Machine-Learning Models to Predict Cathode Performance

Clement Wong [1,2], Ishan Saraswat [4], Katsuyuki Tomita [3], Spandan Gandhi [4],
[1]Engineering and Public Policy, [2] Mechanical Engineering,
[3]Chemical Engineering, [4]Energy Science, Technology and Policy
Carnegie Mellon University
CITRINE INFORMATICS

Introduction and Objective

Our objective is to build a model predicting cathode performance (i.e. capacity) using a dataset of cathode materials from the Taylor Sparks Group at University of Utah.

- **What features are important to predict cathode performance?**
- **What machine learning method works best for prediction of cathode performance?**

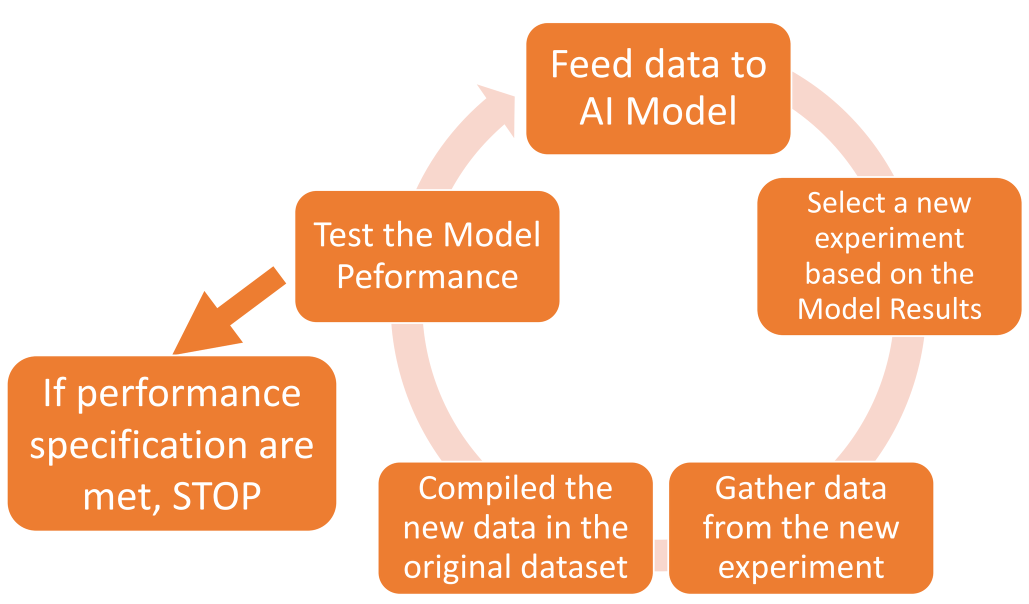
Our Collaborator

Citrine Informatics: Provides a materials data science and artificial intelligence platform that allows users to

1. Store to and extract from material data infrastructure
2. Run machine learning models

Benefits of AI for Materials

- Promotes sharing materials knowledge
- Accelerates material development and discovery through “sequential learning”



Literature Review

Features that affect cathode capacity:

- Material (e.g. chemical composition) [1]
- Design Parameters (e.g. electrode thickness and porosity) [2]

ML algorithms implemented to predict Ni-rich NCM cathode properties [3]

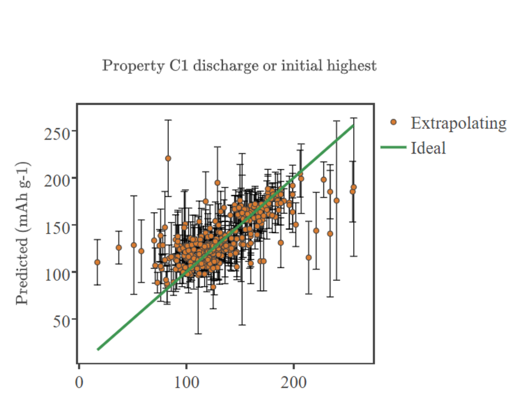
- 13 input features (synthesis parameters, inductively coupled plasma mass spectrometry, X-ray diffraction)
- Extremely Randomized Tree model with AdaBoost algorithm best predicted initial capacity, residual Li, and the cycle life

Results from Citrine Platform

Random Forest Model on Citrination using only features from Taylor Sparks dataset

Important features for capacity

Features	Importance
structure type	26.1%
coulombic efficiency at first cycle	4.7%
mean of Elemental polarizability for formula	4.1%
mean of Number of unfilled p valence electrons for formula	3.4%
Voltage range max	3.0%



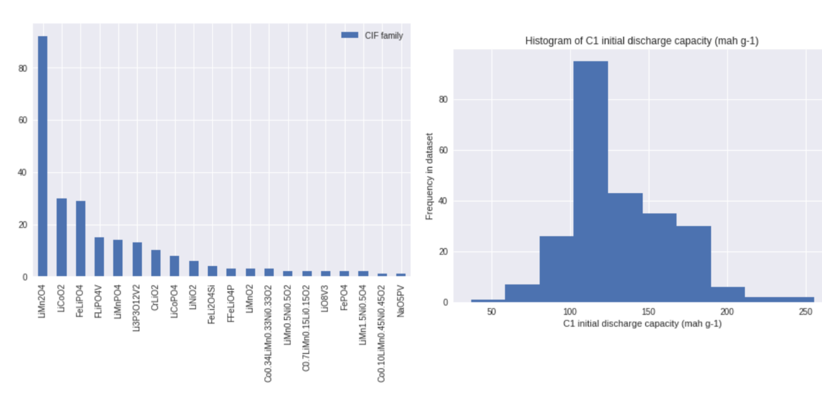
RMSE = 24.4 mAh/g

Our Method

Data Cleaning and Cross Checking

- Removed sparse features /experimental test conditions
- Remove rows which have missing values for important features
- Fill in missing/incorrect data from the research papers

Exploratory Data Analysis (EDA)



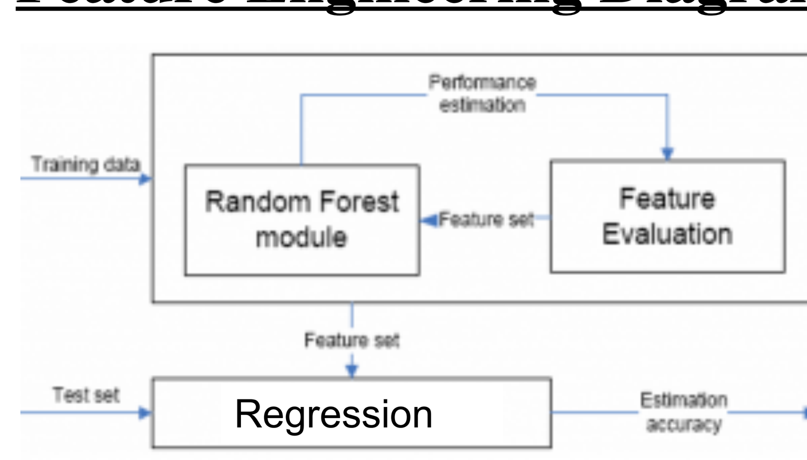
Data Retrieval and Feature Engineering

Open-Source Databases

pymatgen matminer CITRINE INFORMATICS

- Extract important features such as density, electronegativity, and crystal structure (spacegroup)

Feature Engineering Diagram



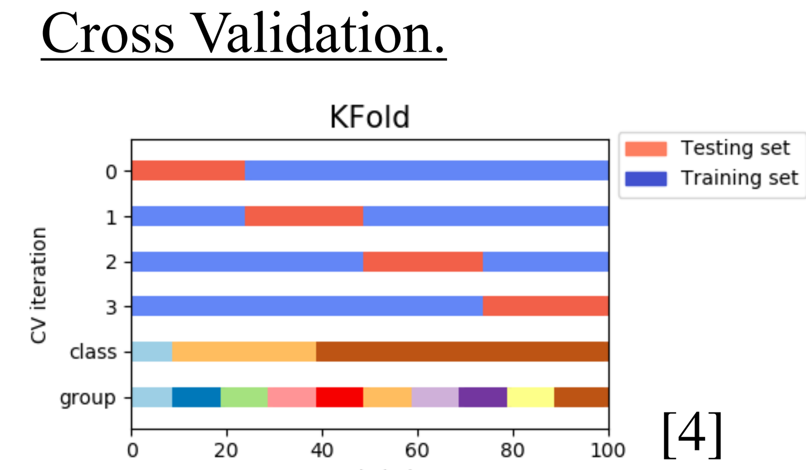
- Conducted Recursive Feature Elimination to narrow to 15 features

Model Selection and Parameter Optimization

Models Implemented

- Linear Regression
- Random Forest
- Decision trees
- Support Vector Regression

Accuracy of the model in terms of RMSE and R^2 score using K Fold Cross Validation.



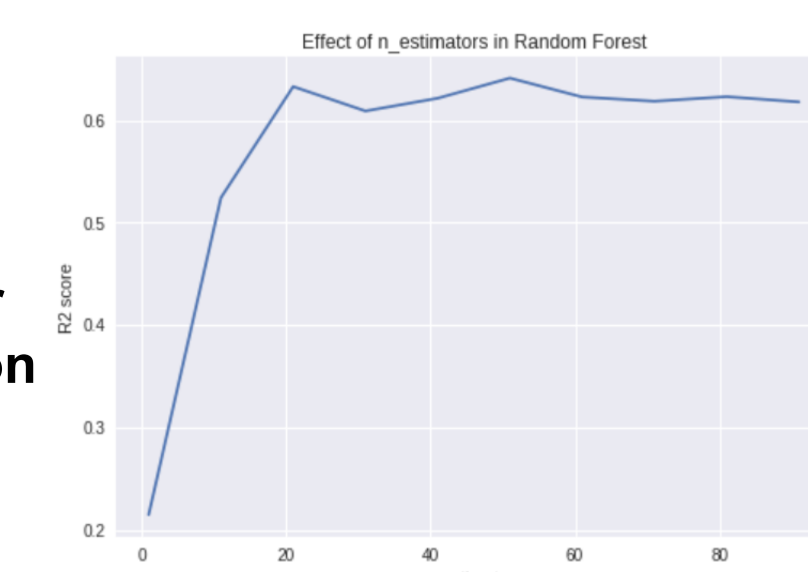
[4]

Our Results

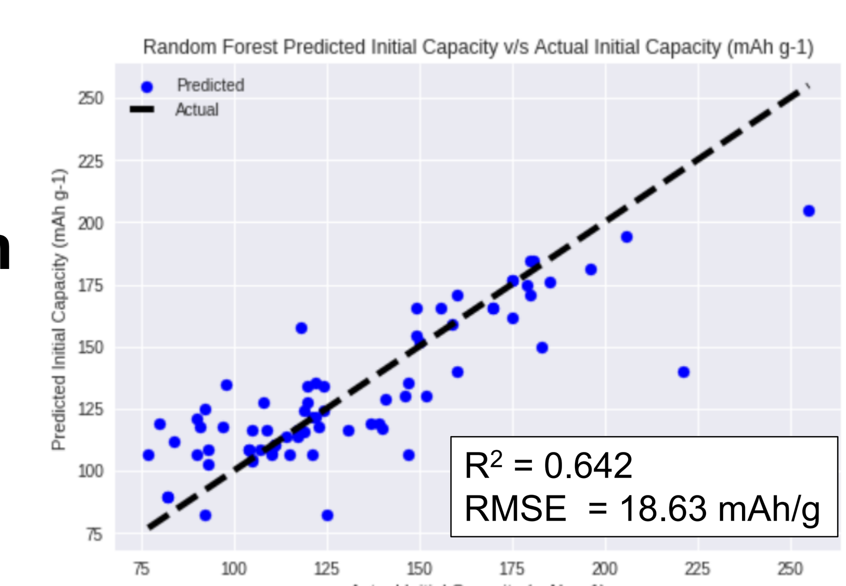
Features Selected As Important From Recursive Feature Elimination

Electrode Composition	Voltage range min (V)	Molecular Mass (g/mol)	Mean of Elemental Crystal Structure (Space Group)
Coating	Voltage range max (V)	Elemental Fraction	Mean of Interatomic distance
Average Electronegativity	Estimated Theoretical Capacity (mAh/g)	Structure Type	

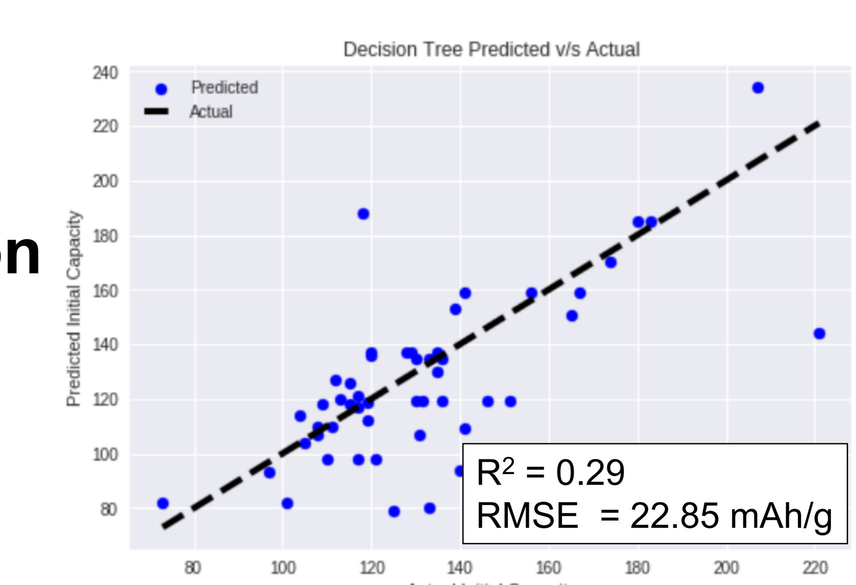
Hyper-parameter Optimization



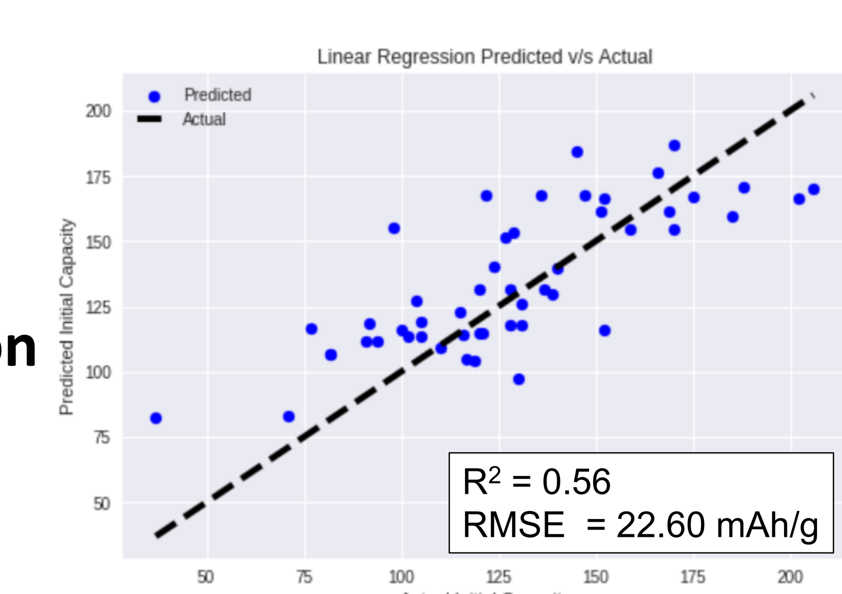
Random Forest



Decision Tree



Linear Regression



Suggestions for Future Work

- Adding more data (Trained on 247 training examples out of the initial 343 in dataset due to missing/incorrect values)
- Test with more material properties to identify other predictors
- Predict other performance properties such as cycle life

Summary

- Random Forest Model best predicted initial capacity
- Founded 15 best features to represent the data using Recursive Feature Elimination
- With more data on cathodes, the model would be more robust

References

[1] A. Eddahech, O. Briat, J.-M. Vinassa. Performance comparison of four lithium–ion battery technologies under calendar aging. Energy, 84 (2015), pp. 542-550
[2] S. Yu, S. Kim, T. Y. Kim, J. H. Nam, and W. I. Cho, Bulletin of the Korean Chemical Society, 34, 79 (2013).
[3] Min, K., Choi, B., Park, K. & Cho, E. Machine learning assisted optimization of electrochemical properties for Ni-rich cathode materials. Sci. Rep. 8, 15778 (2018)
[4]“Scikit-Learn.” *1.4. Support Vector Machines - Scikit-Learn 0.19.2 Documentation*, scikit-learn.org/stable/index.html.