WAYNE STATE
College of Engineering

IE7860 Winter 2022
Intelligent Analytics
Term Project

ESTIMATING THE REMAINING USEFUL LIFE OF LITHIUM-ION BATTERIES

By
Md Reza
Kirankumar Sivarajan

Abstract:
This project aims to predict lithium-ion batteries' remaining useful life (RUL) using recurrent neural network (RNN), long short-term memory (LSTM), and gated recurrent unit (GRU). Estimating RUL in the presence of the capacity regeneration phenomenon, we consider multiple measurable data from battery management systems such as voltage, current, and temperature charging profiles as their patterns change with the battery's aging. Unlike the traditional prediction approach that leverages a one-to-one structural system that matches the input layer with the output layer, we have leveraged a many-to-one structure to be flexible. In addition, we have experienced better generalization with this approach as the number of features was substantially reduced. Finally, given the data and cycle, our goal was to build different ANN models to predict the battery's remaining useful life (RUL) and evaluate the models' performance with model-driven metrics and outcomes.

Keywords:
Remaining Useful Life, RUL, Lithium-ion battery, long short-term memory, recurrent neural network, Gated Recurrent Unit, SHAP, Explainable AI, RUL Cycle Error.

Introduction:
Predicting lithium-ion batteries' useful life is crucial to managing the batteries' health and estimating the remaining useful life. Recent advancements in technology have developed the trends to build AI/ML-based processes in predicting the remaining functional life of the battery to unleash the power of electric vehicles (EVs), home appliances, tools, and energy storage devices in the automotive and manufacturing industries. "However, battery degradation begins immediately after batteries are manufactured, and when 70% or 80% of the initial capacity remains, batteries need to be replaced for safe operation. [9]" Therefore, estimating the battery's remaining useful life is crucial for uninterrupted operation, which could open the windows of possibilities in the space of EVs or other manufacturing industries.

About The Dataset:
"Experiments on Li-Ion batteries. Charging and discharging at different temperatures. Records the impedance as the damage criterion. The data set was provided by the Prognostics CoE at NASA Ames. Please refer to the README file attached to the dataset obtained from NASA's data repository. [2]"

Exploratory Data Analysis (EDA):

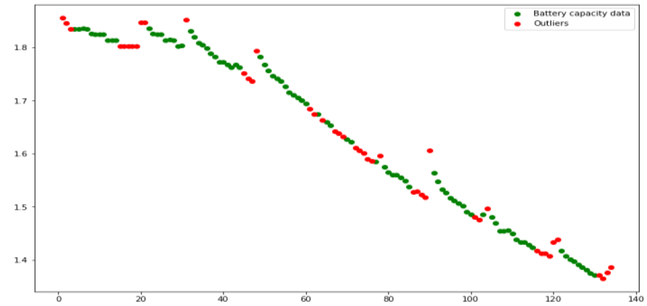
We leveraged voltage and temperature profiles from a given charging and discharge cycle to improve the performance of the proposed machine learning model. In addition to that, it was necessary to extract important features as well as create new features like maximum discharge temperature and maximum charge temperature. Please refer to the snippet on the right to learn more about the extracted dataset.

	cycle	capacity	max_discharge_temp	max_charge_temp
0	1	1.856487	38.982181	27.445134
1	2	1.846327	39.033398	29.341949
2	3	1.835349	38.818797	29.553301
3	4	1.835263	38.762305	29.456340
4	5	1.834646	38.665393	29.481334

Finally, different charts and visualization techniques are leveraged to validate if the datasets are loaded correctly. Please refer to the APPENDIX section for charts and visualizations.

In addition to that, we have leveraged exploratory data analysis (EDA) techniques that include:

- Outliers' detection
- Identify less important feature
- Scaling the data
- Perform data imputation
- Visualize predicted metrics
- Compare predicted metrics



Battery Capacity Degradation:

This dataset consists of three different operating profiles of charging, discharging, and impedance at room temperature. "The constant-current constant-voltage principle performs battery charging; charging is done with the constant current of 1.5A until the voltage reaches the limit of 4.2V. The voltage remains constant until the current drops to 20mA. To quantify battery degradation, state of health (SoH) is defined using capacity. [3]"

$$\mathrm{SoH}(\%) = \frac{Ck}{Co} * 100$$

Per the above equation, the *Co* is the initial capacity, and *Ck* is the measured capacity at cycle *k*. "The remaining useful life of the battery is defined when the remaining capacity is 70 to 80% of the initial capacity, depending on applications. [3]" The NASA datasets appear to have continuous capacity degradation at about 70% and <70%. Therefore, we have decided to use available datasets with maximum retention recorded from the last cycle before the battery degraded.

Data Acquisition from The Charging Process:

"During charging, lithium-ions escape from the electrode particles, and conversely, lithium-ions enter the electrode particles during discharging. As a result, lithium ions are scattered irregularly across the surface of battery particles. The larger the irregularity, the more the battery particles are affected and the shorter battery life. [3]" Thus, to predict the RUL of the battery, the characteristics of the charging or discharging process are identified using the data. However, it wasn't easy to measure or calculate the internal parameters during the discharge because of the distinctive usage pattern. And to capture internal battery parameters change along with aging, we leverage data from the charging cycle.

Methodologies:

This project aims to develop the traditional machine learning models using TensorFlow, Scikit-learn, Keras, SciPy, matplotlib, and seaborn packages. Since this is a time-series/sequential problem, we have decided to build the model with RNN, LSTM, and GRU, where RNN was the baseline model.

Outliers: Outliers are handled by averaging the data over the sampling interval to prevent oscillating short time intervals. Then min-max normalization is performed, dividing the data into training and test set.

Split data into training and testing data sets:
Python function was created for train-test split as well as data-frame conversion. Out of four datasets

- *B0005*, *B0007*, & *B0018* are used for training the models
- Training data: 474 sample

- *B0006* is used for testing the models
- Test data: 158 samples

Model Performance

- Perform permutation and reshape the data to 3 dimensions to improve model performance
- Performance of these models measured utilizing MSE, R-Squared Score, and calculated cycle error (i.e., the difference between actual vs. predicted RUL)

Recurrent Neural Network (RNN):

Recurrent neural networks, RNNs for short, are the variant of the conventional feedforward ANNs that can deal with sequential/time series data. RNNs recognize sequential characteristics of the dataset and use the patterns to predict the following likely scenario. In a nutshell, RNN holds the knowledge about the past to forecast future values. The recurrent neural network was compiled and fitted to the training data, and then the prediction was made on the test data.

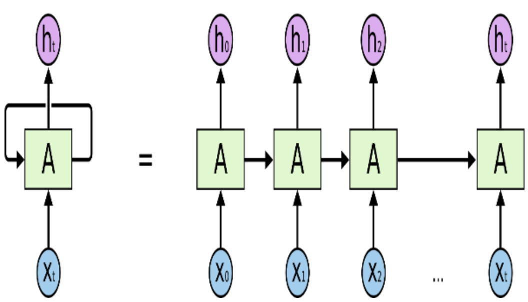
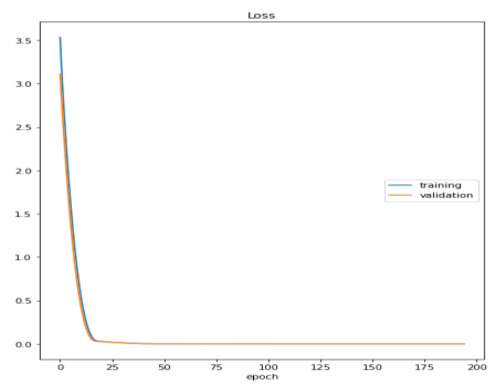
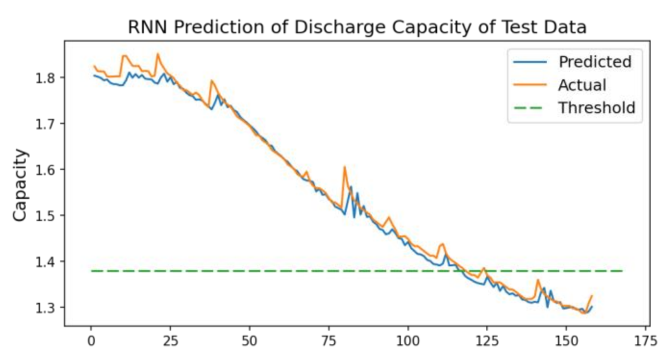


Image source: https://colah.github.io/posts/2015-08-Understanding-LSTMs/



Since this is a time-series/sequential problem, to improve model performance, we have leveraged the linear transfer function with RMSPROP optimizer, MSE loss, early stopping, and monitor value loss to build these time series problems. In addition to that, we have fitted the model with 200 epochs, batch size of 30, and 10% validation splits. As shown in the chart above with RNN, the training and validation loss consistently showed this model's robustness.

Finally, the graph on the left is self-explanatory, showing how the predicted and actual capacity behaves over time and supporting the data. For example, with RNN, the RUL error was 2, which indicates that the model went behind by two-cycle to estimate the RUL of the battery. Also, the MSE and R2 scores show that the model performance is promising.



RNN MSE & R-2 Score

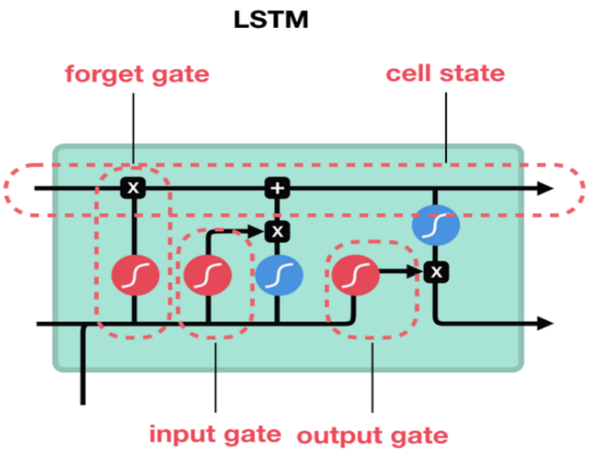
Metrics	Score
MSE	0.0004
R-2	0.9894

RNN RUL Cycle Error

Fail at Cycle Number	Number
Actual	116
Predicted	118
Error of RUL	2

Long Short-Term Memory (LSTM):

"An LSTM network is a *type of recurrent neural network* (RNN) that can learn long-term dependencies between time steps of sequence data, seeking better **context** management. LSTMs are explicitly designed to avoid the long-term dependency problem but require many training parameters. [1]" It turns out that with LSTM compared to RNN, the training and validation loss and model performance improved slightly.



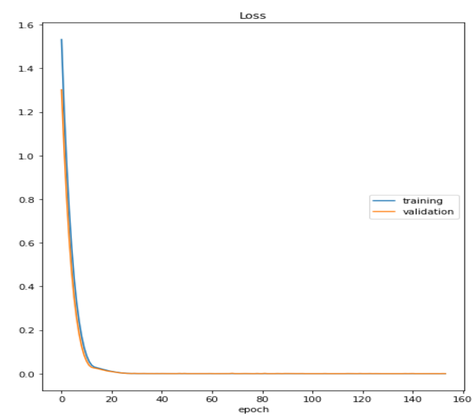


Image source: https://towardsdatascience.com/illustrated-guide-to-lstms-and-gru-s-a-step-by-step-explanation-44e9eb85bf21

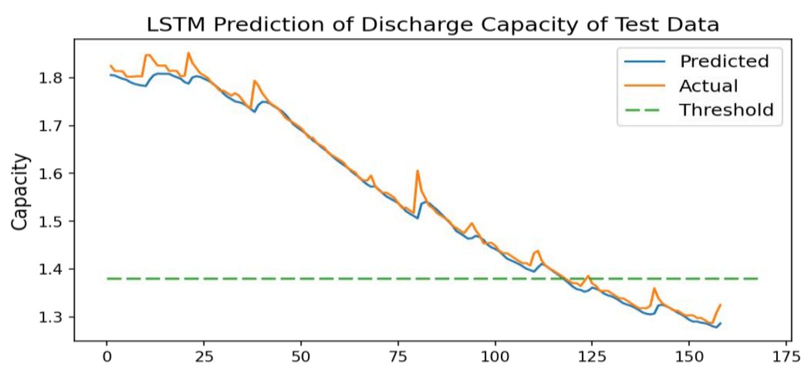
RNN MSE & R-2 Score

Metrics	Score
MSE	0.0003
R-2	0.9907

RNN RUL Cycle Error

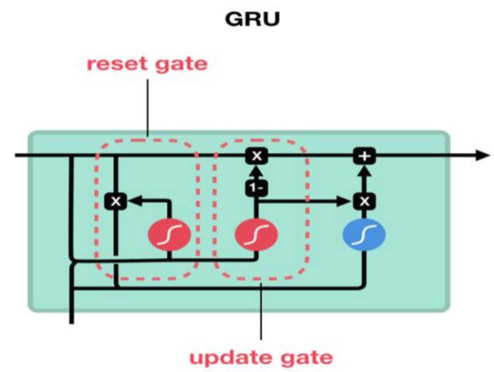
Fail at Cycle Number	Number
Actual	117
Predicted	118
Error of RUL	1

With LSTM based on the predefined threshold (I.e., the dotted line in the chart), the RUL cycle error improved slightly. In addition to that, compared to RNN, the MSE & R-2 scores also improved.



Gated Recurrent Unit (GRU):

"GRU is like an LSTM but simpler. It has no cell state and uses the hidden state to transfer information. It has only two gates: reset and update. GRU has fewer tensor operations and is a little speedier to train than LSTM.[1]" To overcome LSTM dependency on many training parameters, we leveraged GRU, and it turns out GRU outperforms the LSTM.



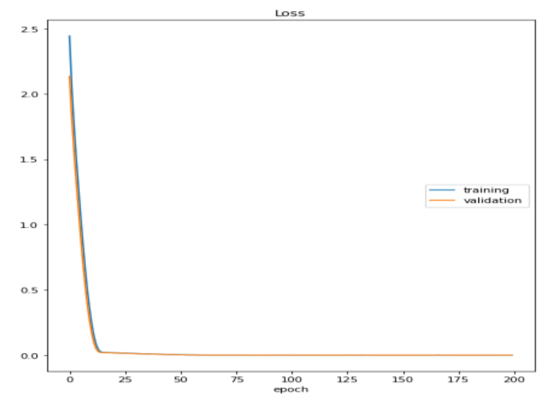


Image source: https://towardsdatascience.com/illustrated-guide-to-lstms-and-gru-s-a-step-by-step-explanation-44e9eb85bf21

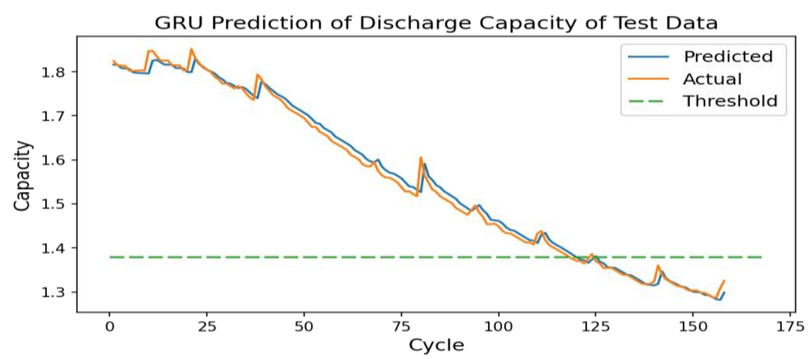
RNN MSE & R-2 Score

Metrics	Score
MSE	0.0002
R-2	0.9930

RNN RUL Cycle Error

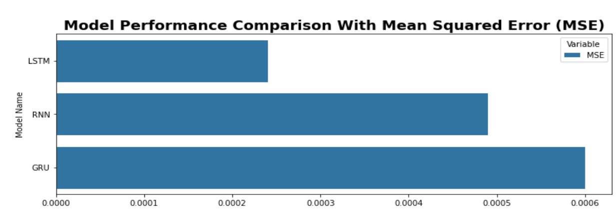
Fail at Cycle Number	Number
Actual	119
Predicted	118
Error of RUL	-1

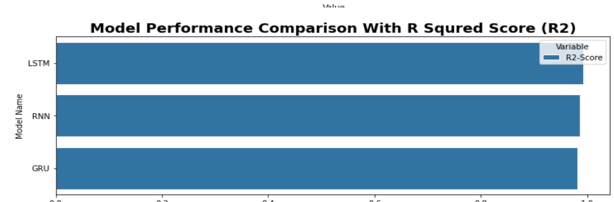
With GRU based on the predefined threshold (I.e., the dotted line in the chart), the RUL cycle error was negative one, which indicates that the model went ahead by one cycle to estimate the RUL of the battery. In addition to that, compared to RNN & LSTM, the MSE & R-2 scores improved slightly.



Model Selection:

Based on the MSE, R Squared Score, and calculated Cycle Error it was challenging to select a specific model for this time series exercise. Therefore, the verdict was no one was the winner; instead, select the model with the highest score and benchmark it. Use remaining models to support and validate the performance of the benchmarked model for estimating the remaining useful life of the battery. For a summarized details on model performance please refer to the model summary table under the appendix.





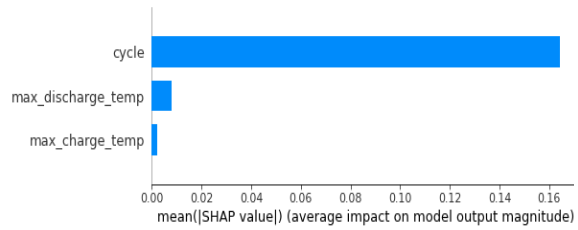
SHAP Explainable AI:

Explain Individual Prediction & Features Importance Using Explainable AI (SHAP):

With the advancement of AI/ML, Model explainability has become an integral part of the machine learning pipeline. Therefore, AI/ML should not be considered a “black box” anymore. Consequently, we have leveraged SHAP *(Shapley Additive Explanations)*, which connects optimal credit allocation with local explanations using the game-theoretic approach and classical Shapley values to explain individual predictions or output of the machine learning model. Some of the SHAP functionalities were leveraged to explain our models’ and predictors are as follows:

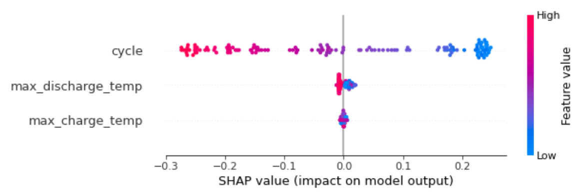
SHAP Value Plot:

Calculates the mean of absolute SHAP values for each feature across all observations. The chart on the left explains the cycle has the highest importance in predicting the RUL, followed by max discharge temperature, and the max charge temperature falls in the lowest bucket.



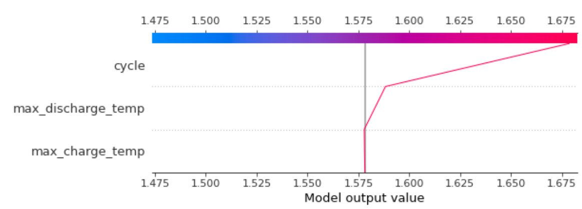
SHAP Summary Plot:

Overcame some of the issues derived from the SHAP Value plot. It combines feature importance with feature effects to highlight important relationships. Again, we observe the same phenomenon.

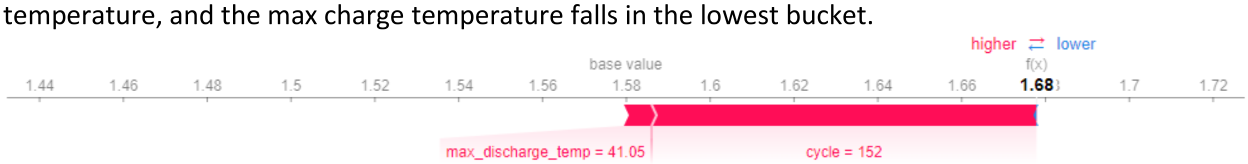


SHAP Decision Plot:

This shows the important features of a model’s output. Again, we observe the same phenomenon.



SHAP Force Plot: This shows how much each feature has increased or decreased the predicted number of rings for a specific observation based on the base value (i.e., 1.58). In this case, the final predicted number is 1.68. Based on that predicted number, we could observe that the explainability has the same phenomenon. The cycle is most important in predicting the RUL, followed by max discharge temperature, and the max charge temperature falls in the lowest bucket.



Conclusion:

This project proposed data-driven RNN, LSTM, & GRU-based ANN to predict lithium-ion batteries' remaining useful life. First, we proposed a many-to-one structure to improve prediction accuracy and showed how the structural changes capture the capacity regeneration phenomenon. And how it reduced the number of parameters for better generalization. We then created a charging profile with Cycle, Maximum Discharge Temperature, and Maximum Charge Temperature and leveraged these key features to predict the remaining useful life of the battery. Our proposed model-based machine learning ANN with the extracted subset of datasets has significantly improved RUL prediction.

Through model-based observation, GRU performed well compared to LSTM and RNN. Still, it was challenging to select a single model for this time-series exercise as these models' MSE, R Squared Scores, and Cycle Error was significantly close. So, the verdict was no one was the winner, but choosing the best model would depend on the dataset's characteristics, use case scenario, and other dependent circumstances. Or another option would be to benchmark the model with the best score and use the remaining models' outcomes to support or validate the performance of the benchmarked model.

REFERENCES:

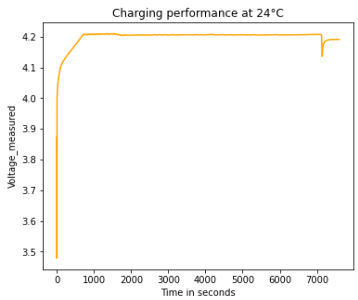
[1] Ratna Babu Chinnam, Ph.D., personal communication, n.d.
[2] B. Saha and K. Goebel (2007). "Battery Data Set", NASA Ames Prognostics Data Repository (https://ti.arc.nasa.gov/tech/dash/groups/pcoe/prognostic-data-repository/), NASA Ames Research Center, Moffett Field, CA.
[3] Park, Kyungnam & Choi, Yohwan & Choi, Won & Ryu, Hee-Yeon & Kim, Hongseok. (2020). LSTM-Based Battery Remaining Useful Life Prediction with Multi-Channel Charging Profiles. IEEE Access. PP. 1-1. 10.1109/ACCESS.2020.2968939.
[4] S. Voronov, E. Frisk and M. Krysander, "Data-driven battery lifetime prediction and confidence estimation for heavy-duty trucks.
[5] J. Qu, F. Liu, Y. Ma and J. Fan, "A neural-network-based method for RUL prediction and SOH monitoring of lithium-ion battery.
[6] X. Wang, C. Hu, X. Si, Z. Pang and Z. Ren, "An adaptive remaining useful life estimation approach for newly developed system based on nonlinear degradation model.
[7] Y. Zhang, R. Xiong, H. He and M. Pecht, "Long short-term memory recurrent neural network for remaining useful life prediction of lithium- ion batteries.
[8] Y. Song, L. Li, Y. Peng and D. Liu, "Lithium-ion battery remaining useful life prediction based on GRU-RNN.
[9] J. L. Zhang and J. Lee, "A review on prognostics and health monitoring of Li-ion battery.

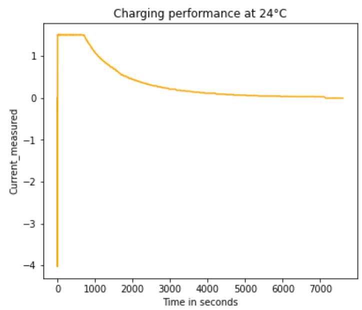
APPENDIX:

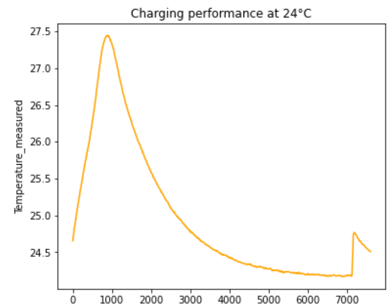
Model Summary

Model Name	MSE	R2-Score
GRU	0.00023712	0.992957
LSTM	0.000314476	0.990659
RNN	0.000358273	0.989358

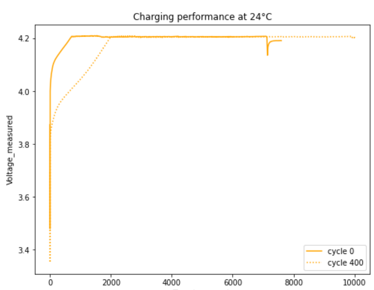
Charging Performance at 24°C:

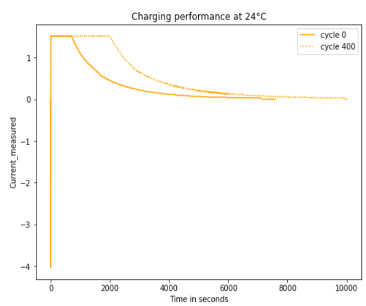


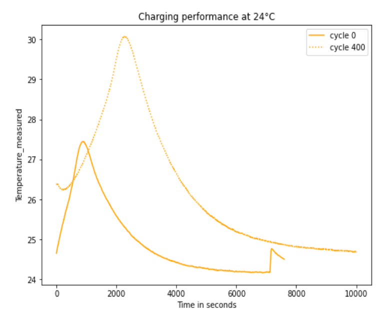




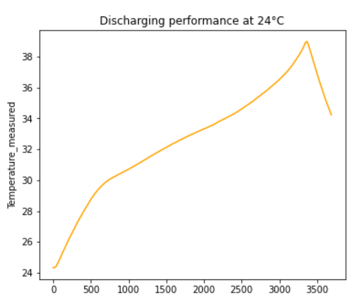
Charging Performance at cycle 0 & cycle 400 with 24°C

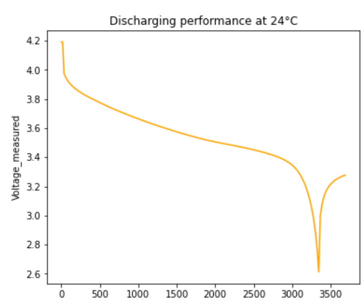


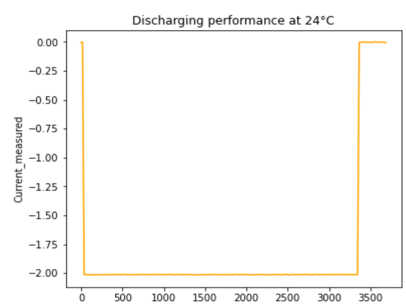




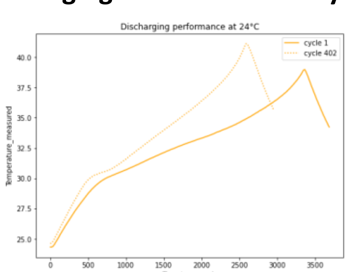
Discharging Performance at 24°C:

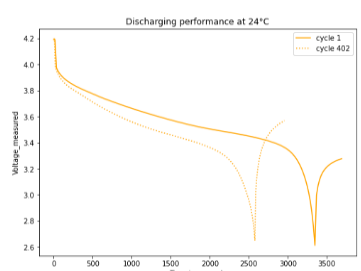


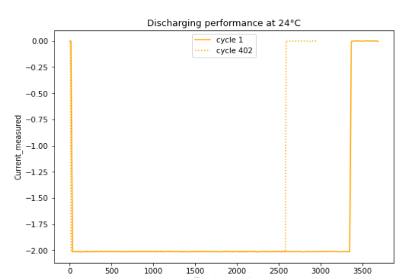




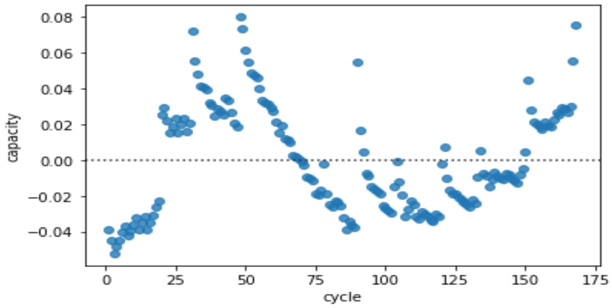
Discharging Performance at cycle 0 & cycle 400 with 24°C



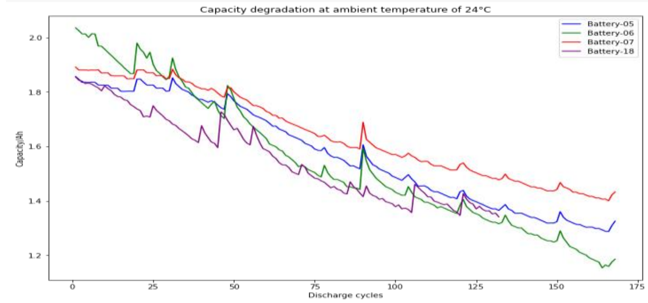




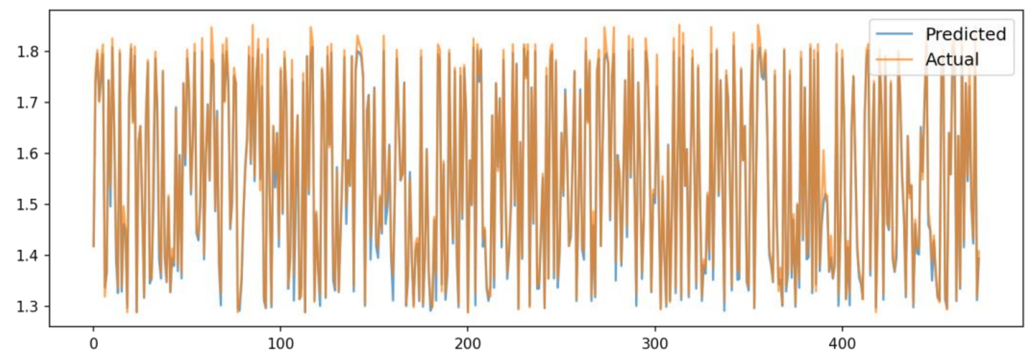
Residual plot for B0005 battery dataset: Capacity vs. Cycle



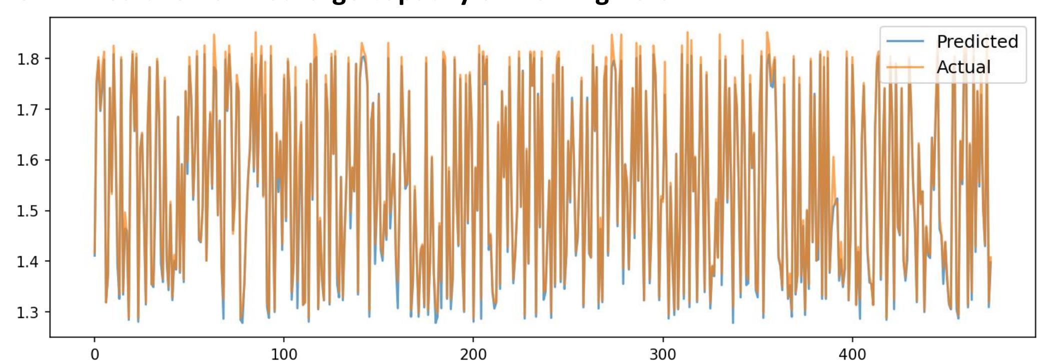
Validation of loading dataset with capacity degradation at the ambient temperature of 24°C



RNN Prediction of Discharge Capacity of Training Data:



LSTM Prediction of Discharge Capacity of Training Data:



GRU Prediction of Discharge Capacity of Training Data: