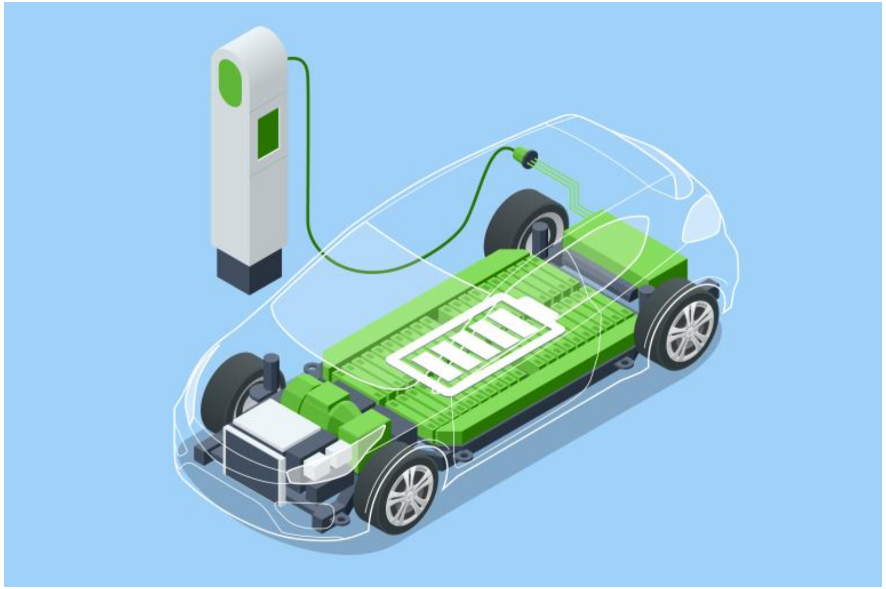
Predicting Li Battery Lifecycle using Charging-Discharging Characteristics

Barath Tirumuruhan
Ritoban Ghosh
Ashutosh Nehete



Background

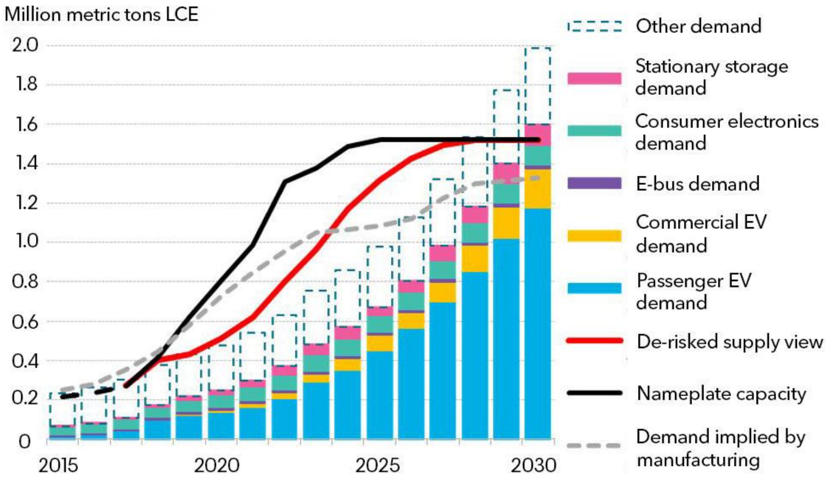
- **4x** increase in the lithium-ion battery demand over the next decade

- For the automotive sector, ability to estimate the battery life-cycle is critical.

- Why?
  - Quality testing
  - Performance prediction
  - Vehicle insurance and resale value
  - Battery second life application

- Can’t we experimentally measure it?
  - Simulating real-time driving: time/cost heavy
  - Materials/chemistry challenge

Figure 1: Global lithium supply and demand forecast, comparing methodologies



Source: BloombergNEF, Avicenne.

Test Duration Estimate	
Chemistry	LFP/Graphite
Capacity	1100 mAh
Charging-Discharging Rate	1C
Time/Cycle	2 hours
Time/2000 cycles	167 days

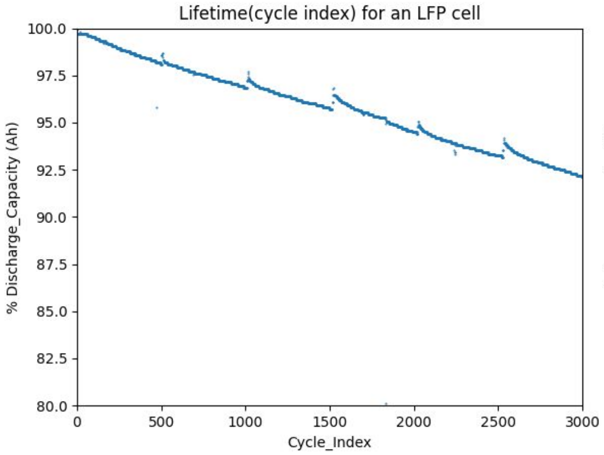
Dataset

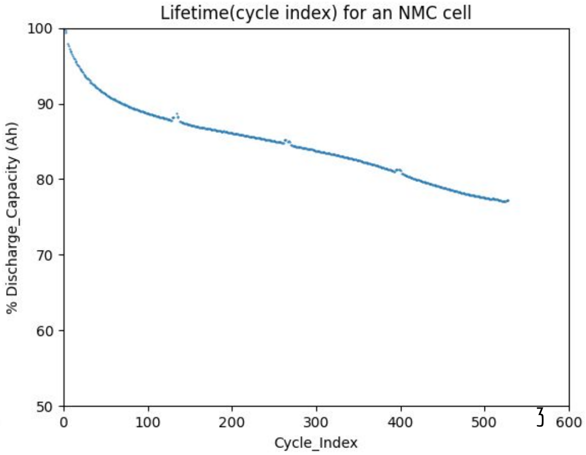
- Sandia National Labs (**SNL**): Commercial batteries degradation studies

- Lifecycle as a function of:
  - Temperature
  - Depth of discharge
  - Discharge Current

- **86** individual cells
  - Time-series data
  - Summary cycle data

Cathode	**LFP**	**NCA**	**NMC**
Chemistry	Li, Fe, P	Li, Ni, Co, Al	Li, Ni, Mn, Co
Manufacturer	A123 Systems	Panasonic	LG Chem
Nominal Capacity (Ah)	1.1	3.2	3
Nominal Voltage (V)	3.3	3.6	3.6
Voltage Range (V)	2 to 3.6	2.5 to 4.2	2 to 4.2
Max Discharge Current (A)	30	6	20
Acceptable Temperature (C)	-30 to 60	0 to 45	-5 to 50
Rated Lifecycle	**2000**	**1000**	**1000**





Data Preprocessing

- **Existing Experimental Features**
    - Battery Rated Capacity (Ah)
    - Temperature
    - Maximum and Minimum SOC
    - Charge rate and Discharge Rate

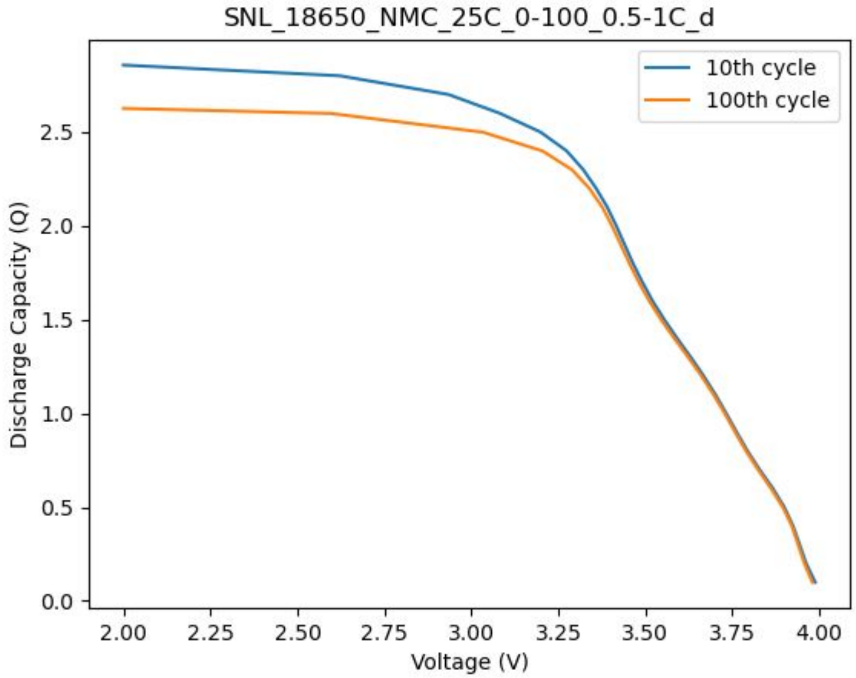
- **Creation of new Features**

$\Delta Q_{10\text{-}100}$ (V)

- Minimum, mean, variance, skewness, kurtosis

Q vs. Cycle

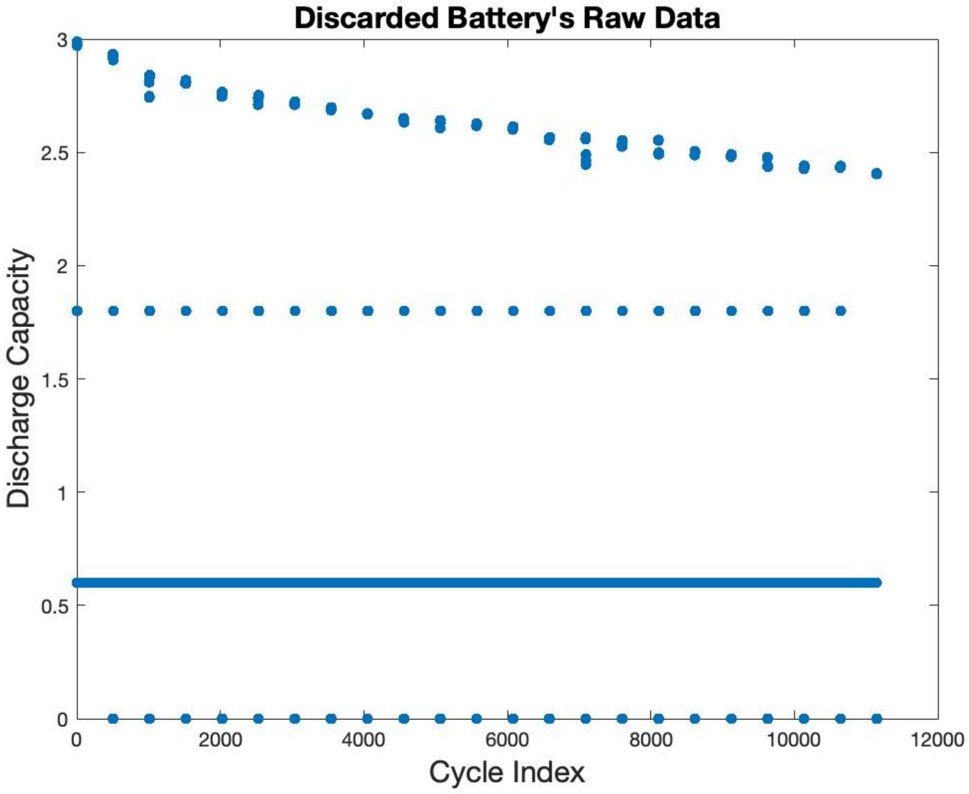
- Two linear fits - cycles (2-100) and (91-100)
- Slopes and Intercepts



Data Preprocessing

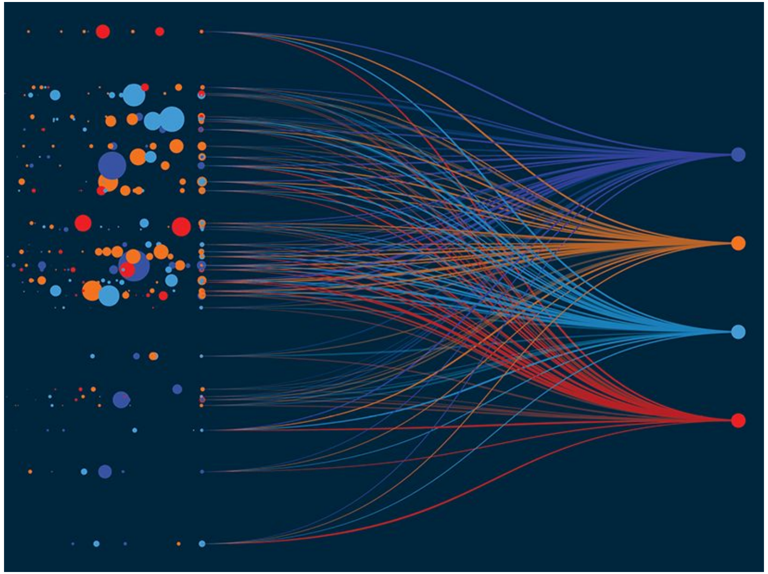
Data Cleaning

- Removal of bad data
  - Discharge capacity is constant
- Outlier removal not effectively implemented.
  - Source of Outliers: Intermittent Tests Performed on the Battery



Methodology Summary

- Chemistry agnostic model
- Feature matrix scaling using **StandardScaler**
- Train : Test = **80 : 20**
- **Stratified** test data split according to cathode chemistry
- Hyperparameter tuning using **GridSearch**
- 5 fold cross-validation using best parameter settings
- Scoring:
    - Regression - RMSE and Absolute error %
    - Classifier - F1



Linear Regression Model

Model	% Error (Training)	% Error (Validation)
Linear L1 LASSO	64.2 +/- 12.3	74.5 +/- 41.9

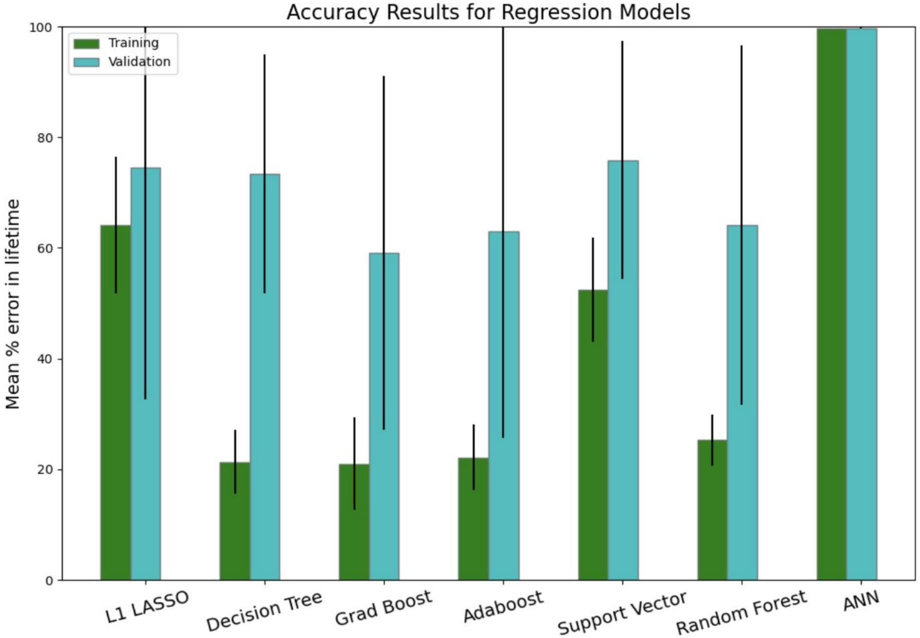
- Linear L1 LASSO model
- Optimised ‘α’ using Grid Search
- Poor prediction and accuracy
    - Failed to capture the non-linear data
    - Nonlinearities in Battery life degradation characteristics

$$J(w) = \sum_{i=1}^{n}(y_i - \hat{y_i})^2 + \alpha \sum_{j=1}^{m}|w_j|$$

Non-linear Regression Models

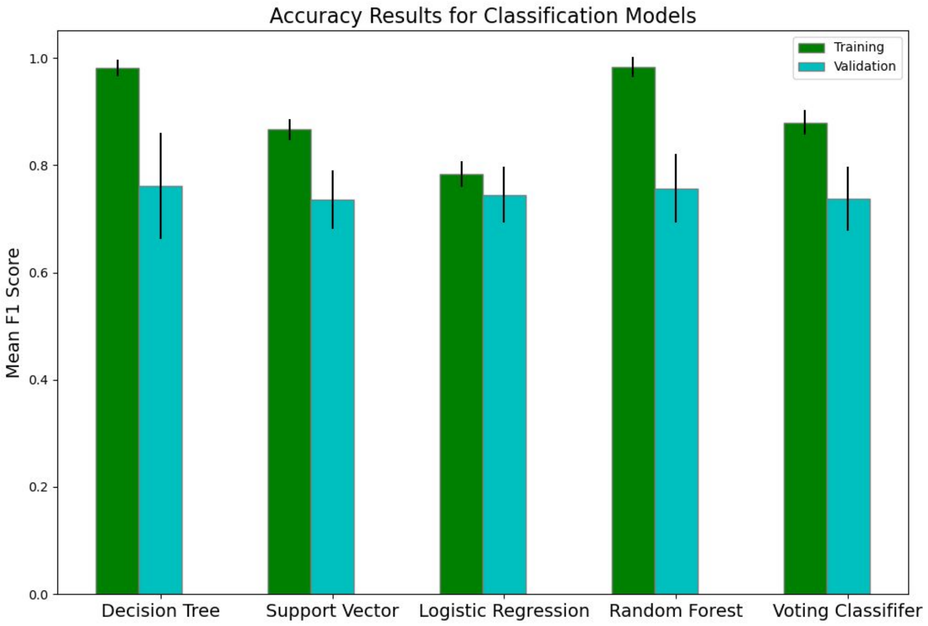
Which models and why?

- Decision Tree & Support Vector Regression perform well with small datasets
- Possible overfitting
- Implemented ensembling methods based on Decision Tree Models
    - No major improvement observed
- ANNs are universal approximators but failed to generalize



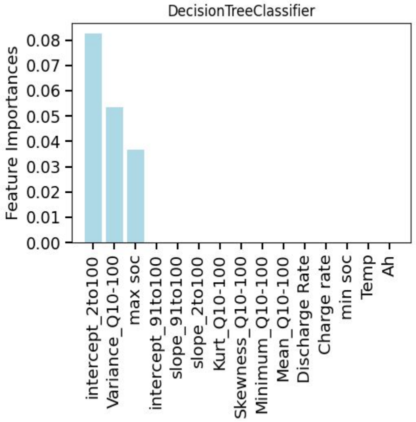
Classifier Models

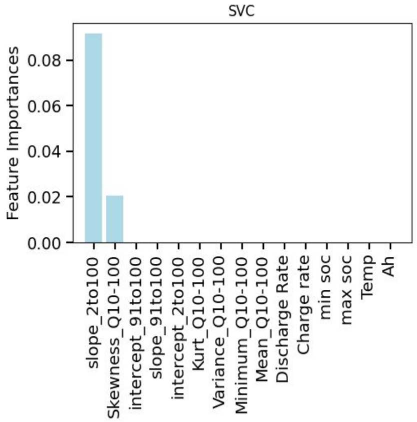
- Is the lifetime of a battery > xyz cycles?
- Use expected lifetimes for prediction
  - LFP: 3000
  - NMC: 600
  - NCA: 400
- Binary Classification, Small-Medium dataset, continuous data, few features: DTC, SVC, LR
- Training F1 scores: highest for DTC and RF
- Validation F1 scores: similar across models
- Test data F1 score: DTC vs. RF

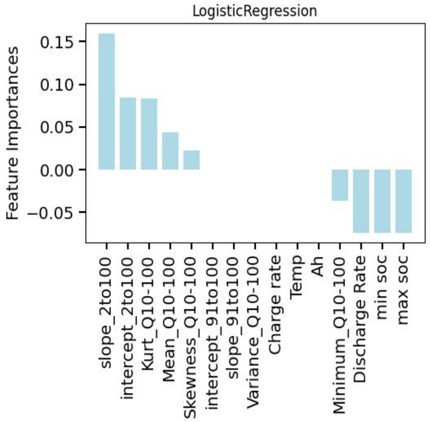


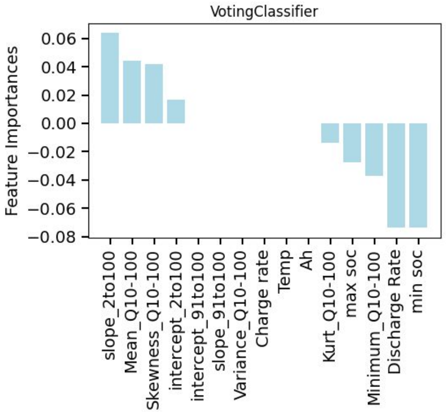
Model	**F1 score (20% Test data)**
Decision Tree Classifier	**0.70**
Logistic Regression	0.82
Support Vector Classifier	0.78
Random Forest Classifier	**0.84**
Voting Classifier	0.71

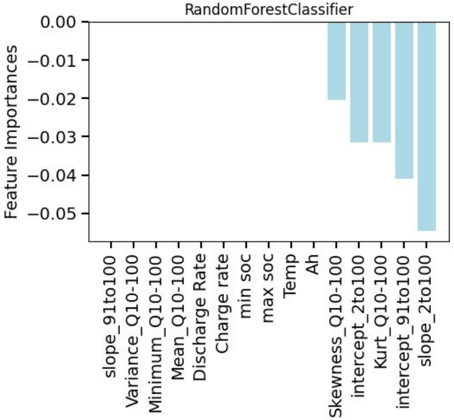
Permutation Importances Analysis











- Negative Importances: $n_{repeats} = 2$
- 2 features had high importance across models
  - 1 linear fit parameter ($Intercept_{2to100}$)
  - 1 statistical feature ($Variance_{Q10\text{-}100}$)
- Random forest had no significant importance for any feature
  - Overfitting possible

Summary

- Linear regression failed to capture non - linear characteristics
- Non-linear regression probably overfitted the dataset
- Classifier models had good training and validation scores
- Decision tree classifier was the best model

Future Directions

- Improve the data preprocessing and outliers removal
- Work with a larger dataset
- Optimise the choice of cycle # for statistical feature generation
  - Earlier cycles
  - Closer cycles
- Deep learning for time series data based forecasting