Feature Engineering for Machine Learning Enabled Early Prediction of Battery Lifetime

Noah H. Paulson[1,*], Joseph Kubal[2], Logan Ward[3], Saurabh Saxena[1], Wenquan Lu[2], Susan J. Babinec[4]

[1]Applied Materials Division, Argonne National Laboratory, Argonne, IL 60439
[2]Chemical Sciences and Engineering Division, Argonne National Laboratory, Argonne, IL 60439
[3]Data Science and Learning Division, Argonne National Laboratory, Argonne, IL 60439
[4]Argonne Collaborative Center for Energy Storage Science (ACCESS), Argonne National Laboratory, Argonne, IL 60439

Abstract

Accurate battery lifetime estimates enable accelerated design of novel battery materials and determination of optimal use protocols for longevity in deployments. Unfortunately, traditional battery testing may take years to reach thousands of cycles. Recent studies have shown that machine learning (ML) tools can predict lithium-ion battery lifetimes from 100 or fewer preliminary cycles, representing only a few weeks of cycling. Until now, conclusions about the efficacy and broad applicability of these predictions across a variety of cathode chemistries have been limited by available experimental information. In this work, we leverage a battery cycling dataset representing six cathode chemistries (NMC111, NMC532, NMC622, NMC811, HE5050, and 5Vspinel), multiple electrolyte/anode compositions, and 300 total carefully prepared pouch batteries to explore feature selection and battery chemistry's role in ML battery lifetime predictions. A mean absolute error (MAEs) of 78 cycles in prediction was seen for a chemistry-spanning test set from 100 preliminary cycles. Furthermore, an MAE of 103 cycles was seen when using only the first cycle. This study represents an in-depth investigation of strategies for feature selection for battery lifetime prediction, ML models' generalization across multiple battery chemistries, and predictions beyond the training set in the chemical space.

keywords: Lithium-ion; battery; life prediction; state of health; machine learning; feature engineering

1. Introduction

For the world to meet its ambitious emissions reduction targets with the urgency demanded by climate change, renewable energy and electrified transportation must rapidly achieve mass adoption. The key enabling technology for both is energy storage, whose performance and capital costs can often be met with the dominant technology, lithium-ion (Li-ion). While capital cost is a useful initial estimating metric, economic viability for deployment investments requires performance effectiveness over the entire asset life. Li-ion lifetime performance information is difficult to obtain with accuracy because its degradation is highly path dependent, with multiple nonlinear and intersecting failure modes across a complex function of chemistries, cell designs, use-cases, and environmental conditions. The complication is especially challenging for stationary storage applications which often require 20 years life with nearly complete daily utilization of available capacity across diverse use cases that are anticipated to evolve over time in response to climate change and grid strategies. Further, accurate Li-ion battery degradation prediction is critical not only for well-established chemistries and cell designs, but equally so in the development of new chemistries and cell designs which offer improved next generation performance or new supply chain options.

Experimental battery life testing requires significant resources and can take months to years to establish performance degradation estimates to a given end of life (EOL) definition. Consequently, predicting battery lifetimes from scarce initial characterization data generated in a short time span is of considerable interest to both academia and industry. Until recently, the state of the art for accelerated battery lifetime prediction methods used capacity and resistance trends over time (cycles) for a particular battery to project forward the battery's degradation profile [1–6]. A variety of mathematical functional models in a filtering framework [1] or machine learning models [2,3,7] have been employed for this task. Unfortunately, these methods have limited accuracy, especially for sudden accelerated capacity decay behavior ("knee point roll-over") seen for many batteries, and may require extensive degradation data to make reasonable predictions [8]. Physics-based models have also been used to predict battery lifetimes; however, parameter identifiability is a formidable problem that requires considerable effort to develop a model for a new battery chemistry or design, especially in the presence of multiple degradation mechanisms and interacting failure modes [9].

A promising option that uses large datasets and machine learning is now emerging. Recent reports demonstrate that machine learning can offer accurate and rapid prediction of degradation [8,10–18]. For example, Severson et al. in 2019 offered one significant demonstration with a dataset of 133 lithium-iron-phosphate (LFP) commercial batteries from A123 Systems [8] predicting battery lifetimes (as cycles to 80% of the nominal capacity) out as far as 2,000 cycles with just the first 100 cycles in different fast-charge scenarios. The study demonstrated that a simple regularized regression approach with *hand-built features* can achieve high accuracy predictions, with a root mean squared error (RMSE) of ~100 cycles. The potentially transformative AI/ML approach continues to develop with reports having diverse strategies and objectives. For example, subsequent papers have demonstrated the prediction of knee-point rollover on the same dataset (but using only 50 initial cycles of testing data) [10] and Strange and dos Reis predict the entire capacity degradation curve with only one preliminary cycle (though with limited accuracy) [19]. Together, these studies represent a considerable advancement over the state of the art, as fewer than 100 charge-discharge cycles may only take ~3 weeks as opposed to the years to reach EOL.

The promise of AI/ML, however, has been attenuated by the scarcity of data - a situation which will ultimately be resolved by creative approaches for more widespread data sharing [20]. Until such a time, ML prediction models using early life data continue to largely be demonstrated using single-origin datasets with objectives reflecting the profile of the available data.

In this work, we utilize a relatively large database of cell data collected over a decade of research on somewhat diverse topics including effect of electrolyte composition and additives, effect of cycling rates, electrode design, cell construction optimization, and formation protocol optimization. The collection allows us to speak to important questions which were previously inaccessible. For example, much ML prior art is focused on fast-charge behaviors in which Li-plating dominates failure and renders other mechanisms somewhat irrelevant. *A contrasting, important question is whether an ML algorithm can effectively predict* *cycle life when a broad mix of failure mechanisms are contributing. With this dataset we are not only able to address this question but additionally to explore the extent to which cathode choices impact a carefully designed all-purpose life-prediction algorithm.* Such a tool will be useful for both research and field deployments in the present data-constrained situation.

Using a series of selection criteria representing this overarching objective, a subset of 300 pouch Li-ion cells was selected from a large historic pool. These cells were constructed with graphite anodes and a broad variety of oxide cathode chemistries spanning $LiNi_{0.33}Mn_{0.33}Co_{0.33}O_2$ (NMC111), $LiNi_{0.5}Mn_{0.3}Co_{0.2}O_2$ (NMC532), $LiNi_{0.6}Mn_{0.2}Co_{0.2}O_2$ (NMC622), $LiNi_{0.8}Mn_{0.1}Co_{0.1}O_2$ (NMC811), $0.5Li_2MnO_3 \cdot 0.5LiNi_{0.375}Mn_{0.375}Co_{0.25}O_2$ (HE5050) [21,22], and $LiNi_{0.5}Mn_{1.5}O_4$ (5Vspinel) [7].–With this database we explored the effectiveness of ML modeling approaches for accurate early predictions when a dominant failure of lithium plating was not significant. We compare strategies for optimizing model accuracy when training with a broad range of cathodes and cell designs and, conversely, how well a model performs when making predictions for unseen cathode chemistries. The training data in this study offers all of the well-chronicled degradation modes [23,24], except for Li-plating (which can distort contributions from the cathode); this possibility is greatly diminished by restricting charging rates to less than 1C. We generate and explore a large set of diverse features for predicting the number of cycles to 80% of the nominal capacity using the voltage and current versus time information collected during single cell cycling experiments at controlled temperatures. Features range from the simple, such as the changes in capacity and the coulombic efficiency, to the more sophisticated, such as the locations and magnitudes of dQ/dV and dV/dQ peaks and the parameters of fitted equivalent circuit models to the cycling data. We explore preferences for various features with different sets of cathode chemistries. Finally, we demonstrate that ML models spanning the oxide cathode chemical space can predict the lifetimes for batteries with cathode chemistries not represented in the training set. *This is especially useful in materials discovery research, where novel cathode compositions are often explored in pursuit of improved performance without expectation of cycling durability.*

The paper is organized as follows: Section 2 describes the details of chemistries, cell designs and cycle protocols for the battery testing dataset, the approach taken to develop and select for features, and the final machine learning approach, Section 3 presents the results of the feature selection and machine learning predictions, Section 4 discusses the results and their relation to existing literature, and Section 5 draws the final conclusions.

2. Design of Battery Dataset

Argonne's lithium-ion dataset consists of over 1000 pouch cells manufactured and tested in the Argonne Cell Analysis, Modeling, and Prototyping (CAMP) facility over the course of 12 years, representing over 40 different pouch cell builds, used for a variety of investigations. At this DOE supported facility, researchers can manufacture full-size prototype battery electrodes and cells (pouch and 18650 formats) using pilot-scale semi-automated equipment in a dry room environment. Battery capacities range from 0.2 to >3 Ah and include a variety of cathode, graphitic anode, and electrolyte compositions. Coating, assembly, and formation protocols mimic industrial processes and are executed only by trained personnel to ensure reproducibility which is tracked with established quality controls techniques. This standardized approach enables efficient transitioning of new advanced battery materials and designs from research to industrial production.

Complex path dependent performance arises from the fact that lithium-ion has many failure mechanisms, which each have different sensitivities to cycling, manufacturing and environmental factors. Different combination of cell chemistry and design, use temperature, and cycling patterns give rise to unique blends of the various failure mechanisms and hence unique cycle life outcomes; this results in high cycle-life-testing complexity. Identifying and understanding these mechanisms continues to be an important area of intense study. Commonly cited failure mechanisms include SEI growth and decomposition, electrolyte decomposition, binder decomposition, current collector corrosion, and dendrites from lithium plating, see several recent papers and reviews [23,25–27]. The need for fast-charging of lithium-ion in electrified mobility has recently resulted in many studies of failures due to lithium plating.

The dataset is designed to reflect the technical objective of this study, which is to design an algorithm that predicts cycle life when a broad mix of failure mechanisms are contributing. Many design variables impact the mix of failure mechanisms, for example electrode design, electrolyte formulation, details of materials selections such as particle size and shape, and cycle protocol; each of these details are important and contribute to cycle life. However, this dataset has a large range of cathode compositions, many of which were designed for improved energy density but not cycle life - which is in contrast to the existing art for a single source of data. As such, the analysis was segmented by cathode chemistry and the effectiveness of a fully optimized feature design and ML-based life prediction approach was evaluated across 6 different cathode chemistries. Three-hundred candidates were selected using the following criteria:

- All anodes have graphite as the active material
- Charging rates are less than or equal to 1C, to minimize contributions of Li plating to battery degradation
- Performance is tested to at least 100 cycles, since this further reduces likelihood of lithium-plating. In future work, the 100-cycle cut-off may be relaxed should only 1st cycle features be used (see Sec. 3.1 Feature Development).

An important point is that, even with this condensed dataset, there still exists a wide range of differences between the cells and the corresponding data. For example, for the same cathode, the electrode properties such as the porosity or loadings are sometimes different; the supplier of the material may be different as well. Changes in voltage windows, electrolyte additives, and cycling procedures all play a role in the degradation, but separating by cathodes was deemed to be the first logical step to test the ML. A summary of the cells is available in the Supplementary Information in Table 3.

Figure 1a shows the breakdown by cathode chemistry of cells meeting the above criteria, and from this each of the different cathode chemistries provides a relatively large range of cycle lives. Figure 1b depicts the voltage vs. capacity discharge plots for fresh cells and those at 80% of initial capacity, using NMC-532 as the proxy for all NMCs. Figure 1c plots the lower and upper cutoff voltages, which were used to limit the charge and discharge operations of the batteries during cycling. Figure 1d shows the range of C-rates employed, with the vast majority at C/2 or lower, and a few at 1C; these rates are unlikely to plate Li metal.

Only cycles with a regular constant-current (and constant-voltage if triggered) charge and discharge were used in this study. Rate tests and other interruptions (including pauses and Hybrid Pulse Power Characterization tests) were removed from the cycle life data by selecting only those cycles where the median current in charge or discharge deviated from the median current across all cycles by less than a factor of 2. In just a few cases, the charge and discharge capacities were not close (within 5%) – in which case the cycle was not used. While these non-standard cycles might degrade the battery at a different rate, they represent a small fraction of the total life of the battery, and therefore the effect of their exclusion is assumed to be negligible, and their contributions reflected in the features (see Sec. 3.1 Feature Development).

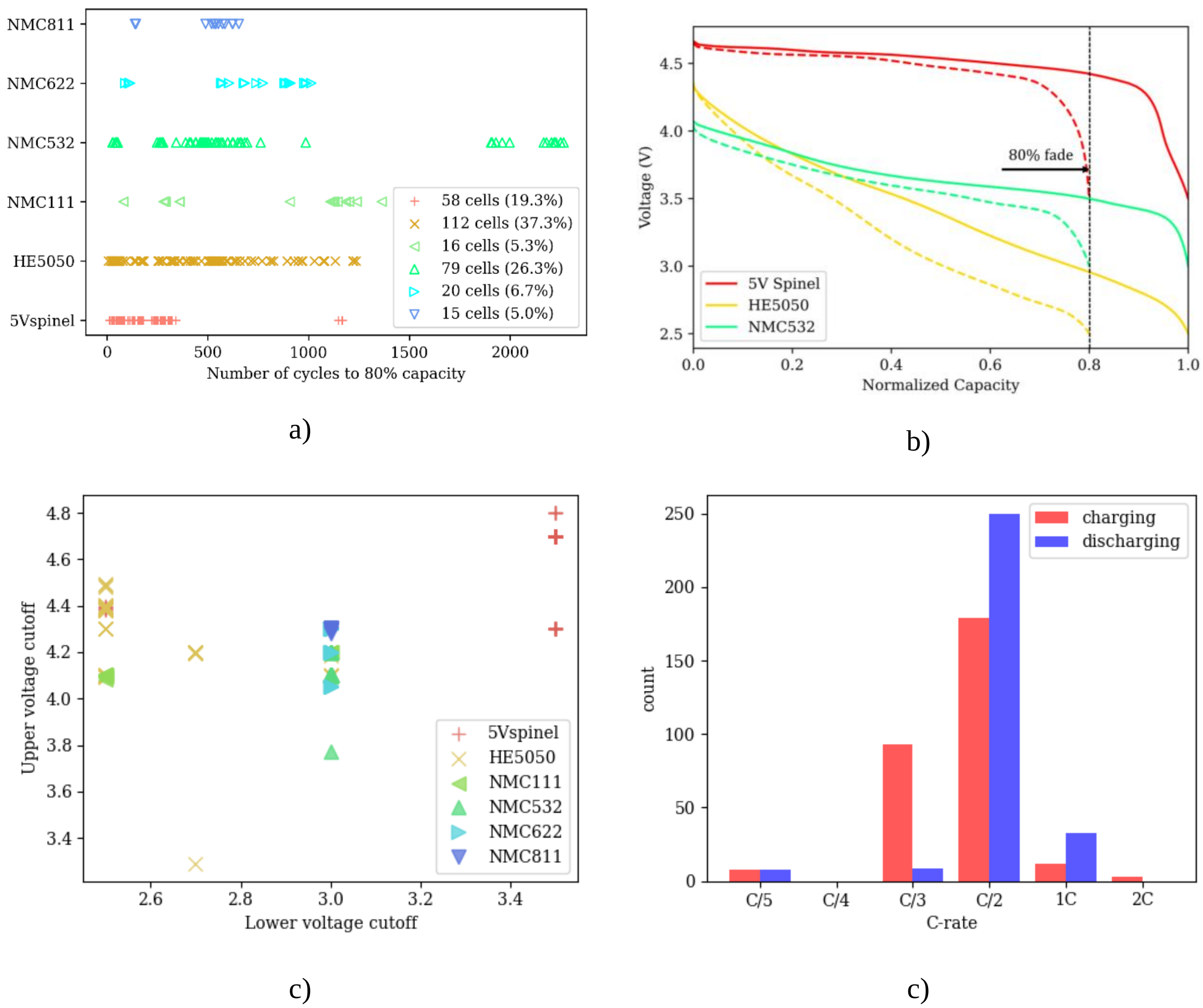


Figure 1: a) A chart showing the percentages, counts, and cycle life distributions for each cathode type, b) voltage vs. normalized discharge capacity initially and at 80% of the original capacity, for the three cathode chemistries, using NMC 532 as a proxy for all NMCs, c) battery chemistries plotted with respect to their lower and upper cutoff voltages, and d) a histogram of the charging and discharging rates for the dataset.

3. Methods

3.1 Feature Development

The literature describes features as attributes or properties of the data that are useful descriptors for machine learning algorithms. Our approach was to describe the voltage and current versus time data with an *exhaustive list of features*, and then down select to those most effective for predictions across broad cathode chemistries. Temperature-based features were considered but temperature measurements were not uniformly available for much of the battery dataset. Many features were inspired by the original Severson life prediction work and by the Pan et al. study addressing battery state of health (SOH) [28]. In this work, feature extraction was a multistep workflow that included file conversion, per cycle feature calculations, and multi-cycle feature calculations. First, the raw MACCOR output files were converted into a convenient hdf5 file format using the open-source Battery Data Toolkit [29]. Features were then calculated and tabulated for each cycle; Table 1 is an exhaustive list of all the feature types employed in this work. The < > brackets indicate a descriptor of the feature that is subject to change, for example ‘ch’ for charge versus ‘di’ for discharge or a specific voltage or SOC. Note that certain features are defined for a voltage range that exceeds that of individual batteries, in which case the value of the feature is given as the value of that feature at the voltage boundary. The name of the feature describes the process used to generate its value. There are several features that require further explanation, including dQ/dV features and features extracted from an equivalent circuit model - details are in the supplementary information. In the case where a feature was undefined, a not a number (NaN) value was given. For example, if for a given dQ/dV curve there was only one peak, the features associated with the second peak would be assigned as NaNs.

Multi-cycle features (which depend on feature values from multiple cycles) were then constructed from most of the per cycle features. As an additional appendage to the features in Table 1 are the y_0, y_{100m0}, and y_{diff} modifiers, which approximately encode the median, derivative with respect to cycles, and curvature of the original features with change in cycles, respectively. These multi-cycle features are calculated as follows,

$$\begin{aligned} y_0 &= med\{F_i, i = 0\,..\,9\} \\ y_{50} &= med\{F_i, i = 44\,..\,53\} \\ y_{100} &= med\{F_i, i = 90\,..\,99\}, \end{aligned} \tag{1}$$

$$y_{100m0} = y_{100} - y_0, \quad (2)$$

$$y_{diff} = (y_{100} - y_{50}) - (y_{50} - y_0), \quad (3)$$

where F_i is a given feature at cycle i and $med\{\,\}$ is the median operation. The median operation over multiple cycles is intended to remove the influence of outliers in the feature values for individual cycles. The cycle dependent features of Table 1 are calculated separately as they are directly inspired by the work of Severson et al. and therefore do not use the y_0, y_{100m0}, and y_{diff} modifiers. *In total, 396 features were created for this study from those in Table 1 combined with the feature generation rules defined by Equations 1-3.*

Table 1: Naming conventions and descriptions of per-cycle and cycle-dependent features are provided. < > brackets indicate that the given quantity is variable. State refers to charging (‘ch’) or discharging (‘di’).

Core Features	Description
cycle_number	cycle number
Imed_<state>	median current
Vavg_<state>	average voltage
Q_<state>	cumulative capacity
E_<state>	cumulative energy
Qeff	coulombic efficiency
Eeff	energy efficiency
Inorm_<state>	median current normalized by lifetime median current
SOH	state of health – capacity as a fraction of initial capacity
t_<state>	charge or discharge time
tCC_<state>	constant current charge or discharge time
tCV_<state>	constant voltage charge or discharge time
tCCvsCVfrac_<state>	tCC divided by total time
TIEVC_<state>	time interval during equal voltage change (3.5V-4.0V)
Vstart_<state>	voltage at start of segment
Vend_<state>	voltage at end of segment
dVdtStart_<state>	voltage derivative with respect to time at start of segment
dVdtEnd_<state>	voltage derivative with respect to time at end of segment
VoltsAt<SOC>pctSOC_<state>	voltage at specified SOC
dVdtat<SOC>pctSOC_<state>	voltage derivative with respect to time at specified SOC
VCET<time>s_<state>	voltage change during equal time between t=0 and specified time
OCVat<SOC>pctSOC	open circuit voltage at specified SOC from equivalent circuit model fit
Rat<SOC>pctSOC	resistance at specified SOC from equivalent circuit model fit
SOCat<voltage>V_<state>	SOC at specified voltage
dQdVpeak<num.>loc_<state>	location of specified peak in dQdV curve
dQdVpeak<num.>mag_<state>	amplitude of specified peak in dQdV curve
dQdVpeak<num.>area_<state>	area under specified peak in dQdV curve
dQdVvalley<num.>loc_<state>	location of specified valley in dQdV curve
dVdQvalley<num.>mag_<state>	amplitude of specified valley in dQdV curve
dVdQpeak<num.>loc_<state>	location of specified peak in dVdQ curve
dVdQpeak<num.>mag_<state>	amplitude of specified peak in dVdQ curve
dVdQvalley<num.>loc_<state>	location of specified valley in dVdQ curve
dVdQvalley<num.>mag_<state>	amplitude of specified valley in dVdQ curve
dSOCdCyc<statistic>_<state>	specified statistic applied to change in SOC vs voltage distribution between cycles 100 and 0
dSOHdCycCyc<cycle>	change in SOH vs change in cycle at specified cycle

3.2 Machine Learning and Feature Selection

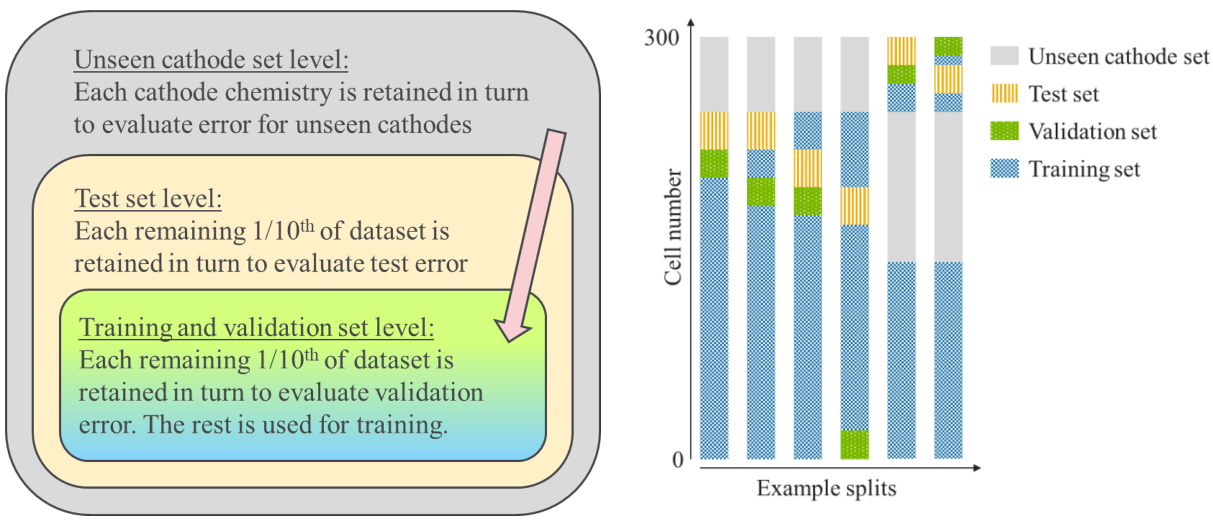
The electrochemical and computational scientific objectives in this study are tightly intertwined - the overall machine learning strategy consists of feature generation, specifics of the models used, and feature selection strategies based on preliminary life prediction metrics. The electrochemical objective is to devise efficient algorithms which accurately predict life across a range of oxide cathode chemistries. This eases the life prediction bottleneck in energy storage research - especially materials discovery. The computational objective is to create an efficient approach to broadly explore and then rank a feature space directed towards this life prediction goal - if too many features are needed, the algorithm will be overly complex, which may limit broad utility. *As the focus of this study is on the comparison of machine learning performance with different feature sets and cathode chemistries, we do not perform an exhaustive comparison of all possible machine learning approaches and feature selection strategies. Instead, we chose well performing machine learning models and a single feature selection strategy.*

The 300 available batteries were divided into different sets to enable rigorous training and validation of the overall machine learning approach. These are named the *unseen cathode set*, the *test set*, the *validation set*, and the *training set*. The purposes of these different sets are delineated as follows:

- *Unseen cathode set:* to evaluate predictive performance on cathode chemistries not used to develop the machine learning model
- *Test set:* to evaluate predictive performance of batteries not used to develop the machine learning model for cathode chemistries similar to the training set
- *Validation set*: to evaluate and compare different machine learning algorithms and feature sets in the process of training the machine learning model
- *Training set*: to train the machine learning model

Instead of having a fixed split between the sets, batteries were iteratively swapped between sets to ensure that each battery served in each set at least once. This procedure is described as follows and is illustrated in Figure 2a and Figure 2b.

At the outermost level, the entire battery dataset was split into six divisions according to cathode chemistry. In practice, this meant that batteries with one cathode chemistry at a time served as the *unseen cathode set* and were reserved for evaluating the predictive performance on a novel cathode chemistry. In cases where the chemistry generalizing performance was not of interest, this outer loop was omitted. At the next level down, the remaining dataset was split into 10 random, equal-sized divisions. Again, one division at a time was retained as the *test set*. In the most inner level, the remaining data was again split into 10 random, equal-sized divisions. One division at a time was retained in this case as the *validation set* with the rest reserved as the *training set*. In the machine learning literature, this inner level is called K-fold cross-validation (CV) which allows the model designer to tune and select an appropriate ML model. This optimized ML model is then used for performance evaluation on the *test set* and *unseen cathode set*. The inner level has 10 iterations, the intermediate level has 10 iterations, and the outer level has 6 iterations for a total of 600 unique data set splits (as illustrated in Figure 2b).



a) b)

Figure 2: a) A schematic illustrating the multi-level strategy for evaluating performance employed in this work is shown. From the outer through to the inner loops less data is available for a given evaluation. b) Example dataset splits are displayed with different batteries serving in the unseen cathode set, test set, validation set, and training sets.

Numerous machine learning models were explored, including lasso regression, random forests, Gaussian process regression, and artificial neural network regression; of these, only Nu Support Vector Regression (NuSVR) and Extremely Randomized Trees (Extratrees) regression [30,31] adequately balanced predictive performance and computational cost. For example, random forest regression was outperformed by Extratrees, and while neural networks made accurate predictions, the training time prohibited certain classes of feature selection algorithms (including sequential feature selection). For brevity, only the results from the two most successful life prediction ML models – Extratrees and NuSVR – have been discussed in this report. NuSVR, available through the open-source Sci-kit Learn python package [32], is a version of support vector regression where a nu parameter controls the number of support vectors employed. Extratrees, also available through Sci-kit Learn, is an ensemble decision tree approach where a random subset of candidate features is used in each ensemble member in addition to a randomized threshold for the splitting rule. The details of these approaches are beyond the scope of this work but are described in the literature [31]. Hyperparameter settings for the two ML models were determined through trial and error and were employed throughout the remainder of this work.

Strategy in feature selection reflects the objective - too many features lead to a model that is too complex and tailored to the training data, which means it is unlikely to generalize to new data. It is necessary to down select features to achieve good overall predictive performance and to enable the drawing of physical insights. A suite of feature selection strategies was explored; simple approaches studied included univariate feature selection via Pearson correlations and feature-importance-based selection strategies, while sophisticated approaches included feature-based clustering [33] and genetic algorithms [34]. The different approaches have advantages and disadvantages. For example, univariate feature selection techniques are simple and computationally inexpensive, but ignore the combined effect of multiple features in improving the predictive performance and do not remove highly correlated features. Feature-based clustering, on the other hand, automatically removes highly similar features, but in its most simple form does not consider the features' correlation with the property of interest. Sequential feature selection (SFS) offered a good balance of resulting model performance and computational cost, and for these reasons was chosen for feature selection in this study. This method describes cycling through the available features one at a time to identify the next feature which provides the best performance in the metric of interest – in this case the CV mean squared error (MSE). SFS was effective at finding a minimum set of high-performing features, but required many ML performance evaluations, the total number of which is described by the following equation,

$$N_r = \sum_{i=1}^{N_{f.d.}} N_{f.t.} - i \tag{4}$$

where N_r is the number of ML performance evaluations, $N_{f.d.}$ is the number of features desired, and $N_{f.t.}$ is the total size of the original feature pool. Twenty steps of sequential feature selection were performed for a given data and model configuration as this was typically sufficient to identify the feature set that minimized the CV mean squared error (MSE). *This corresponds to 7,730 performance evaluations and 77,300 ML training runs due to the 10-fold cross validation strategy*. We compare the features selected by SFS to features with high Pearson correlation scores in the following section. We did not use Pearson correlation scores for feature selection in this study because preliminary analyses revealed the model prediction errors to be considerably higher than for SFS.

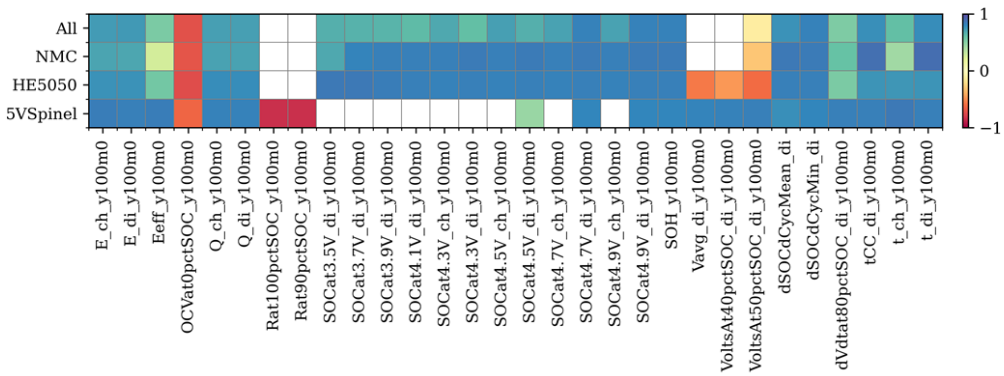
4. Results

4.1 Feature Selection

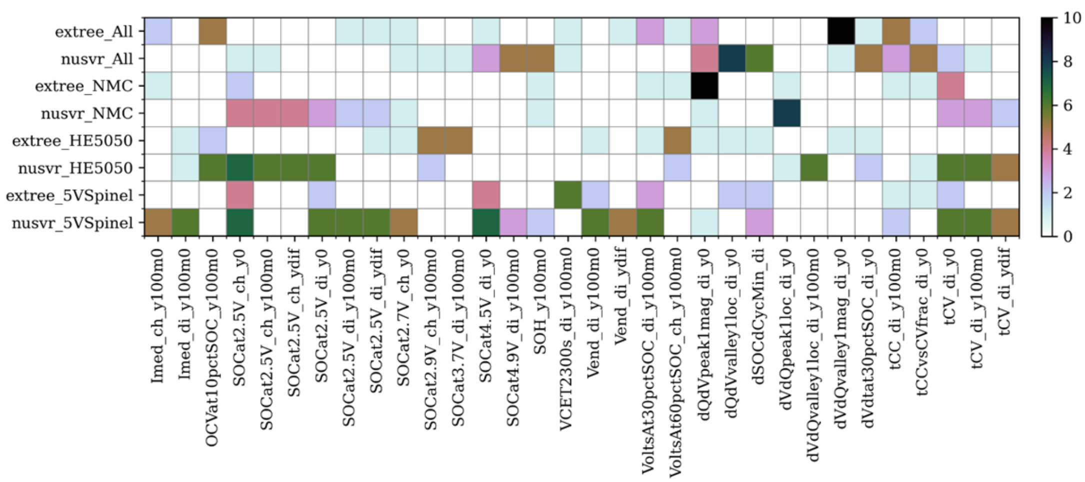
This section presents the results of univariate and multivariate selection approaches which were used to identify the most useful features. Again, the overall electrochemical goal is to devise a cathode-nonspecific algorithm and to understand which features are broadly useful versus useful only for certain cathodes; the data set was designed for this purpose. An important caution is that high performing features may not be physically motivated but can be useful for reflection once feature utility is established. *Thus, it would be ideal to find signals that apply to a range of chemistries and conditions and not just to a small subset.*

While Pearson correlation coefficients (which represent the linear correlation between two datasets in the range of -1 to 1) were not used for feature selection in this work, they can still provide insights about the importance of features and are offered for this analysis. We note that the Pearson correlation coefficient approach may not capture all important features. Figure 3a shows the top 15 features for each set of cathode chemistries ranked by the magnitude of the Pearson correlation coefficient. It spans the entire battery cycling dataset without regard to divisions between training, validation, and test sets since these were found to have no impact on feature selection in the machine learning. With few exceptions, all the highly correlated features have the *y100m0 appendage, indicating that the change in feature values between the first and 100th cycle is most directly correlated with remaining life*. Features codifying SOC as a function of voltage, charge, and discharge times are well represented. The features that were most positively correlated with cycle life *across cathode sets* were the difference in SOC at 4.7V during discharge between cycles 100 and 0 (SOCat4.7V_di_y100m0), the difference in SOC at 4.9V during discharge between cycles 100 and 0 (SOCat4.9V_di_y100m0), the difference in SOH between cycles 100 and 0 (SOH_y100m0), and the minimum of the change in SOC (across all SOC) during discharge between cycles 10 and 100 (dSOCdCycMin_di) features. Interestingly, the dQ/dV and dV/dQ peak and valley features do not show high linear correlation with cycle life. The only highly negatively correlated feature, displayed in warm color, is the difference in OCV at 0% SOC between cycles 100 and 0 (OCVat0pctSOC_y100m0), indicating that a positive change at the low SOC OCV is correlated with low cycle life. An early increase in OCV could be explained by increased accessible capacity as wetting improves (a well-known early cycling phenomena) or to another break-in process. Interestingly, the difference in R at 100% SOC between cycles 100 and 0 (Rat100pctSOC_y100m0) and the difference in R at 90% SOC between cycles 100 and 0 (Rat90pctSOC_y100m0) are highly correlated for the 5VSpinel cathode, but uncorrelated for the other chemistries which have a less significant voltage fade. Additionally, the difference in average voltage during discharge between cycles 100 and 0 (Vavg_di_y100m0), is positively correlated with life for 5V Spinel, negatively correlated for HE5050, yet uncorrelated for NMC and the 'all' cathodes. This matches with the expectation for HE5050 which is well known for voltage fade during cycling - suggesting that the average discharge voltage feature might correlate with cycle life [35]. These examples, among others, indicate that *the usefulness of some features is cathode specific, while other features are correlated with battery life across cathodes.* This result has implications for which features are important in predicting within chemistry, and which features are necessary for accurate broad utility and predictions with new chemistries.

Figure 3b shows the number of times a given feature was present in the 10 test-set CV runs for the four cathode sets. Only features that were present 5 or more times in at least one of the cathode sets are shown for readability purposes. In contrast to Figure 3a, y_0, y_100m0, and y_diff features are all represented in the selected features. Furthermore, dQ/dV and dV/dQ features are well represented, which has long been proposed as a key indicator for voltage and capacity fade [36,37]. Of particular interest is the first valley magnitude of the dV/dQ curve during discharge at cycle 0 (dVdQvalley1mag_di_y0) feature, which is present in all CV sets for Extratrees but *only with all cathodes*. This finding offers a clear positive answer to one central question of whether there is one feature which is uniquely strongly valued and predictive across multiple chemistries. Equally interesting is that this same feature does not perform well in individual cathode sets. Represented in both Figure 3a and Figure 3b are the change in SOC at 3.7V during discharge between cycle 100 and 0 (SOCat3.7V_di_y100m0), the change in SOC at 4.9V during discharge between cycles 100 and 0 (SOCat4.9Vdi_y100m0), the minimum of the change in SOC (across all SOC) during discharge between cycles 10 and 100 (dSOCdCycMin_di), and the change in constant current discharge time between cycles 100 and 0 (tCC_di_y100m0).



a)



b)

Figure 3: a) Pearson correlations are shown for the top fifteen ranked features for each cathode set. Pearson correlations with a p-value exceeding 0.05 (where the probability of any correlation being due to random noise is too high) are shown in white. b) The number of instances of a feature present out of ten CV sets are shown for all model types and cathode chemistry sets. Only features where at least one of the model types had 5 of the 10 or more CV instances of this feature are shown.

4.2 Machine Learning Performance

This section presents the performance of the two machine learning algorithms on different cathode chemistry and feature sets. Conclusions are offered on the overall performance of each algorithm and on when the use of a reduced feature set is preferred over the full feature set, and if so, which feature set this should be.

Figure 4 plots an analysis where different ML models were trained for different cathode sets and restricted feature sets. The restricted feature sets are categorized by a major feature type – for example all dQ/dV related features or all SOC at a given voltage features. The ‘all’ feature set contains all 396 features. Both the mean and standard deviation across the test-set CV runs are offered for the mean absolute errors (MAE). Empty cells indicate feature sets that were empty for a given chemistry due to excessive NaNs in the original features. Figure 4a and Figure 4b show validation MAE means and standard deviations, respectively; Figure 4c and Figure 4d show test MAE means and standard deviations, respectively. Figure 4 note-worthy observations include:

- The validation and test MAE means and standard deviations are quite similar. This is because both validation and test data are drawn from the same pool of batteries (i.e., are from the same statistical distribution of data).
- A variety of features sets and models yield accurate performance on both the validation and test sets, *which suggests that there is no one dominant feature set.* All feature sets, excluding dVdQ, R, dSOCdCyc and dSOHdCyc, provide accurate predictions.
- The mean and standard deviation validation MAE for the ‘All’ cathode set for Extratrees are 65 and 1.6 cycles, respectively, and for NuSVR are 72 and 5.5 cycles, respectively. Based only on the mean error Extratrees is preferred, but the scatter in the performance means that sometimes NuSVR may outperform Extratrees.
- While the use of all features and the Extratrees algorithm results in the lowest validation MAE for the all-cathode category, the dV/dt feature set has a lower test MAE of 75 cycles to the ‘All’ set’s 78 cycles. Also unusual is the case of the Extratrees models for the HE5050 cathode set, where the validation and test errors for the dV/dt feature set are lower than for the ‘All’ feature set. This shows that in some cases, a reduced feature pool may exhibit enhanced performance versus a larger pool of features.

Generally, the inclusive ‘All’ feature set results in the best performance by only a small margin, but if for a particular study only a reduced feature set is available it should not be a large detriment to the ML performance. Furthermore, the Extratrees algorithm exhibits better performance than NuSVR across the validation and test sets, but the improvement is typically small and depends on which cathode and feature sets used. Figure 5a and b provide parity plots for predictive performance on the test set for the Extratrees and NuSVR algorithms, respectively.

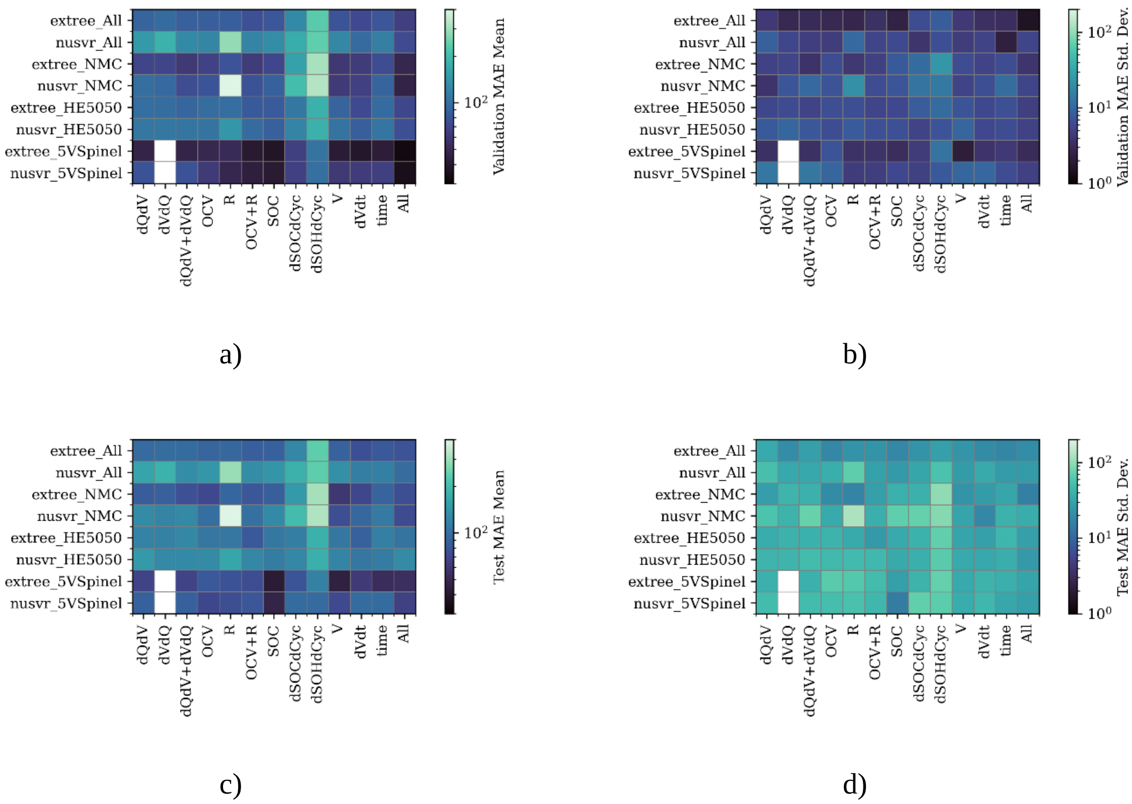
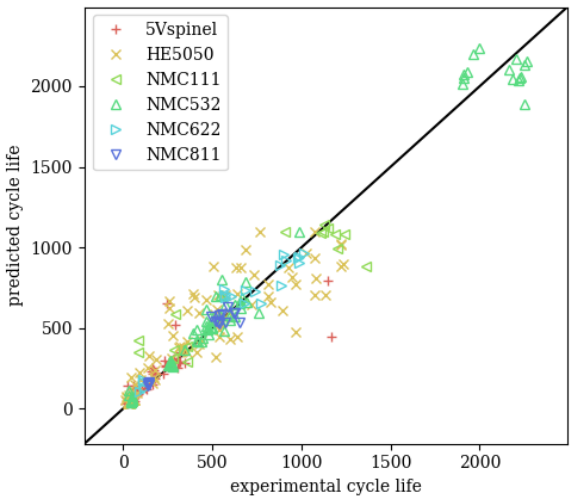
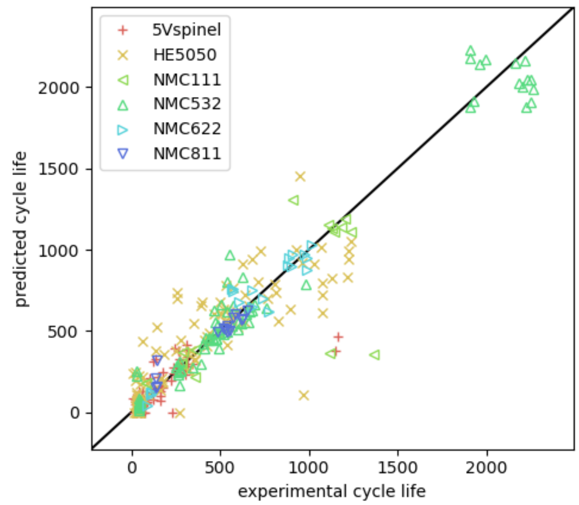


Figure 4: The mean and standard deviations of the validation and test MAE across CV sets are plotted versus feature set (x-axis) and the ML algorithm and cathode set (y-axis) as follows, a) mean validation MAE, b) standard deviation validation MAE, c) mean test MAE, and d) standard deviation test MAE. On the x-axis 'All' refers to the use of all feature sets, and on the y-axis it refers to using all cathode chemistries at the same time.

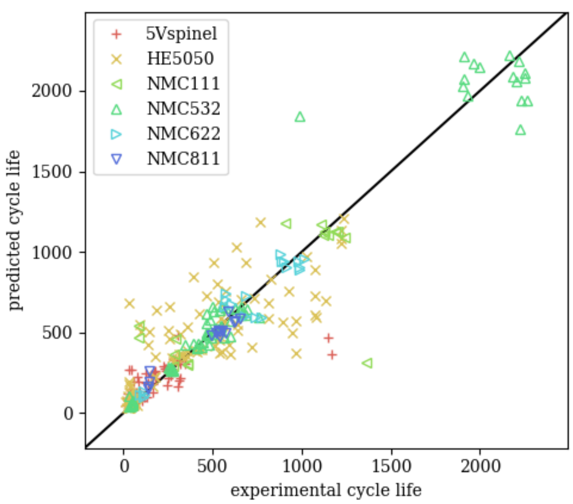
It is also of considerable potential impact to test the predictive accuracy when using only the first cycle to calculate the features. Figure 5c and d plot the predicted versus experimental cycle lives for both ML algorithms using only the first cycle of testing data. The performance is still within 32% of that of the original feature set, with Extratrees and NuSVR exhibiting test MAEs of 103 and 114 cycles, respectively. Accuracy is lower than for the case when all 100 preliminary cycles are available but do provide a valuable life prediction and rapid screening tool. Selected feature types for this rapid screening differed in some regards with respect to Figure 4b, with many important dQ/dV and dV/dQ features, and many important dV/dt features. Performance in predicting from unseen chemistries was also degraded with MAEs of 371 and 333 cycles for Extratrees and NuSVR, respectively.



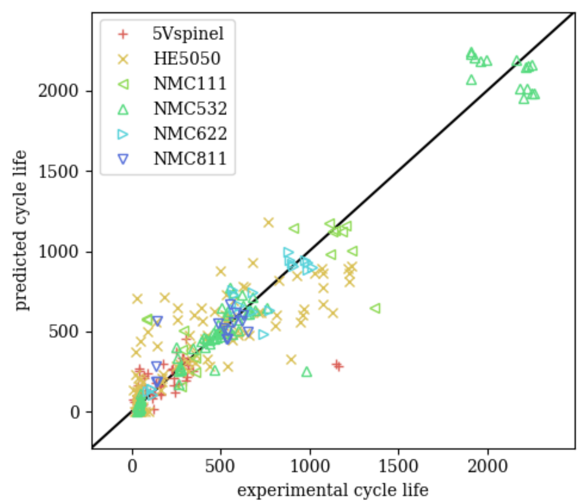
a)



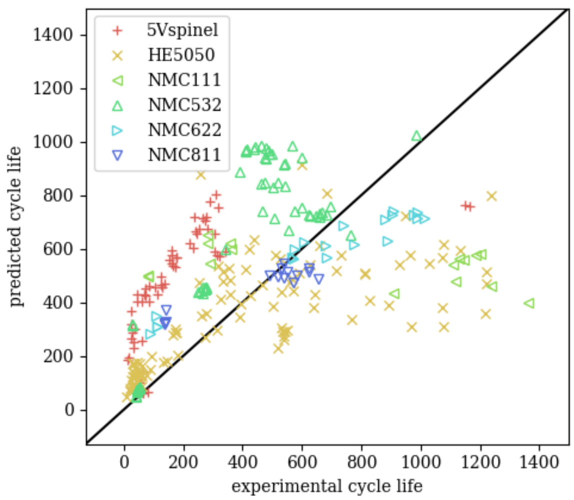
b)



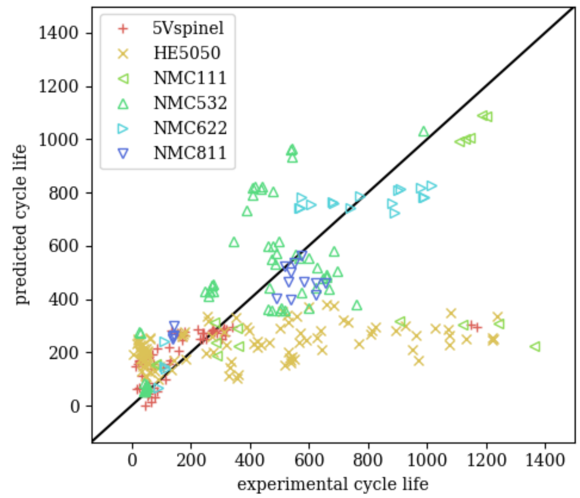
c)



d)



e)



f)

Figure 5: Parity plots comparing experimental and predicted cycle lives in the test set using the first 100 cycles of testing data are shown for a) the Extratrees algorithm, and b) the NuSVR algorithm. Parity plots comparing experimental and predicted cycle lives in the test set using only the first cycle of testing data are shown for c) the Extratrees algorithm, and d) the NuSVR algorithm. Parity plots comparing experimental and predicted cycle lives for unseen cathode chemistries are shown for e) the Extratrees algorithm, and f) the NuSVR algorithm.

Additionally, this study was structured to offer a comparison of ML performance across novel cathode chemistries. *This has tremendous future impact for materials discovery research as it reduces the typical bottleneck of characterization of cycle performance on future chemistries.* Predictive performance on new chemistries was evaluated through an outer loop where each cathode chemistry was swapped into the evaluation set in turn. Table 2a and Table 2b tabulate the predictive performance for each cathode chemistry for the Extratrees and NuSVR algorithms, respectively. The ‘All’ cathode set refers to the MAE across all cells reserved for evaluating performance on unseen chemistries. Note that in the presented results, cells with experimental cycle lives greater that 1500 cycles are excluded, as these are only present for the NMC532 cathode case and therefore these ML algorithms cannot be expected to predict values greater than they have previously seen. Both algorithms demonstrate strong performance on some cathodes and weak performance on others. The Extratrees algorithm shows good performance for the NMC532, NMC622, NMC811, and HE5050 cathodes, while NuSVR has strong performance for NMC532, NMC622, NMC811, and 5Vspinel cathodes. Neither perform well on NMC111, though NuSVR has the edge. The NMC111 cells were produced over a decade and represent diverse cell design and materials differences. Therefore, it is possible that the machine learning algorithms would struggle to predict performance across the 16 NMC111 cells without that diversity mirrored in the training dataset.

Table 2: The a) Extratrees, and b) NuSVR performance on unseen chemistries is tabulated for different cathode sets via mean absolute error (MAE), mean squared error (MSE), root mean squared error (RMSE), and R-squared (R^2). 'All' cathode set refers to all cells reserved for evaluating performance on unseen chemistries. Metrics in which one algorithm out-performs the other are highlighted.

a)

Cathode Set	MAE	MSE	RMSE	R^2
NMC111	507	294003	542	0.163
NMC532	221	82128	287	0.773
NMC622	156	33295	182	0.945
NMC811	98	14977	122	0.879
HE5050	181	67247	259	0.528
5Vspinel	318	116145	341	0.391
All	230	88138	297	0.314

b)

Cathode Set	MAE	MSE	RMSE	R^2
NMC111	287	205726	454	0.396
NMC532	152	39756	199	0.532
NMC622	117	18225	135	0.821
NMC811	104	14772	122	0.749
HE5050	276	134300	366	0.234
5Vspinel	91	31343	177	0.311
All	191	81693	286	0.330

Figure 5a and b plot the experimental versus predicted performance for the Extratrees and NuSVR algorithms, respectively, in out-of-training predictions for unseen cathode chemistries. Extratrees is predictive for the HE5050 cathode set below 500 cycles, but plateaus after this point, while NuSVR predicts a life of around 200 cycles for all cells. Overall, NuSVR has higher predictive accuracy for unseen cathode chemistries.

We also investigate whether training the machine learning algorithms with a broad or narrow range of cathode chemistries increases their predictive ability for new chemistries. To this end, we evaluated out of chemistry predictive performance for a reduced training set consisting of only NMC cathode chemistries. Interestingly, the out of chemistry predictive performance for NMC cathode chemistries decreased considerably for both the Extratrees and NuSVR algorithms. The predictive performance was especially degraded for the NuSVR algorithm, with all MAE scores increasing by over 100 cycles versus the more expansive training set. *This result implies that the presence of HE5050 and 5VSpinel cathodes in the training set improves predictive ability for new NMC cathode chemistries - which is somewhat unexpected.*

This behavior may be partially explained by Figure 3, which shows considerable overlap in selected features between HE5050, 5VSpinel, and NMC cathode chemistry sets, especially for the initial SOC at 2.5V during charge or discharge (SOCat2.5V_ch_y0 and SOCat2.5V_di_y0, respectively), the initial CV discharge time (tCV_di_y0), and the difference in CV discharge time between cycles 100 and 0 (tCV_di_y100m0).

5. Discussion

This study met its objectives by investigating the effect of 396 distinct features on the performance of machine learning algorithms in predicting the cycle lives of 300 expertly made Li-ion batteries with graphite anodes and six different metal oxide cathodes. The remainder of this section discusses the results of the feature selection process and its effect on the predictive performance of two machine learning algorithms for different cathode chemistries both in and outside of the training datasets.

Feature selection:

Sequential feature selection resulted in models with the best performance, but was computationally expensive, requiring 463,800 performance evaluations and 4,638,000 ML training runs to study performance for both similar and dissimilar cathode chemistries to the training data. The reason for this large number of model training runs is the cross-validation procedure employed to evaluate a model's predictive performance across the entire training dataset. The relationship between optimal feature choice, the breadth of chemistry addressed, and the number of cycles used for the life prediction was explored. Certain features performed well across cathode chemistries, for example, the first valley magnitude of the dV/dQ curve during discharge at cycle 0 was used by the Extratrees ML model across all cathodes. In many cases, restricted features sets also yielded strong performance (e.g., equivalent circuit model OCV+R versus SOC features and Voltage versus SOC features) but in others very poor predictive performance was seen (e.g., features based only on the equivalent circuit model resistance versus SOC). The distribution of important feature types was similar between cases where 1 and 100 preliminary cycles were used to define the features. The primary difference was an increase in the number of dQ/dV, dV/dQ, and dV/dt features when only 1 preliminary cycle was used.

Predictive capabilities for cathodes in the training set:

The Extratrees algorithm outperformed NuSVR in cases where the same cathode chemistries were in the training and test sets. Predictive performance was strong across cathode chemistries when 100 cycles of data were used to build features, with an MAE of 78 cycles when all cathodes were in the training set. Predictive performance was strongest when only 5Vspinel was in the training set with an MAE of 54 cycles, and the weakest when only HE5050 was in the training set with an MAE of 103 cycles. Predictive performance was somewhat degraded when only one cycle was used to build features. In this case, an MAE of 103 was achieved when all cathodes were in the training set and the Extratrees algorithm was used. Predictive performance was strongest when only NMC cathodes were in the training set with an MAE of 86 cycles, and the weakest when only HE5050 was in the training set with an MAE of 125 cycles. The success of ML when using only first cycle features points to the combined strength of a comprehensive feature portfolio and effective feature selection techniques towards making accurate predictions.

Predictive capabilities for unseen cathode chemistries:

By segmenting batteries into training, validation, test, and "unseen cathode" sets, a single cathode chemistry was isolated at a time, enabling the evaluation of the models' predictive performance on cathode chemistries not employed to train the models. Predictive performance on the unseen cathode set was strong for some cathode chemistries, but weak for others. The strongest prediction was for 5Vspinel with an MAE of 91, and the weakest was for NMC111 with an MAE of 287, both with the NuSVR algorithm. This is likely to be a result of specialized features being selected for the training data that are not predictive for unseen chemistries. This points towards the importance of identifying features or feature sets that are "universally" predictive. The NuSVR models generally outperformed Extratrees models for the unseen cathode sets. A possible explanation is that the NuSVR algorithm is underfitting the data, yielding stronger predictive performance outside of the domain of the training data (i.e., a new cathode chemistry). Interestingly, predictive performance on unseen NMC cathode chemistries was improved by the presence of HE5050 and 5Vspinel cathodes in the training set. Even though these cathode formulations are dramatically different than NMC, they improve the ML model's ability to learn these "universally" predictive features. This has positive implications for materials discovery research. Predictive performance for the unseen cathode set was not strong when only single-cycle features were used, with MAEs of 371 and 333 cycles for Extratrees and NuSVR, respectively.

6. Conclusions

This study successfully met the objective of devising an algorithm which is effective for a broad range of Li-ion cathode chemistries, using minimal initial cycling data - as little as one cycle - even on chemistries previously unseen during algorithm development. Our approach was to explore effectiveness of various feature selection and ML modeling approaches in a carefully designed dataset, using a powerful computational facility. These results will impact not only deployment needs but also materials discovery research to provide the next generation of solutions.

A database of 300 expertly made Li-ion batteries with six different cathode chemistries drawn from a rich pool at a US national lab was used. Approximately 400 expert-informed quantities were developed to featurize the ML input space; 7,730 performance evaluations and 77,300 ML training runs were used to optimize. The sequential feature selection approach enabled powerful predictive performance using two ML algorithms: ExtraTrees and NuSVR - interpolative performance was strong. Optimized algorithm/features combinations offer excellent predictive quality across a broad range of cathode chemistries, with as little as one initial cycle. Specifically: (1) A model trained on all cathode chemistries with 100 initial cycles offers predictions with a mean absolute error (MAE) of 78 cycles (2) An MAE of 103 cycles was seen on the same dataset when features were restricted to only the first cycle. (3) Performance on cathode chemistries unseen by the ML algorithms during the training phase was also demonstrated, with MAEs of 117, 104, and 91 cycles seen for NMC622, NMC811, and 5Vspinel datasets. Overall, the ExtraTrees algorithm was preferred when predicting for cathode chemistries in the training set, while NuSVR was preferred for predicting for unseen cathode chemistries (i.e., cathodes not in the training set).

AI/ML for Li-ion life prediction is an emerging field of study with great potential impact that has been shackled by the lack of abundant data. The success of this challenging study was made possible by a data set that, although small in scope, was sufficient to support the objectives, especially in combination with significant computational resources. The developed modeling approach is promising in accelerating the development of new battery chemistries and designs. In future work, it would be valuable to compare the results of this approach to those of a significantly larger and more diverse dataset, and to extend to different temperatures and use profiles.

Acknowledgements

The authors acknowledge financial support from Laboratory Directed Research and Development (LDRD) funding from Argonne National Laboratory, provided by the Director, Office of Science, of the U.S. Department of Energy under Contract No. DE-AC02-06CH11357. We also gratefully acknowledge the computing resources provided by the Laboratory Computing Resource Center on the Bebop supercomputer at Argonne National Laboratory. The electrodes, pouch cells, and 18650 cells were produced at the U.S. Department of Energy’s (DOE) CAMP (Cell Analysis, Modeling and Prototyping) Facility, Argonne National Laboratory. The CAMP Facility is fully supported by the DOE Vehicle Technologies Office (VTO).

Data Availability

The raw data required to reproduce these findings will be made available upon request.

References

[1] A. El Mejdoubi, H. Chaoui, H. Gualous, P. Van Den Bossche, N. Omar, J. Van Mierlo, Lithium-ion batteries health prognosis considering aging conditions, IEEE Trans. Power Electron. 34 (2019). https://doi.org/10.1109/TPEL.2018.2873247.

[2] D.D. Susilo, A. Widodo, T. Prahasto, M. Nizam, State of health estimation of lithium-ion batteries based on combination of gaussian distribution data and least squares support vector machines regression, in: Mater. Sci. Forum, 2018. https://doi.org/10.4028/www.scientific.net/MSF.929.93.

[3] B. Saha, K. Goebel, S. Poll, J. Christophersen, Prognostics methods for battery health monitoring using a Bayesian framework, IEEE Trans. Instrum. Meas. 58 (2009). https://doi.org/10.1109/TIM.2008.2005965.

[4] X. Zhong, L. Liu, Y. Jiang, X. Wang, L. Wang, G. Zhuang, X. Li, D. Mei, J.G. Wang, D.S. Su, Synergistic Effect of Nitrogen in Cobalt Nitride and Nitrogen-Doped Hollow Carbon Spheres for the Oxygen Reduction Reaction, ChemCatChem. 7 (2015) 1826–1832. https://doi.org/10.1002/cctc.201500195.

[5] I. Bloom, B.W. Cole, J.J. Sohn, S.A. Jones, E.G. Polzin, V.S. Battaglia, G.L. Henriksen, C. Motloch, R. Richardson, T. Unkelhaeuser, D. Ingersoll, H.L. Case, An accelerated calendar and cycle life study of Li-ion cells, J. Power Sources. 101 (2001) 238–247. https://doi.org/https://doi.org/10.1016/S0378-7753(01)00783-2.

[6] K. Jalkanen, J. Karppinen, L. Skogström, T. Laurila, M. Nisula, K. Vuorilehto, Cycle aging of commercial NMC/graphite pouch cells at different temperatures, Appl. Energy. 154 (2015) 160–172. https://doi.org/https://doi.org/10.1016/j.apenergy.2015.04.110.

[7] D. Liu, Y. Bai, S. Zhao, W. Zhang, Improved cycling performance of 5 v spinel LiMn 1.5Ni 0.5O 4 by amorphous FePO 4 coating, J. Power Sources. 219 (2012).

https://doi.org/10.1016/j.jpowsour.2012.07.058.

[8] K.A. Severson, P.M. Attia, N. Jin, N. Perkins, B. Jiang, Z. Yang, M.H. Chen, M. Aykol, P.K. Herring, D. Fraggedakis, M.Z. Bazant, S.J. Harris, W.C. Chueh, R.D. Braatz, Data-driven prediction of battery cycle life before capacity degradation, Nat. Energy 2019 45. 4 (2019) 383–391. https://doi.org/10.1038/s41560-019-0356-8.

[9] A. Aitio, S.G. Marquis, P. Ascencio, D. Howey, Bayesian parameter estimation applied to the Li-ion battery single particle model with electrolyte dynamics, in: IFAC-PapersOnLine, 2020. https://doi.org/10.1016/j.ifacol.2020.12.1770.

[10] P. Fermín-Cueto, E. McTurk, M. Allerhand, E. Medina-Lopez, M.F. Anjos, J. Sylvester, G. dos Reis, Identification and machine learning prediction of knee-point and knee-onset in capacity degradation curves of lithium-ion cells, Energy AI. 1 (2020). https://doi.org/10.1016/j.egyai.2020.100006.

[11] M. Dubarry, D. Beck, Analysis of synthetic voltage vs. Capacity datasets for big data li-ion diagnosis and prognosis, Energies. 14 (2021). https://doi.org/10.3390/en14092371.

[12] P. Gasper, K. Gering, E. Dufek, K. Smith, Challenging Practices of Algebraic Battery Life Models through Statistical Validation and Model Identification via Machine-Learning, J. Electrochem. Soc. 168 (2021) 20502.

[13] W. Li, N. Sengupta, P. Dechent, D. Howey, A. Annaswamy, D.U. Sauer, One-shot battery degradation trajectory prediction with deep learning, J. Power Sources. 506 (2021) 230024. https://doi.org/https://doi.org/10.1016/j.jpowsour.2021.230024.

[14] S. Greenbank, D. Howey, Automated feature extraction and selection for data-driven models of rapid battery capacity fade and end of life, IEEE Trans. Ind. Informatics. (2021) 1. https://doi.org/10.1109/TII.2021.3106593.

[15] V. Sulzer, P. Mohtat, S. Lee, J.B. Siegel, A.G. Stefanopoulou, Promise and challenges of a data-driven approach for battery lifetime prognostics, ArXiv Prepr. ArXiv2010.07460. (2020).

[16] Y. Ma, L. Wu, Y. Guan, Z. Peng, The capacity estimation and cycle life prediction of lithium-ion batteries using a new broad extreme learning machine approach, J. Power Sources. 476 (2020) 228581. https://doi.org/https://doi.org/10.1016/j.jpowsour.2020.228581.

[17] S. Shen, V. Nemani, J. Liu, C. Hu, Z. Wang, A Hybrid Machine Learning Model for Battery Cycle Life Prediction with Early Cycle Data, in: 2020 IEEE Transp. Electrif. Conf. Expo, 2020: pp. 181–184. https://doi.org/10.1109/ITEC48692.2020.9161647.

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

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

[20] L. Ward, S. Babinec, E.J. Dufek, D.A. Howey, V. Viswanathan, M. Aykol, D.A.C. Beck, B. Blaiszik, B.-R. Chen, G. Crabtree, others, Principles of the Battery Data Genome, ArXiv Prepr. ArXiv2109.07278. (2021).

[21] M.C. Smart, C. Hwang, F.C. Krause, J. Soler, W.C. West, B. V. Ratnakumar, K. Amine, Wide Operating Temperature Range Electrolytes for High Voltage and High Specific Energy Li-Ion Cells, ECS Trans. 50 (2013). https://doi.org/10.1149/05026.0355ecst.

[22] D. Mohanty, A.S. Sefat, J. Li, R.A. Meisner, A.J. Rondinone, E.A. Payzant, D.P. Abraham, D.L. Wood, C. Daniel, Correlating cation ordering and voltage fade in a lithium-manganese-rich lithium-ion battery cathode oxide: A joint magnetic susceptibility and TEM study, Phys. Chem. Chem. Phys. 15 (2013). https://doi.org/10.1039/c3cp53658k.

[23] C.R. Birkl, M.R. Roberts, E. McTurk, P.G. Bruce, D.A. Howey, Degradation diagnostics for lithium ion cells, J. Power Sources. 341 (2017) 373–386. https://doi.org/10.1016/j.jpowsour.2016.12.011.

[24] B.-R. Chen, M.R. Kunz, T.R. Tanim, E.J. Dufek, A machine learning framework for early detection of lithium plating combining multiple physics-based electrochemical signatures, Cell Reports Phys. Sci. 2 (2021) 100352. https://doi.org/https://doi.org/10.1016/j.xcrp.2021.100352.

[25] J. Vetter, P. Novák, M.R. Wagner, C. Veit, K.-C. Möller, J.O. Besenhard, M. Winter, M. Wohlfahrt-Mehrens, C. Vogler, A. Hammouche, Ageing mechanisms in lithium-ion batteries, J. Power Sources. 147 (2005) 269–281. https://doi.org/https://doi.org/10.1016/j.jpowsour.2005.01.006.

[26] V. Agubra, J. Fergus, Lithium Ion Battery Anode Aging Mechanisms, Materials (Basel). 6 (2013) 1310–1325. https://doi.org/10.3390/ma6041310.

[27] S.-K. Jung, H. Gwon, J. Hong, K.-Y. Park, D.-H. Seo, H. Kim, J. Hyun, W. Yang, K. Kang, Understanding the Degradation Mechanisms of LiNi0.5Co0.2Mn0.3O2 Cathode Material in Lithium Ion Batteries, Adv. Energy Mater. 4 (2014) 1300787. https://doi.org/https://doi.org/10.1002/aenm.201300787.

[28] H. Pan, Z. Lü, H. Wang, H. Wei, L. Chen, Novel battery state-of-health online estimation method using multiple health indicators and an extreme learning machine, Energy. 160 (2018). https://doi.org/10.1016/j.energy.2018.06.220.

[29] L. Ward, J. Kubal, B. Blaiszik, Battery Data Toolkit, (2021). https://github.com/materials-data-facility/battery-data-toolkit.

[30] P. Geurts, D. Ernst, L. Wehenkel, Extremely randomized trees, Mach. Learn. 63 (2006). https://doi.org/10.1007/s10994-006-6226-1.

[31] J. Platt, others, Probabilistic outputs for support vector machines and comparisons to regularized likelihood methods, Adv. Large Margin Classif. 10 (1999).

[32] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, others, Scikit-learn: Machine learning in Python, J. Mach. Learn. Res. 12 (2011) 2825–2830.

[33] Q. Song, J. Ni, G. Wang, A fast clustering-based feature subset selection algorithm for high-dimensional data, IEEE Trans. Knowl. Data Eng. 25 (2013). https://doi.org/10.1109/TKDE.2011.181.

[34] N.H. Paulson, B. Gould, S.J. Wolff, M. Stan, A.C. Greco, Correlations between thermal history and keyhole porosity in laser powder bed fusion, Addit. Manuf. 34 (2020). https://doi.org/10.1016/j.addma.2020.101213.

[35] S.E. Trask, Y. Li, J.J. Kubal, M. Bettge, B.J. Polzin, Y. Zhu, A.N. Jansen, D.P. Abraham, From coin cells to 400 mAh pouch cells: Enhancing performance of high-capacity lithium-ion cells via modifications in electrode constitution and fabrication, J. Power Sources. 259 (2014) 233–244. https://doi.org/10.1016/j.jpowsour.2014.02.077.

[36] H. Kuhlenbeck, S. Shaikhutdinov, H.J. Freund, Well-ordered transition metal oxide layers in model catalysis - A series of case studies, Chem. Rev. 113 (2013) 3986–4034. https://doi.org/10.1021/cr300312n.

[37] I. Bloom, A.N. Jansen, D.P. Abraham, J. Knuth, S.A. Jones, V.S. Battaglia, G.L. Henriksen, Differential voltage analyses of high-power, lithium-ion cells: 1. Technique and application, J. Power Sources. 139 (2005) 295–303. https://doi.org/10.1016/J.JPOWSOUR.2004.07.021.

Supplementary Information

S1. Battery component properties

Table 3: Summary of anode, cathode, and electrolyte properties.

Cell						Anode					Cathode				
Group	Anode	Cathode	Electrolyte	Electrolyte additive	Target capacity (Ah)	Supplier	Product	Thickness (µm)	Loading (mg/cm²)	Porosity (%)	Supplier	Product	Thickness (µm)	Loading (mg/cm²)	Porosity (%)
1	graphite	5Vspinel	Gen 2	NONE	0.3	Conoco-Phillips	A12	86	5.8	35	NEI	SP-10	73	14.9	43
2	graphite	5Vspinel	Gen 2	NONE	0.3	Conoco-Phillips	A12	86	5.8	35	NEI	SP-10	73	14.9	43
3	graphite	5Vspinel	Gen 2	NONE	0.3	Conoco-Phillips	A12	91	6.4	35	NEI	SP-10	59	14.3	34
4	graphite	5Vspinel	Gen 2	NONE	0.3	Conoco-Phillips	A12	59	12.5	32	NEI	SP-10	62	14.8	35
5	graphite	5Vspinel	Gen 2	NONE	0.4	Gelon	G15, MCMB	62	15.1	33	NEI	SP-10	46	6.4	35
6	graphite	5Vspinel	Gen 2	LiDFOB, LiBOB	0.4	Gelon	G15, MCMB	62	15.1	33	NEI	SP-10	46	6.4	35
7	graphite	5Vspinel	Gen 2	LiDFOB, LiBOB	0.4	Gelon	G15, MCMB	62	15.1	33	NEI	SP-10	46	6.4	35
8	graphite	5Vspinel	Gen 2	LiDFOB, LiBOB	0.4	Gelon	G15, MCMB	62	15.1	33	NEI	SP-10	46	6.4	35
9	graphite	5Vspinel	Gen 2	X	0.4	Gelon	G15, MCMB	62	15.1	33	NEI	SP-10	46	6.4	35
10	graphite	HE5050	Gen 2	NONE	0.4	Conoco-Phillips	A12	86	5.8	35	Toda	TodaHE5050	68	14.5	42
11	graphite	HE5050	Gen 2	NONE	0.3	Conoco-Phillips	A12	86	5.8	35	Toda	Toda	63	14.3	38
12	graphite	HE5050	Gen 2	X	0.3	Conoco-Phillips	A12	86	5.8	35	Toda	Toda	63	14.3	38
13	graphite	HE5050	Gen 2	NONE	0.4	Conoco-Phillips	A12	161	11.2	36	In-house	#N/A	117	12.1	44
14	graphite	HE5050	Gen 2	LiDFOB	0.3	Conoco-Phillips	A12	86	5.8	35	Toda	TodaHE5050	63	14.3	38
15	graphite	HE5050	Gen 2	LiBOB	0.3	Conoco-Phillips	A12	86	5.8	35	Toda	TodaHE5050	63	14.3	38
16	graphite	HE5050	Gen 2	LiDFOB, LiBOB	0.3	Conoco-Phillips	A12	86	5.8	35	Toda	TodaHE5050	63	14.3	38
17	graphite	HE5050	Gen 2	NONE	0.3	Conoco-Phillips	A12	86	5.8	35	Toda	TodaHE5050	63	14.3	38
18	graphite	HE5050	Gen 2	NONE	0.3	Conoco-Phillips	A12	86	5.8	35	Toda	TodaHE5050	63	14.3	38
19	graphite	HE5050	Gen 2	NONE	0.4	Conoco-Phillips	A12	161	11.2	36	Merf	#N/A	121	13	43
20	graphite	HE5050	Gen 2	NONE	0.3	Conoco-Phillips	A12	86	5.8	35	Toda	TodaHE5050	63	14.3	38
21	graphite	HE5050	Gen 2	LiDFOB	0.3	Conoco-Phillips	A12	86	5.8	35	Toda	TodaHE5050	63	14.3	38
22	graphite	HE5050	Gen 2	LiDFOB, LiBOB	0.3	Conoco-Phillips	A12	86	5.8	35	Toda	TodaHE5050	63	14.3	38
23	graphite	HE5050	Gen 2	LiDFOB, LiBOB	0.3	Conoco-Phillips	A12	86	5.8	35	Toda	TodaHE5050	63	14.3	38
24	graphite	NMC111	Gen 2	NONE	0.3	Conoco-Phillips	A12	100	7.3	33	BASF	BASF111	100	13.9	33
25	graphite	NMC111	Gen 2	NONE	0.3	Conoco-Phillips	A12	109	7.4	33	BASF	BASF111	121	14.1	33
26	graphite	NMC111	Gen 2	NONE	0.3	Conoco-Phillips	A12	109	7.4	33	BASF	BASF111	121	14.1	33
27	graphite	NMC111	Gen 2	NONE	0.3	Conoco-Phillips	A12	100	7.3	33	BASF	BASF111	67	15.2	29
28	graphite	NMC111	Gen 2	NONE	0.1	Conoco-Phillips	A12	20.5	3.0	32	Toda	TodaNMC111	25	6	35
29	graphite	NMC111	Gen 2	NONE	0.4	Conoco-Phillips	A12	105	15.6	32	Toda	Toda111	51	4	#N/A
30	graphite	NMC111	Gen 2	NONE	0.4	Conoco-Phillips	A12	105	15.6	32	Toda	Toda111	#N/A	#N/A	#N/A
31	graphite	NMC523	Gen 2	NONE	0.3	Conoco-Phillips	A12	100	6.7	38	Toda	Toda532	101	12.7	38
32	graphite	NMC523	Gen 2	NONE	0.3	Conoco-Phillips	A12	100	6.7	38	Toda	Toda532	101	12.7	38
33	graphite	NMC523	Gen 2	FEC	0.3	Hitachi	MagE3	60	8.9	32	Toda	Toda532	101	12.7	38
34	graphite	NMC523	Gen 2	NONE	1	Conoco-Phillips	A12	76	11.3	31	Toda	Toda532	101	12.7	38
35	graphite	NMC622	Gen 2	NONE	0.4	Hitachi	MagE3	105	15.6	32	ECOPRO	NMC622	78	20.69	32
36	graphite	NMC811	Gen 2	NONE	0.4	Hitachi	MagE3	105	15.6	32	Toda	#N/A	#N/A	#N/A	#N/A
37	graphite	NMC811	Gen 2	NONE	0.1	Conoco-Phillips	A12	43	6.1	35	Targray	NMC811	37	9.6	37

Definition of terms:

- LiDFOB: Lithium difluoro (oxalate) borateLiBOB: Lithium bis(33xalate)borate
- FEC: Fluoroethylene Carbonate
- Electrolytes with an X correspond to an electrolyte that the supplier of the data wanted to keep confidential
- #N/A are where the information was either not gathered or the spread was too large to be meaningful

S2. Feature Calculation

In this section we describe the calculation of the non-trivial or non-obvious physics-inspired features employed in this work.

dVdtStart_<state>, dVdtEnd_<state>, dVdtat<SOC>pctSOC_<state> features:

The dV/dt derivative quantity was simply calculated as the difference in voltage divided by the time difference between adjacent timesteps. Figure 6 illustrates these quantities in addition to the related voltage features.

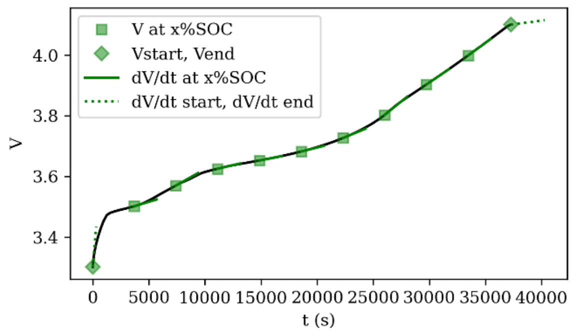
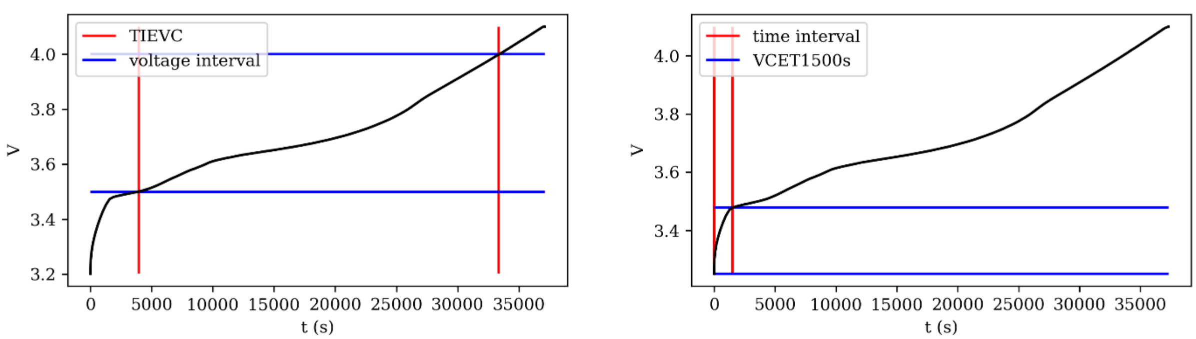


Figure 6: Vstart_ch, Vend_ch, Vat<SOC>pctSOC_ch, dVdtStart_ch, dVdtEnd_ch, and dVdtat<SOC>pctSOC_ch features are visualized for a voltage versus time curve.

TIEVC and VCET features

The calculation of TIEVC and VCET features is illustrated in Figure 7.



a) b)

Figure 7: a) TIEVC and b) VCET features are plotted on a voltage versus time curve for an example cycle.

OCVat<SOC>pctSOH and Rat<SOC>pctSOC features

The OCV and resistance features were calculated using a simple, single resistor equivalent circuit model with SOC dependent resistance. During fitting, 25 SOC levels with higher density in low and high SOC ranges were employed for the OCV and R levels. The ECM fitting residual is given as follows,

$$res_i = V_i - \left(OCV_{lin.}(SOC_i) + I_i * R_{lin.}(SOC_i)\right), \tag{S1}$$

Where i is the timestep index and $OCV_{lin.}$ and $R_{lin.}$ are functions that return the linear interpolation with SOC between the fitted OCV and R values. The scipy Powell minimization scheme was employed in ECM fitting with a guess for the OCV and R parameters set to the fitted values from the previous cycle. During fitting, an empirically developed regularization scheme damped oscillations in the fitted OCV and R curves. The total cost function and regularization are given as follows,

$$Cost = \frac{1}{N}\sum_{i}^{N} {res_i}^2 + reg \tag{S2}$$

$$reg = \frac{1}{P-1}\sum_{j}^{P-1}\left(OCV_{j+1} - OCV_j\right)^2 + \frac{0.02}{P-1}\sum_{j}^{P-1}\left(R_{j+1} - R_j\right)^2 \tag{S3}$$

Where j indexes the OCV and R parameters and the scaling factor of 0.02 was empirically determined. After fitting, the resistance values were normalized by the battery nominal capacity. Linear interpolation was then used to estimate OCV and R at the desired SOC levels (every 10%). Figure 8 shows the quality of fit and the raw OCV and R parameters for an example NMC532 battery.

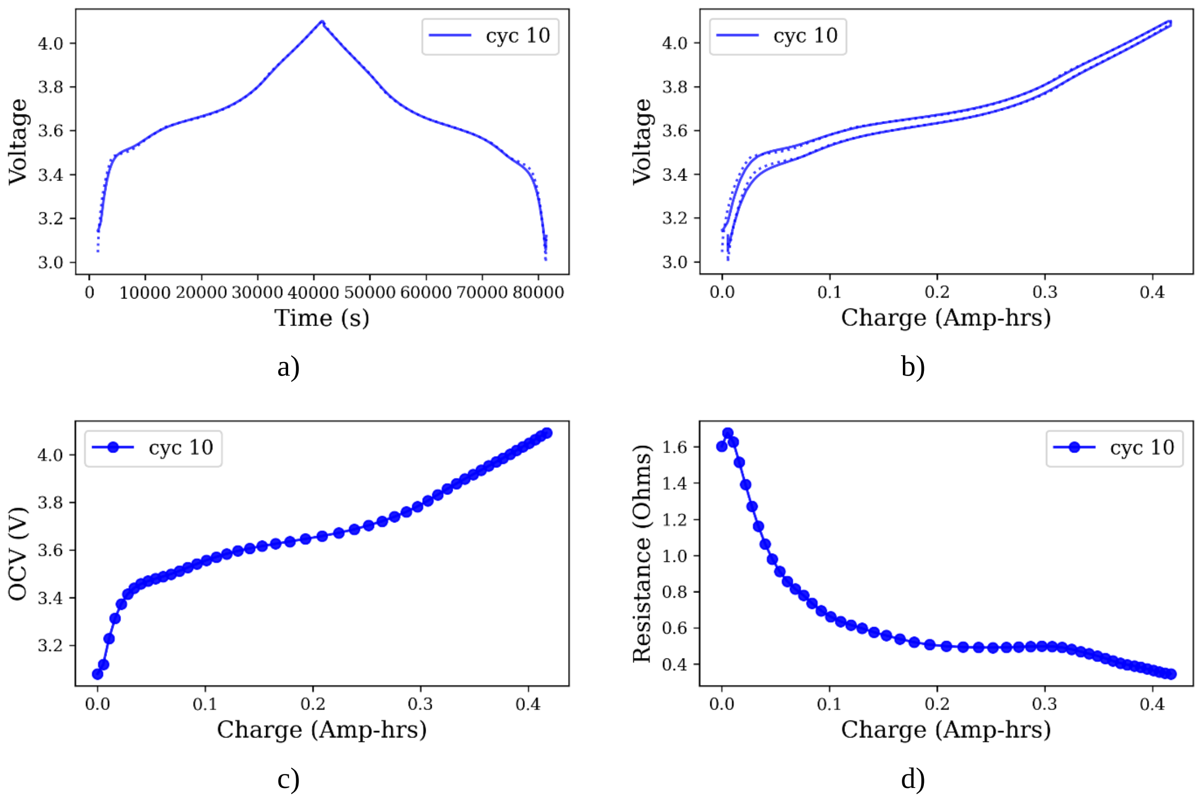


Figure 8: Measured and ECM predicted voltage is plotted versus a) time, and b) charge for an example cycle, and the raw c) OCV, and d) R ECM parameters are plotted versus charge.

dQ/dV and dV/dQ features:

First, the dQ/dV and dV/dQ curves were calculated from constant current segments during charge or discharge. In the case of calculating dQ/dV, both Q and V were sorted by V, duplicate V points were removed, and dV points close to zero were removed. In the case of dV/dQ, only dQ points close to zero needed to be removed. In both cases, quadratic spline interpolation was employed to smooth the curves. The scipy find_peaks algorithm was employed to identify the locations and values of all peaks (valleys) with a prominence that exceeded 5% of the maximum (minimum) peak (valley) height (depth). In the case of dQ/dV the peak areas were also calculated by fitting gaussian-type curves to the identified peaks and computing the area under the curve. A maximum of two peaks or valleys were computed for both dQ/dV and dV/dQ. If fewer than two peaks or valleys were computed, the missing features were assigned NaN values. These quantities are plotted in Figure 9 versus the dQ/dV and dV/dQ curves for an example cycle from an NMC532 battery.

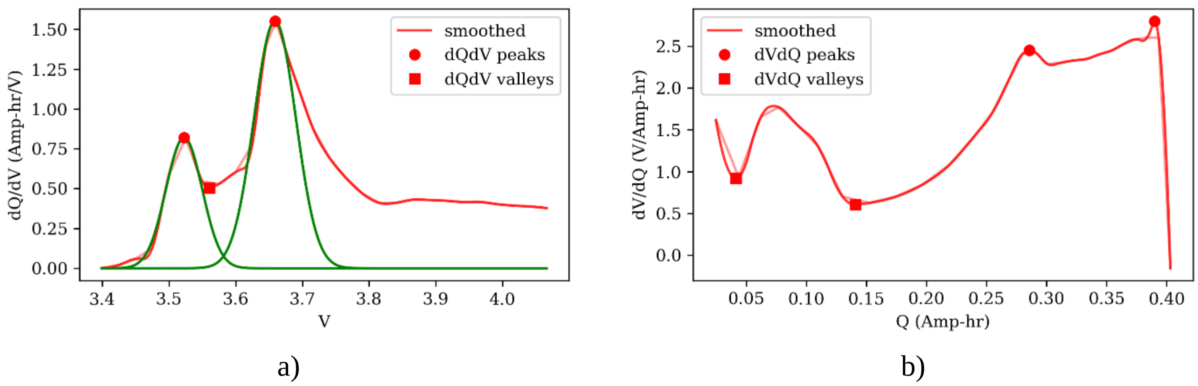


Figure 9: a) dQ/dV is plotted versus voltage alongside the fitted peaks, valleys, and areas for an example cycle. b) dV/dQ is plotted versus charge alongside the fitted peaks and valleys for an example cycle.

dSOHdCycCyc<cycle>

In calculating the dSOHdCycCyc<cycle> feature, first 1000 SOH points were interpolated between the first and hundredth cycle. The derivative of the SOH with respect to changing cycles was then calculated for adjacent steps in this interpolated domain. The final features at the first and last cycle were calculated as the median of the dSOH/dCyc quantities for the first and last 10% of the derivative points, respectively.