Early Prediction of Lithium-ion Battery End-of-Life

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

***Abstract-* These instructions give you basic**

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 (EV's), 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[]. Accurate end-of-life prediction is essential for reliable operation of lithium-ion batteries [2]. Battery failure can have devastating consequences when used in important operations such as grid-scale support and large industrial systems. Malfunctions can be avoided by identifying premature failures and scheduling battery management and replacement. End-of-life is influenced by a number of factors intrinsic to the design and manufacturing process.

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

Data-driven modeling alternative to modeling physical decay of battery degradation. And prediction of end-of-life is a promising method. The use of artificial intelligence and machine learning methods can be applied without the need for a deep understanding of battery life. All of the complicated electrochemical processes can be considered a black box, with all of the behavior of battery decay captured in the data. Data-driven modeling is mechanism-agnostic, meaning machine learning models can capture all of the internal physics and operating conditions with just data collected such as voltage (V), resistance (IR), and temperature (T). Fluctuations and complex nonlinear. The effects such as lithium plating and electro degradation result in complex nonlinear fluctuations in data. Machine learning

Pattern recognition
The challenges in data-driven modeling, sparsity of data

And the need to reduce manufacturing cost result in
with highly nontrivial behavior.

Meanwhile, it is also a challenging task since capacity degradation of batteries is a complex and nonlinear process influenced by internal physics and operating condition

And have proven to be robust for the transition to clean energy,

prevent thermal runaway
Battery R&D is an active field focused on improving
intractable to model

second-life applications.
worldwide clean energy sustainable and you factoring
Second Life market

multiple concurrent degradation mechanisms
challenging to make a trade-off
when the battery should be removed or replaced
Interpretability Limitations in
thermal stress a

modeling black box
without figuring out the complex electrochemical proces
highly variable graph

Fig. 1. Description

II. Battery Data Set

A. *Figures and Tables*

In 2019, Severson et al. [] at MIT, produced and published the largest open-source database of lithium-ion battery lifecycle. 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 in an environmental chamber so that the relative temperature changes could be observed.

The cells were tested under a variety of conditions, chosen to mimic industrial and commercial use. 72 charging protocols were used

Voltage (V), resistance (IR), and temperature (T) were recorded and

Severson et al. also provided pre-processing calculation and codes including training/test split for reproducibility. Several pre-calculated features of the data are provided. Particularly Qdlin,. from the collected data 20 features were developed suitable for machine learning.

Attia et al. [] produced a follow-up paper that acted as a benchmark for the data set. They focused strictly on the voltage capacity relationship and was able to marginally increase the given abilities of models voltage vs discharge capacity curves $\Delta Q100-10(V)$.

The target of the machine learning is the end-of-life

the authors have published preprocessing code which includes the same test/trains splits allowing consistent results compared relatively high correlation coefficients

capacities under various fast-charging conditions ranging from 3.6 to 6 C in an

environmental chamber (30_C). The cells were then charged from 0% to 80%

SOC with one-step or two-step charging profiles (e.g., 6 C charging from 0% to

40% SOC, followed by a 3 C charging to 80% SOC). Then, all cells were charged

from 80% to 100% SOC with 1 C CCCV step to 3.6 V and subsequently discharged

with 4 C CCCV step to 2.0 V, and the cut-off current was set to C/50. During the

cycling test, the cell temperature was recorded and internal resistance was

obtained at 80% SOC.

Feature Engineering

This work also calculated other engineered features. In order to capture the timeseries information linear regression answer or vector regression use tabular feature representation. However, the predictive power of timeseries can be utilized instm

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 dependent explanatory variables. Coefficients of the linear equation are updated by minimizing a cost function. The most widely used cost function in regression analysis is the Least-Mean-Squares update rule. When the linear regression algorithm is applied to the training set, the result is an equation of a line that hopefully can be used to predict the target variable with only knowing the features.

In practice, a technique called regularization is applied that adjust the cost functi stic net regularization is a combination of 2 widely used methods called Lasso regularization which uses the L1 norm, and Ridge regularization which uses the L2 norm. By including both the L2 and L1 norm with coefficients The on, allowing better generalization of the regression model. Ela

Severson et al. [] variance model consisted of univariate simple linear regression.

Priors

Minimizes the objective function

$$\mathcal{L}(\boldsymbol{w}, \lambda_1, \lambda_2) = ||\boldsymbol{y} - \mathbf{X}\boldsymbol{w}||^2 + \lambda_2||\boldsymbol{w}||_2^2 + \lambda_1||\boldsymbol{w}||_1 \quad (1)$$

$$\min_{w} \frac{1}{2n_{samples}} ||X_w - Y||_2^2 + \alpha\rho||W||_1 + \frac{\alpha(1-\rho)}{2} ||W||_2^2$$

$$L_{enet}(\hat{\beta}) = \frac{\sum_{i=1}^{n}(y_i - x_i^T\hat{\beta})^2}{2n} + \lambda\left(\frac{1-\alpha}{2}\sum_{j=1}^{m}\hat{\beta}_j^{\ 2} + \alpha\sum_{j=1}^{m}|\hat{\beta}_j|\right)$$

$$\mathcal{L}(\omega, \alpha, \lambda) = ||y - X\omega||_2^2 + \alpha\lambda||\omega||_1 +$$

Avoid placing them in the middle

A. *Support Vector Regression*

Linear regression is a highly restricted algorithm and is unable to predict nonlinear relationships. However, nonlinear calculations can be prohibitive and computationally expensive. To overcome this, Support Vector Machines (SVM's) implement feature mapping techniques and represents with a different set of features. Position nonlinear kernel trick projects lower dimension to higher dimensional without the computationally expensive explosion support Vector machines used in classification finds the largest margin hyper plane, this hyper Lane represents a linear projection of

kernel

kernel

tol

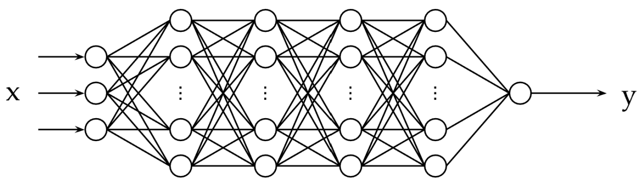
C

Regularization parameter. The strength of the regularization is inversely proportional to C. Must be strictly positive. The penalty is a squared l2 penalty.

gamma{'scale', 'auto'} or float, default='scale'

B. *Multilayer Perceptron*

Position figures and tab

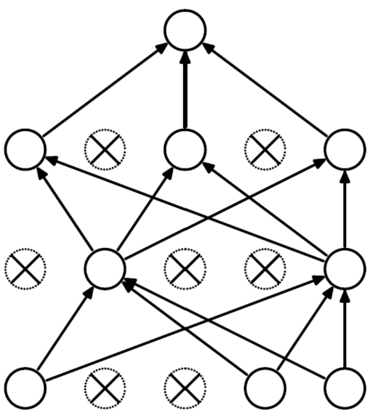


C. *Long-Sort-Term-Memory Recurrent Neural Network*

Position figures and tables at the

C. *Convolutional Neural Network*

Position figures and tables at the t



$$f^{(t)} = \sigma\left(W^{fX} x^{(t)} + W^{fh} h^{(t-1)} + b_f\right)$$
$$i^{(t)} = \sigma\left(W^{iX} x^{(t)} + W^{ih} h^{(t-1)} + b_i\right)$$
$$g^{(t)} = \tanh\left(W^{gX} x^{(t)} + W^{gh} h^{(t-1)} + b_g\right)$$
$$c^{(t)} = g^{(t)} * i^{(t)} + c^{(t-1)} * f^{(t)}$$
$$o^{(t)} = \sigma\left(W^{oX} x^{(t)} + W^{oh} h^{(t-1)} + b_o\right)$$
$$h^{(t)} = \tanh\left(c^{(t)}\right) * o^{(t)}$$

IV. Experiment

Position figures and tables at

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

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

Figure axis labels are often a source

V. Results & Discussion

Position figures and tables at the

TABLE I
Type Sizes for Papers

Type Size (pts.)	Appearance		
	Regular	Bold	Italic
6	Table captions,[a] table superscripts		
8	Section titles,[a] references, tables, table names,[a] first letters in table captions,[a] figure captions, footnotes, text subscripts, and superscripts		
9		Abstract	
10	Authors' affiliations, main text, equations, first letters in section titles[a]		Subheading
11	Authors' names		
24	Paper title		

[a]Uppercase

VI. Conclusion & Future Work

Position figures and t

equations

$$a + b = c. \tag{1}$$

Symbols in your

Data & Code

The data is Publicly available (https://www.nature.com/articles/s41560-019-0356-8#data-availability)
data.matr.io/1/projects/5c48dd2bc625d700019f3204
The author provides 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

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] 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

[3] 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

[4] 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

[] 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

[5] 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

[5] 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

[5] 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

[5] 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

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

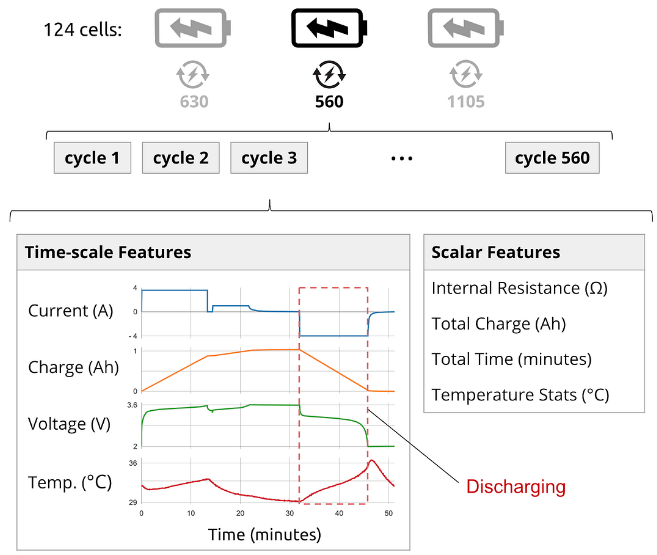
[5] Fei, Z., Yang, F., Tsui, K.-L., Li, L., & Zhang, Z. (2021). Early prediction of battery lifetime via a machine learning based framework. *Energy*, 225, 120205. doi.org/10.1016/j.energy.2021.120205

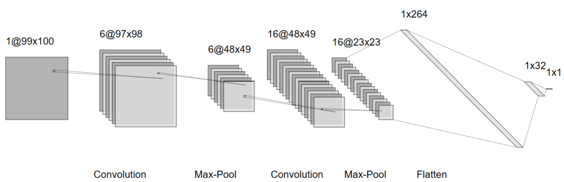
[5] Zhang, Y., Xiong, R., He, H., & Pecht, M.G. (2018). Long Short-Term Memory Recurrent Neural Network for Remaining Useful Life Prediction of Lithium-Ion Batteries. *IEEE Transactions on Vehicular Technology, 67*, 5695-5705. doi.org/10.1109/TVT.2018.2805189

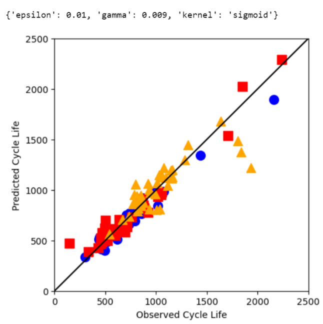
[] Edge, J. S., O'Kane, S., Prosser, R., Kirkaldy, N. D., Patel, A. N., Hales, A., Ghosh, A., Ai, W., Chen, J., Yang, J., Li, S., Pang, M.-C., Bravo Diaz, L., Tomaszewska, A., Marzook, M. W., Radhakrishnan, K. N., Wang, H., Patel, Y., Wu, B., & Offer, G. J. (2021). Lithium ion battery degradation: What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

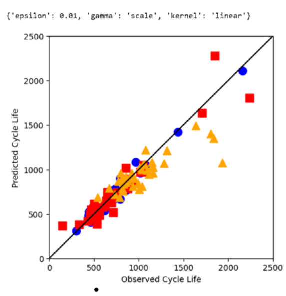
[] Murphy, K. P. (2022). *Probabilistic machine learning: An introduction*. MIT Press. probml.ai

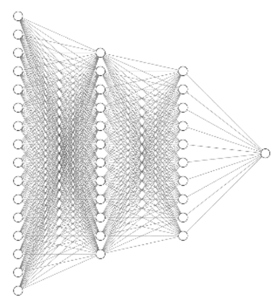
ElasticNet- univariate			
Support vector regression			
Multilayer perceptron			
LSTM			
CNN			











	Model	RMSE - Train	RMSE - Primary test	RMSE - Secondary test	MPE - Train	MPE - Primary test	MPE - Secondary test
0	Discharge model	78.9	91.9	171.8	9.5	14.5	9.5
1	Full model	53.7	118.7	199.7	6.6	12.0	11.1