage curves. The "discharge model" used data from only the discharge capacity values, interpolated at evenly spaced voltages (Table I).

TABLE I
SUMMARY FEATURES FROM DISCHARGE VOLTAGE CURVES USING THE 10TH AND 100TH CYCLES (ΔQ100-10(V))

- Minimum ($\Delta Q_{100\text{-}10}(V)$)
- Variance ($\Delta Q_{100\text{-}10}(V)$)
- Skewness ($\Delta Q_{100\text{-}10}(V)$)
- Kurtosis ($\Delta Q_{100\text{-}10}(V)$)
- Discharge capacity of cycle 2
- Max discharge capacity - Discharge capacity of cycle 2

Discharge Capacity - Voltage Relationship Data

The discharge capacity-voltage relationship can be represented as time series data [9]. Each cycle of the battery can represent a time step. Each voltage along the discharge curve represents the features of the learning model. Figure 3 demonstrates the decaying capacity as the battery ages (increased cycle count).

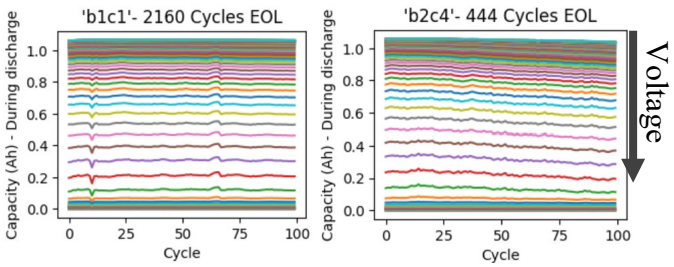

Fig. 3. Discharge Capacity-Voltage Curves
Comparison of batteries with high and low end-of-life

III. Machine Learning Models

A. *Elastic Net Regression*

Ordinary linear regression approximates the relationship between a target variable (i.e., number of cycles until end-of-life) and a set of explanatory variables. Coefficients of the linear equation are updated by minimizing a cost function (commonly Least-Mean-Squares). Regularization is applied to the cost function, allowing better generalization of the regression model. Elastic net regularization is a combination of 2 widely used methods: Lasso regularization (using the L1 norm), and Ridge regularization (using the L2 norm). By including both the L2 and L1 norm, shown in equation 1, the cost function penalizes the coefficients more and discourages learning an overfitted model.

$$\mathcal{L}(\omega, \alpha, \lambda) = \left| ||Y - X\omega||^2 + \alpha(\lambda||\omega||_1 + (1-\lambda)||\omega||_2^2) \right. \quad (1)$$

The parameters α and λ are chosen by cross-validation and control the trade-off between fitting the training data well and generalizing to new, unseen data. Severson et al. used elastic net regression on the 3 proposed models and is the first benchmark of the data set.

B. *Support Vector Regression*

Linear regression and elastic net regularization are restricted to linear representation of data. However, nonlinear curve fitting can be computationally expensive and prohibitive. To overcome this, support vector machines (SVMs) implement feature mapping where the features are transformed to a higher dimensional feature space. The trick to overcoming the expensive computation is the use of kernel functions in feature mapping which avoid explicit inner product calculations. Two common kernel functions are the sigmoid (eq. 2) and radial basis function (eq. 3).

$$\text{RBF Kernel:} \quad K(x, y) = e^{-}(\gamma||x - y||^2) \quad (2)$$

$$\text{Sigmoid Kernel:} \quad K(x, y) = tanh(\gamma.x^T y + r) \quad (3)$$

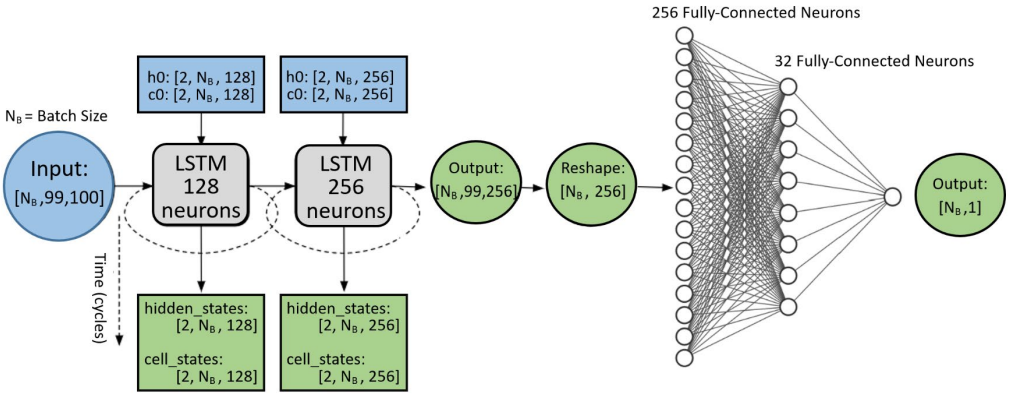

Fig. 5. Architecture of Long-Short-Term-Memory Recurrent Neural network

Support vector regression is a version of SVMs that predicts a real value (i.e., the end-of-life). In this work, 2 additional hyperparameters are chosen by cross-validation: epsilon (ε) and gamma (γ). Epsilon defines a distance in which there is no penalty in the cost function, where low values make the model more robust to outliers. Gamma is the coefficient of the RBF and sigmoid kernels (eq. 2, eq. 3).

C. *Multilayer Perceptron*

Multilayer perceptrons capture complex behavior with fully connected layers of artificial neurons (Fig. 4). Each neuron learns nonlinear behavior through the mechanism of activation functions (i.e., rectified linear unit). Learning is achieved by updating the weights and biases of each neuron, iteratively, with the backpropagation algorithm. The updates are governed by an optimization algorithm, commonly Stochastic-Gradient-Descent or Adaptive-Momentum-Estimation (Adam). Multilayer perceptron architecture has an interesting influence on the ability to learn. Regularization is implemented in the multilayer perceptron model by optimizing a coefficient (alpha) that determines the relative strength of the L2 norm term.

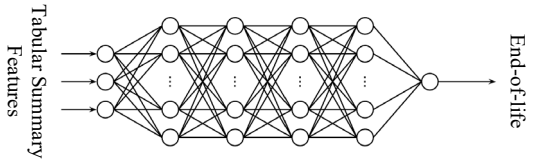

Fig. 4. Multilayer perceptrons are fully connected artificial neural networks

D. *Long-Sort-Term-Memory Recurrent Neural Network*

Recurrent neural networks are designed to process time series data and capture trends in training data. Long-Sort-Term-Memory Recurrent Neural Networks (LSTMs) have weights and biases, like other neural networks, but also have hidden and cell states (Fig. 5) which help capture long-term dependencies (including the ability to process variable length sequences). These additional memory units are updated with a series of calculations in each elementary unit of LSTMs (called a cell). Below are the calculations that determine each cell state (c_t) and hidden state (h_t). LSTMs also update with backpropagation and optimization algorithm (Adam). Weight decay is a hyperparameter that controls the strength of L2 regularization.

Xu et al. [9] demonstrated that LSTMs are capable of learning with less data (using data from only the first 60 cycles) when trained on the discharge capacity-voltage data.

$$
\begin{aligned}
i_t &= \sigma(W_{ii}x_t + b_{ii} + W_{hi}h_{t-1} + b_{hi}) \\
f_t &= \sigma(W_{if}x_t + b_{if} + W_{hf}h_{t-1} + b_{hf}) \\
g_t &= \tanh(W_{ig}x_t + b_{ig} + W_{hg}h_{t-1} + b_{hg}) \\
o_t &= \sigma(W_{io}x_t + b_{io} + W_{ho}h_{t-1} + b_{ho}) \\
c_t &= f_t \odot c_{t-1} + i_t \odot g_t \\
h_t &= o_t \odot \tanh(c_t)
\end{aligned}
$$

Fig. 6. The sequence of calculations at each cell in LSTM [10] (input gate, forget gate, cell gait, & output gate)

E. *Convolutional Neural Network*

The discharge capacity-voltage data does not have variable length sequences. Each sequence of discharge capacities is the length of the cycles used in prediction. A 2D representation of the discharge capacity-voltage data can be used to train a convolutional neural network (Fig. 7). Convolutional layers apply filters to the 2D data that help extract complicated patterns, and max-pooling layers reduce the dimensionality.

The "discharge" model in Severson et al. used features that come strictly from the discharge capacity-voltage relationship (see Table I). Since these features can be derived directly from the same 2D data, this proposes a variation in the architecture whereas the tabular summary features are concatenated with the flattened features of the CNN (see Fig. 7).

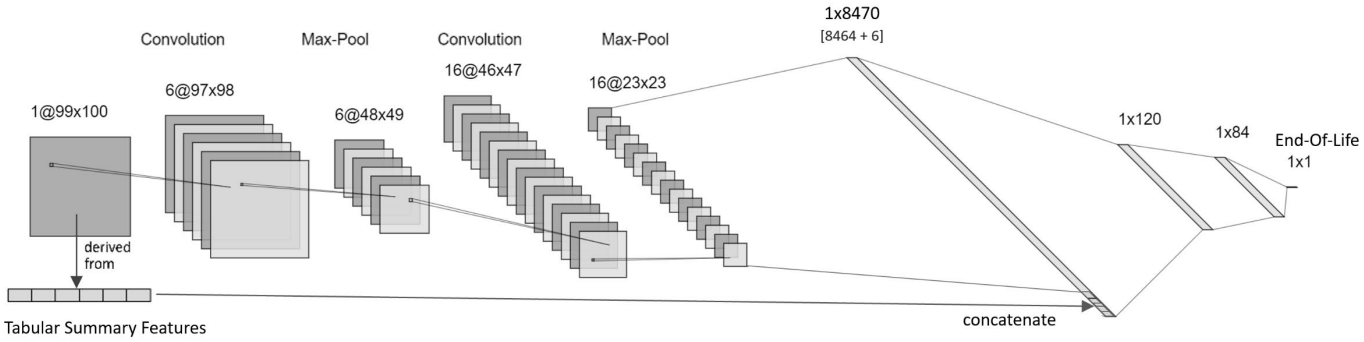

Fig. 7. Architecture of Convolutional Neural network

IV. Results & Discussion

The calculated end-of-life of lithium-ion batteries is evaluated by two metrics, the same as the benchmark studies [2,8]. The root-mean-square-error (RMSE) and the mean absolute error (MAE) compare each prediction to the observed value and are commonly used in statistics.

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|\widehat{y}_i - y_i\right| \tag{4}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\widehat{y}_i - y_i\right)^2} \tag{5}$$

Table II lists the results of the original paper [2] and compares the support vector regression and multilayer perceptron models, which are trained on the same tabular summary data. These models achieved similar predictive power.

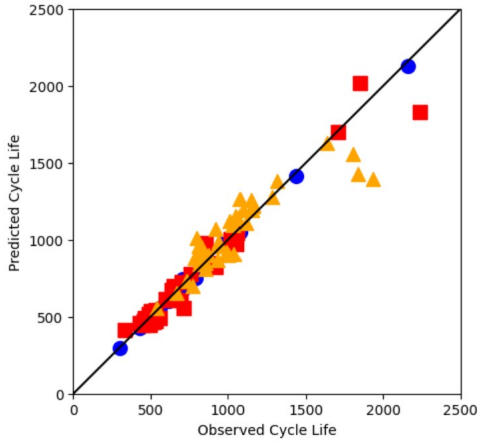

Fig. 8. Predicted vs Observed End-Of-Life
CNN - first 100 cycles

TABLE II
RESULTS FROM TABULAR SUMMARY DATA

MODEL	RMSE TRAIN	RMSE TEST	RMSE TEST #2	MPE TRAIN	MPE TEST	MPE TEST #2
Severson et al. Elastic Net "Variance" model	(103)	(138)	(196)	(14)	(13)	(11)
"Discharge" model	(76)	(86)	(173)	(10)	(10)	(9)
"Full" model	(51)	(100)	(214)	(6)	(8)	(11)
Support Vector Regression - "Discharge" features	79	79	172	10	10	9
Support Vector Regression - "Full" features	54	115	200	7	9	11
Multilayer Perceptron - "Discharge" features	**67**	**105**	**159**	**8**	**8**	**8**
Multilayer Perceptron - "Full" features	39	138	176	4	8	10

TABLE III
RESULTS FROM DISCHARGE CAPACITY-VOLTAGE DATA

MODEL	RMSE TRAIN	RMSE TEST	RMSE TEST #2	MPE TRAIN	MPE TEST	MPE TEST #2
Xu et al. [9] LSTM 100 cycles	(42)	(91)	(215)			
60 cycles	(48)	(88)	(212)			
40 cycles	(51)	(127)	(224)			
LSTM 100 cycles	35	101	193	1	11	14
60 cycles	38	316	220	2	18	17
40 cycles	49	127	218	3	12	17
Attia et al. [8] Convolutional Neural Network	(17)	(72)	(204)			
Convolutional neural network 100 cycles	**16**	**83**	**141**	**1**	**6**	**8**
60 cycles	14	184	203	1	8	12
40 cycles	11	80	206	1	6	11

The LSTM models also achieved results similar to the previous literature. Table III presents the results from the discharge capacity-voltage data. The convolutional neural network (with the tabular data concatenation) proved to have better learning capabilities from the same data. Figure 8 demonstrates the end-of-life predictions, with most of the error in high-lifecycle batteries.

Conclusion & Future Work

Lithium-ion battery R&D is rapidly growing and machine learning will certainly be a valuable tool in the future. End-of-life prediction has proven to be a challenging task in battery R&D. Dozens of papers have been published, attempting to create machine learning models from the Severson et al. data set [11–15].

While this work focuses on the discharge capacity-voltage relationship, the Severson et al. data set includes voltage, temperature, and current data (Fig 2). Two notable publications have used this additional information to predict end-of-life with only one cycle of data [16,17]. Each of these publications use convolutional neural networks with advanced and optimized architecture. One trade-off with using additional data is that capacity data "comes for free" [8], and is easier to obtain, whereas data from temperature requires additional equipment and may be prohibitive in future studies.

Data & Code

DATA AVAILABILITY

data.matr.io/1/projects/5c48dd2bc625d700019f3204

DATA PRE-PROCESSING CODE

(https://www.nature.com/articles/s41560-019-0356-8#code-availability)
github.com/rdbraatz/data-driven-prediction-of-battery-cycle-life-before-capacity-degradation

CODE

https://github.com/JamesChapmanNV/MachineLearning_Lithium-IonBattery_End-Of-Life

REFERENCES

[1] United States, Congress, Department of Energy, & Derasmo, W. R. (2021). National Blueprint for Lithium Batteries. https://energy.gov/eere/vehicles/articles/national-blueprint-lithium-batteries

[2] Severson, K. A., Attia, P. M., Jin, N., Perkins, N., Jiang, B., Yang, Z., Chen, M. H., Aykol, M., Herring, P. K., Fraggedakis, D., Bazant, M. Z., Harris, S. J., Chueh, W. C., & Braatz, R. D. (2019). Data-driven prediction of battery cycle life before capacity degradation. *Nature Energy*, *4*(5), 383–391. doi.org/10.1038/s41560-019-0356-8

[3] Hu, X., Xu, L., Lin, X., & Pecht, M. (2020). Battery Lifetime Prognostics. *Joule*, 4, 3-20. doi.org/10.1016/j.joule.2019.11.018

[4] Dragonfly Energy. (2022, December 16). *What is thermal runaway in batteries?* Dragonfly Energy. Retrieved April 16, 2023, from https://dragonflyenergy.com/thermal-runaway/

[5] Edge, J. S., O'Kane, S., Prosser, R., Kirkaldy, N. D., Patel, A. N., Hales, A., Ghosh, A., Ai, W., Chen, J., Yang, J., Li, S., Pang, M.-C., Bravo Diaz, L., Tomaszewska, A., Marzook, M. W., Radhakrishnan, K. N., Wang, H., Patel, Y., Wu, B., & Offer, G. J. (2021). Lithium ion battery degradation: What you need to know. *Physical Chemistry Chemical Physics*, *23*(14), 8200–8221. https://doi.org/10.1039/d1cp00359c

[6] Li, Y., Liu, K., Foley, A. M., Zülke, A., Berecibar, M., Nanini-Maury, E., Van Mierlo, J., & Hoster, H. E. (2019). Data-driven health estimation and lifetime prediction of lithium-ion batteries: A review. *Renewable and Sustainable Energy Reviews*, 113, 109254. doi.org/10.1016/j.rser.2019.109254

[7] Gao, T., & Lu, W. (2021). Machine learning toward advanced energy storage devices and systems. *iScience,* 24(1), 101936. doi.org/10.1016/j.isci.2020.101936

[8] Attia, P. M., Severson, K. A., & Witmer, J. D. (2021). Statistical learning for accurate and interpretable battery lifetime prediction. *Journal of The Electrochemical Society*, *168*(9), 090547. doi.org/10.1149/1945-7111/ac2704

[9] Xu, P., & Lu, Y. (2022). Predicting Li-ion Battery Cycle Life with LSTM RNN. *arXiv preprint* arxiv.org/abs/2207.03687

[10] Sak, H., Senior, A. W., & Beaufays, F. (2014). Long Short-Term Memory Based Recurrent Neural Network Architectures for Large Vocabulary Speech Recognition. CoRR, abs/1402.1128. Retrieved from http://arxiv.org/abs/1402.1128

[11] Ansari, S., Ayob, A., Lipu, M. S. H., Hussain, A., & Md Saad, M. H. (2022). Remaining useful life prediction for lithium-ion battery storage system: A comprehensive review of methods, key factors, issues and future outlook. *Energy Reports*, 8, 12153-12185. doi.org/10.1016/j.egyr.2022.09.043

[12] Ng, M. F., Zhao, J., Yan, Q., Chen, B., Wang, Q., Xia, Y., ... & Shen, Z. J. (2020). Predicting the state of charge and health of batteries using data-driven machine learning. *Nature Machine Intelligence*, 2(3), 161-170. doi.org/10.1038/s42256-020-0156-7

[13] Jin, S., Sui, X., Huang, X., Wang, S., Teodorescu, R., & Stroe, D. I. (2021). Overview of machine learning methods for lithium-ion battery remaining useful lifetime prediction. *Electronics*, 10(24), 3126. doi.org/10.3390/electronics10243126

[14] Fei, Z., Yang, F., Tsui, K.-L., Li, L., & Zhang, Z. (2021). Early prediction of battery lifetime via a machine learning based framework. *Energy*, 225, 120205. doi.org/10.1016/j.energy.2021.120205

[15] Hong, J., Lee, D., Jeong, E.-R., & Yi, Y. (2020). Towards the swift prediction of the remaining useful life of lithium-ion batteries with end-to-end deep learning. *Applied Energy*, *278*, 115646. https://doi.org/10.1016/j.apenergy.2020.115646

[16] Strange, C., & dos Reis, G. (2021). Prediction of future capacity and internal resistance of Li-ion cells from one cycle of input data. *Energy and AI*, *5*, 100097. https://doi.org/10.1016/j.egyai.2021.100097

[17] Hsu, C.-W., Xiong, R., Chen, N.-Y., Li, J., & Tsou, N.-T. (2022). Deep neural network battery life and voltage prediction by using data of one cycle only. *Applied Energy*, *306*, 118134. https://doi.org/10.1016/j.apenergy.2021.118134

[18] Murphy, K. P. (2022). *Probabilistic machine learning: An introduction*. MIT Press. probml.ai