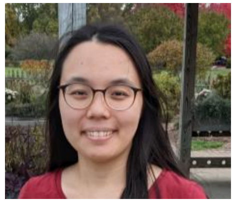
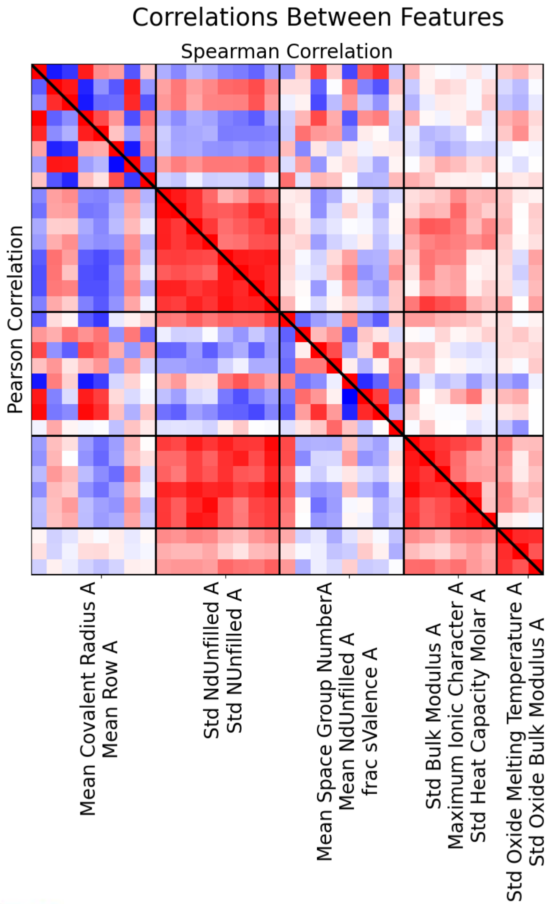
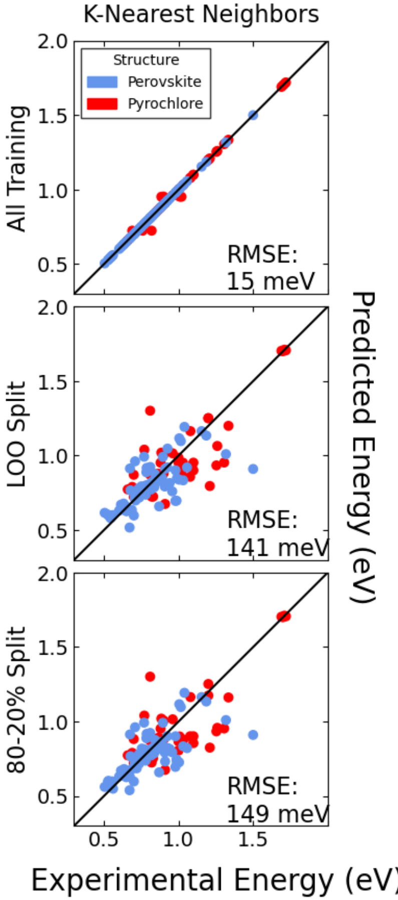
Explainable Machine Learning Models for Predicting Oxygen Activation Energies in Perovskites and Pyrochlores

Grace M. Lu [1], and Dallas R. Trinkle [1]

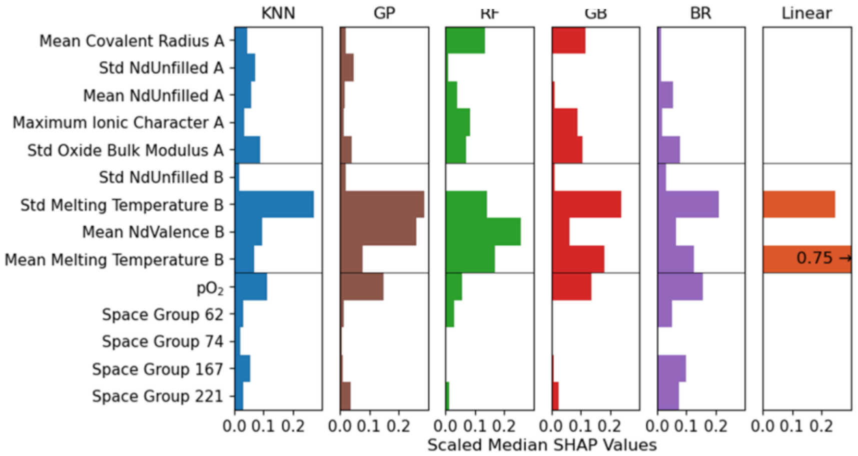
[1] Department of Materials Science and Engineering, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, US



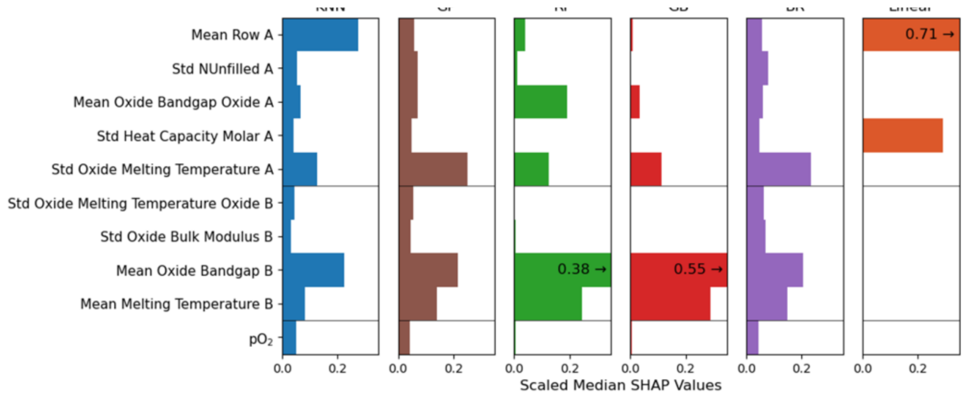




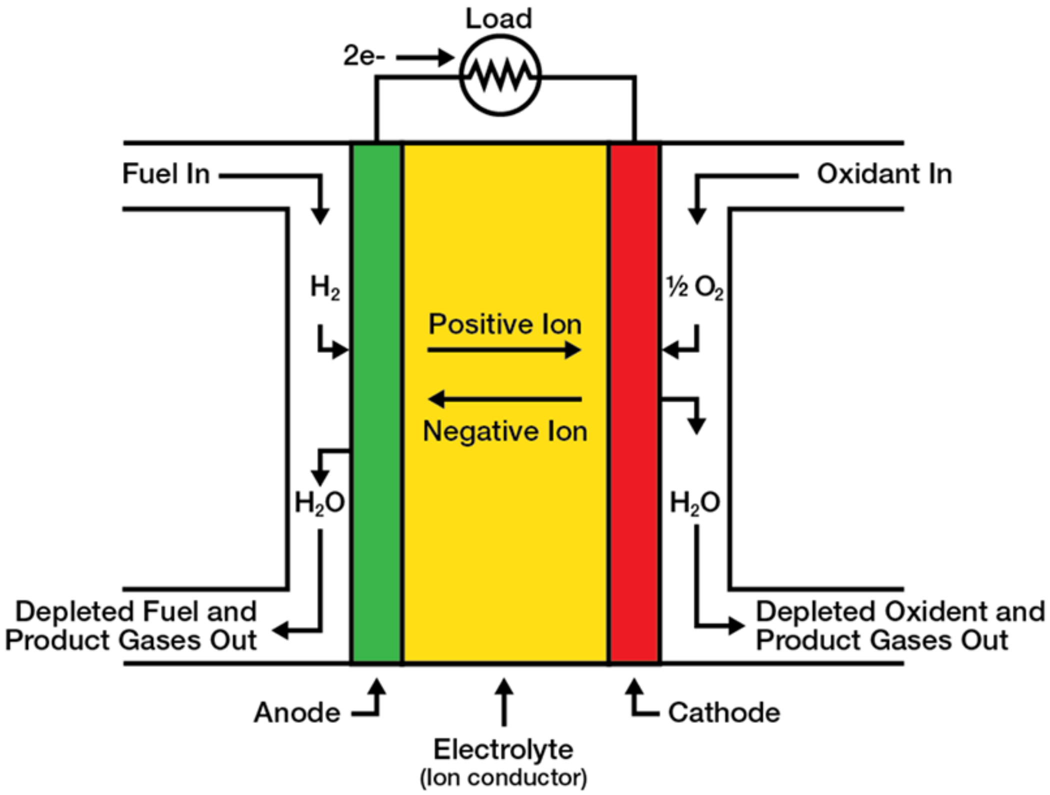
Perovskite



Single Vacancy Pyrochlore



Solid Oxide Fuel-Cells

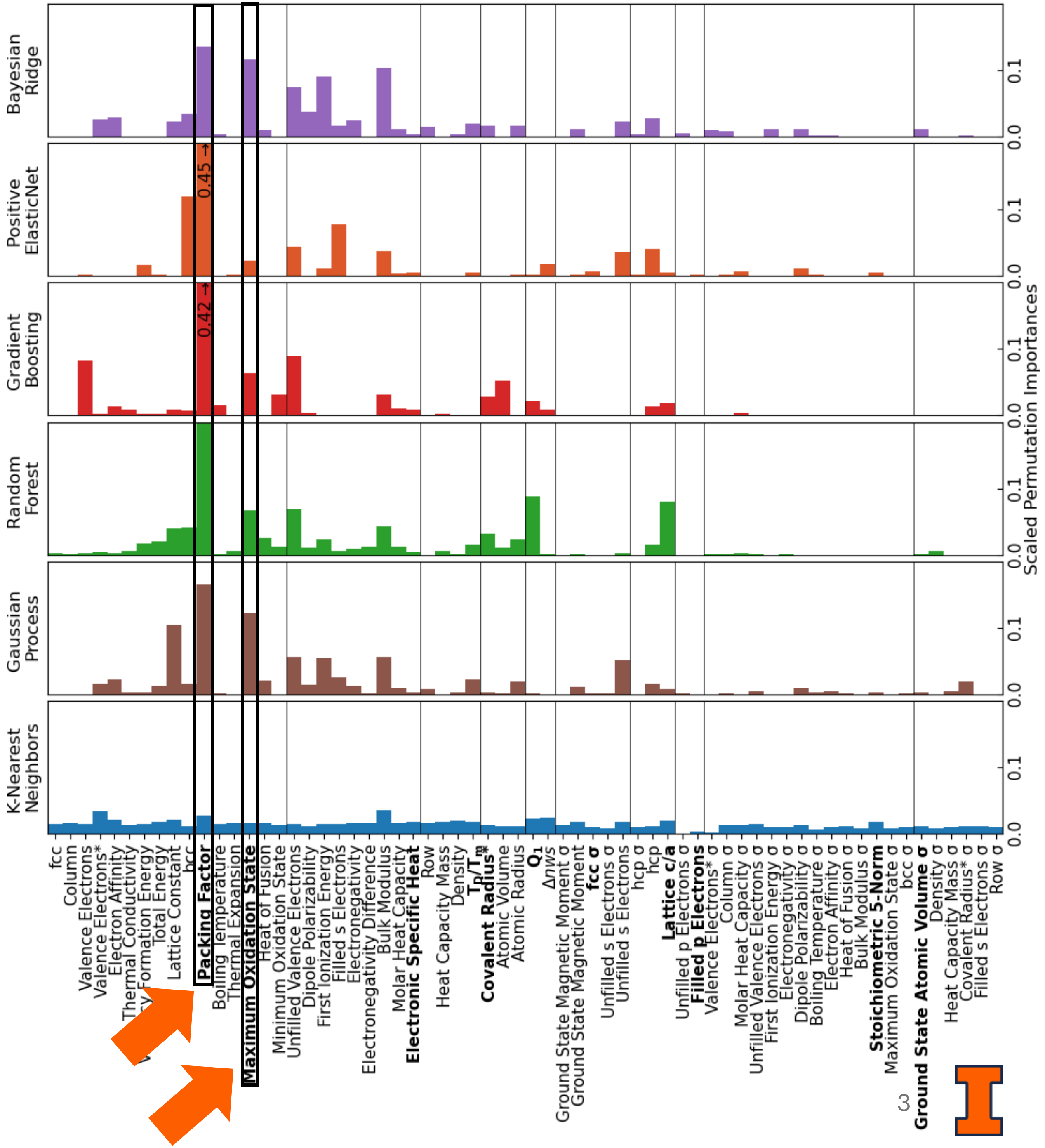


SOFC Operating Principle. *National Energy Technology Laboratory*. https://netl.doe.gov/carbon-management/sofc/operating-principle

- Motivation: Reduce operating temperature (500^{o}C) – improve oxygen diffusion kinetics
- Large experimental space – rapid screening of new materials
- Goal: Obtain physical insight into oxygen diffusion using a grouped analysis of machine learning models
  - Why ML? More complicated relationships than just correlation
  - Why grouping? Negate inconsistencies between models

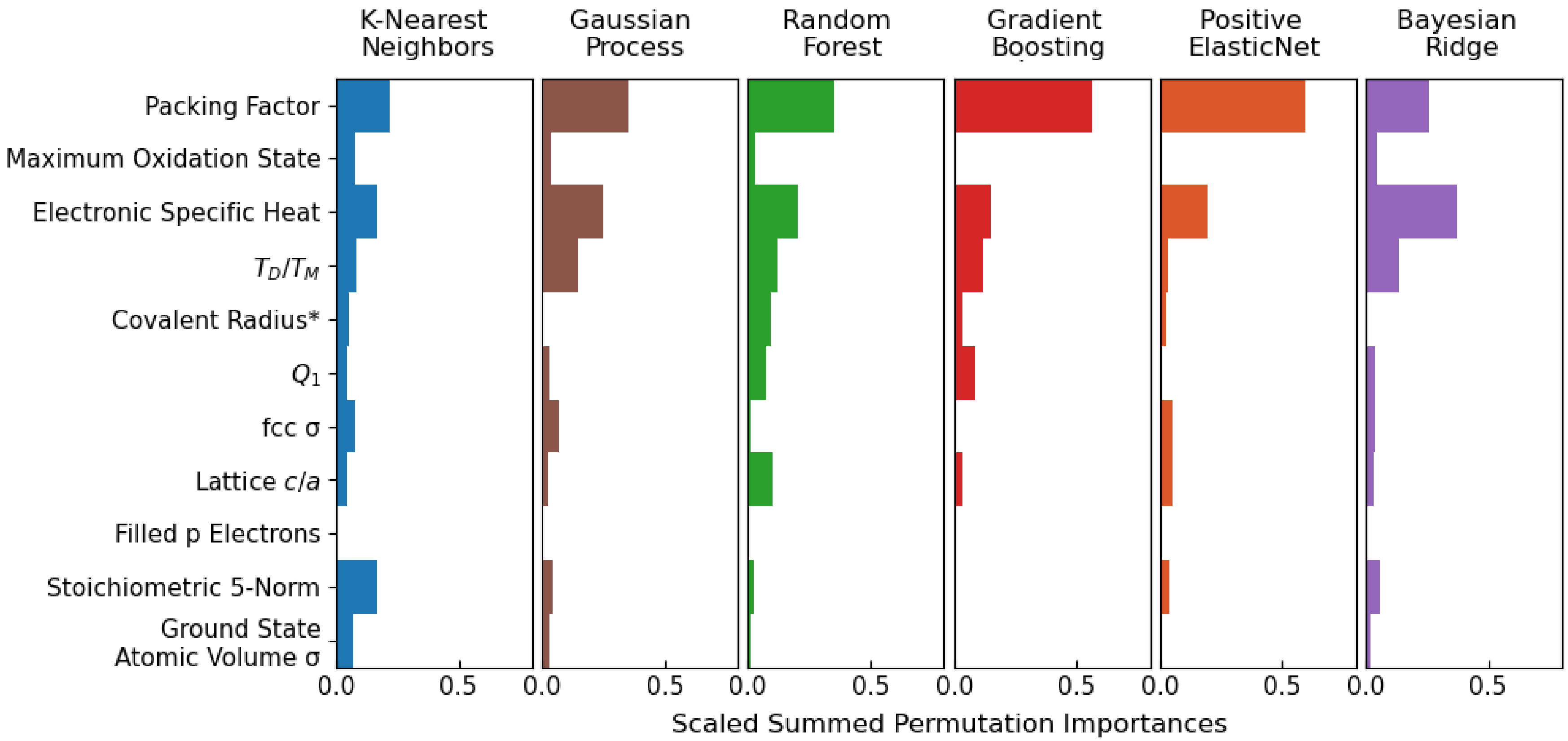
Individual Feature Importances for Hydrogen Diffusion in Metals[1]

- Permutation importances: shuffled feature impacts R^2
- Packing factor is the most important feature
- Inconsistent features between models
- Top 3 feature of each model: 9 different features



[1] **GML**, Witman, M., Agarwal, S., Stavila, V., Trinkle, D. R. (2023)

Grouped Features Enable Interpretability for H Diffusion [1]



- **Packing factor**: Diffusion pathways
  - Structural
  - Energies
- **Electronic specific heat**: H interaction with bulk
  - Electronic properties
  - Bulk modulus

- Similar features to Flynn and Stoneham's self-trapping model [2]

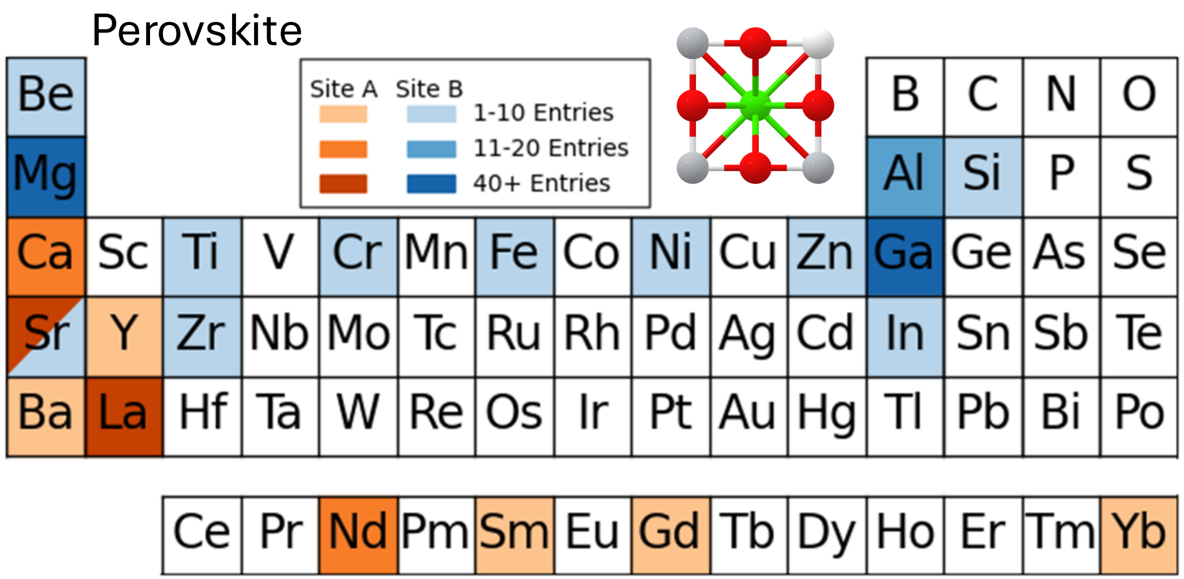
$$Q = \frac{M\omega_D^2 d^2}{360}\left(\frac{1+\nu}{1-\nu}\right)^2\left(\frac{\delta V}{\Omega}\right)^2 \Phi(q_m d, \eta)$$

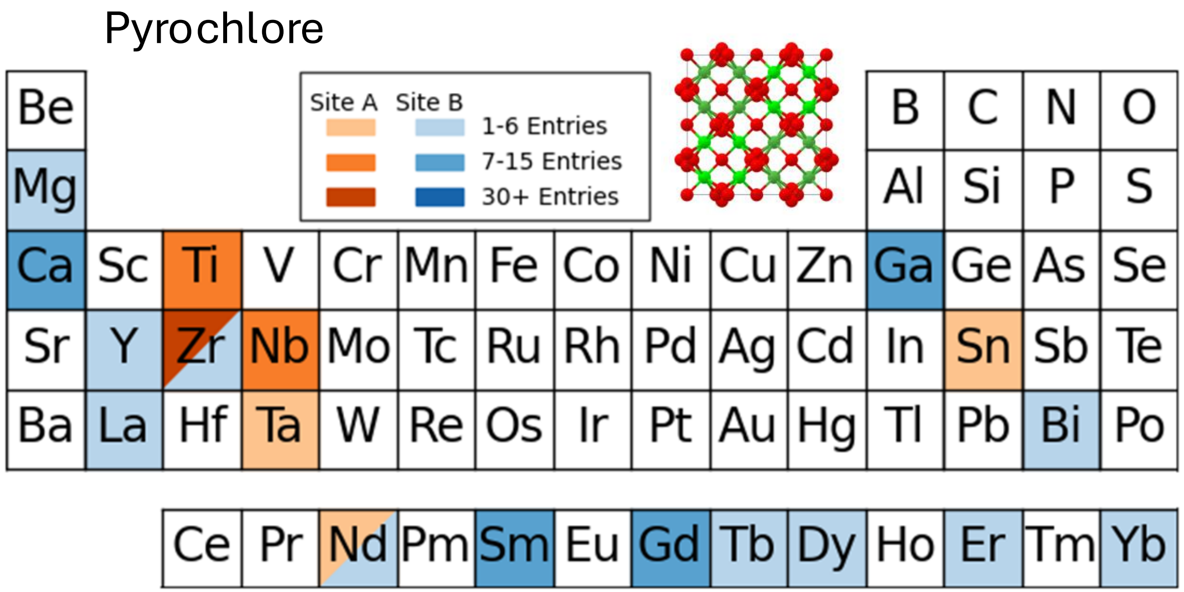
M: Atomic mass
w_D: Debye frequency
d: Jump distance
ν: Poisson's ratio
δV: Lattice dilation due to defect
Ω: Solvent atomic volume

q_m: Radius of Debye sphere
η: fraction of volume change due to dipole component of the strain field

[1] **GML**, Witman, M., Agarwal, S., Stavila, V., Trinkle, D. R. (2023)
[2] Flynn, C. P., and Stoneham, A. M. (1970)

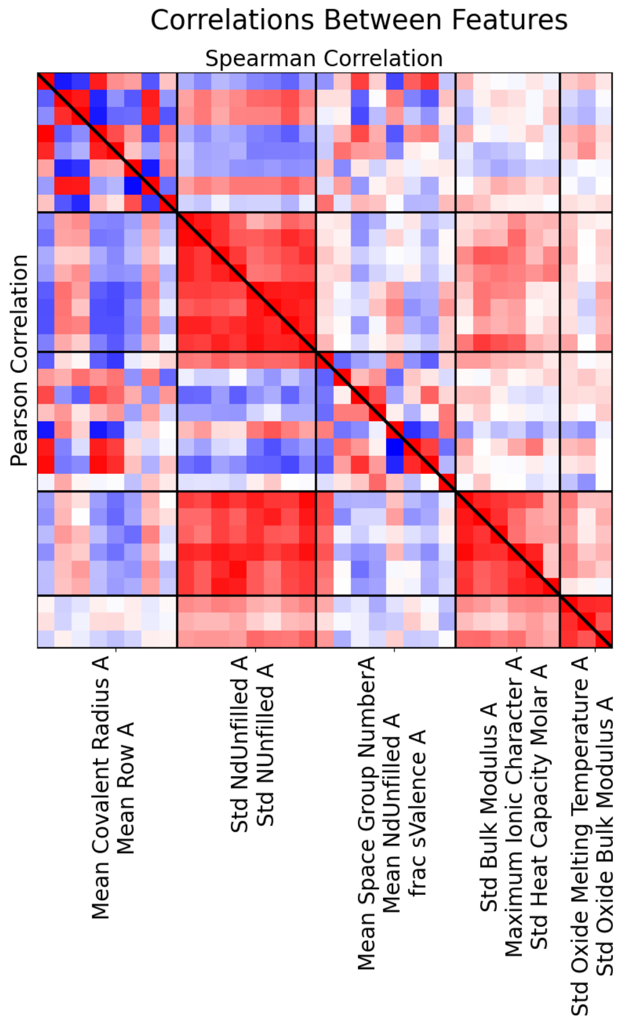
Database (Perovskites and Pyrochlores)

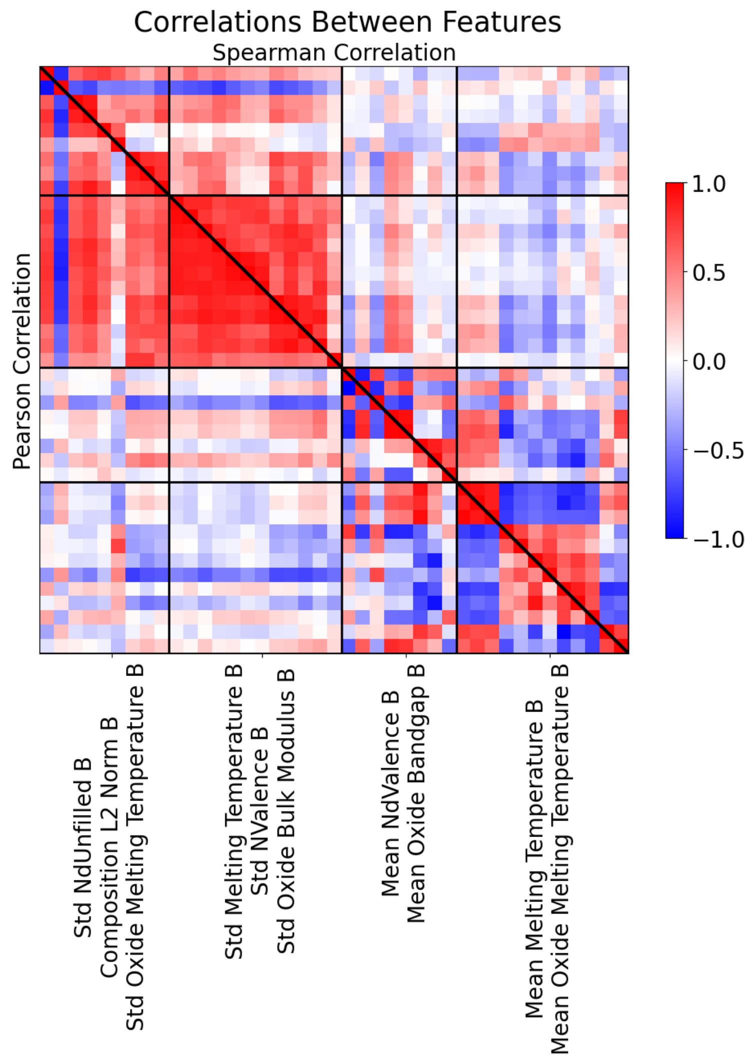




- 76 Perovskites (ABO_3) and 42 Pyrochlores ($A_2B_2O_7$) – oxides with the most experimental measurements
- Number of times each element appears in the database
- La (62), Ga (60), and Sr (57) most prevalent perovskite elements
- Zr (34) most common pyrochlore element

Easily-Obtained Features

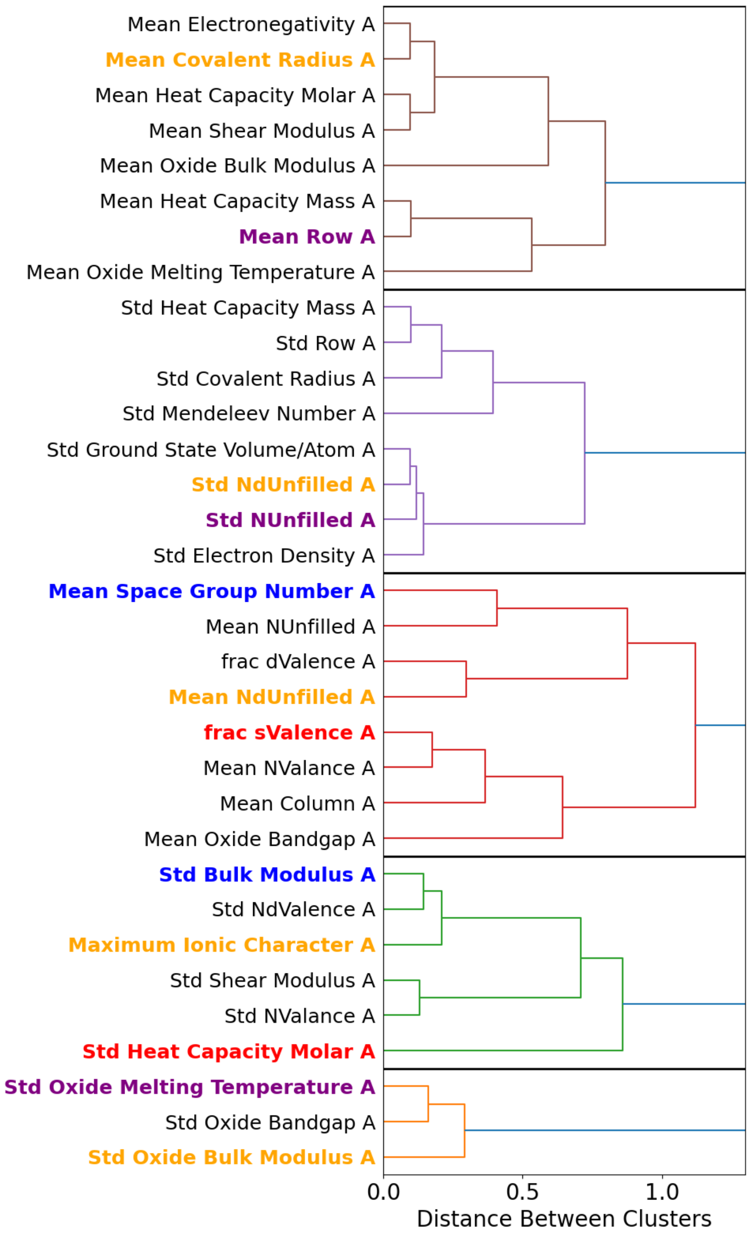


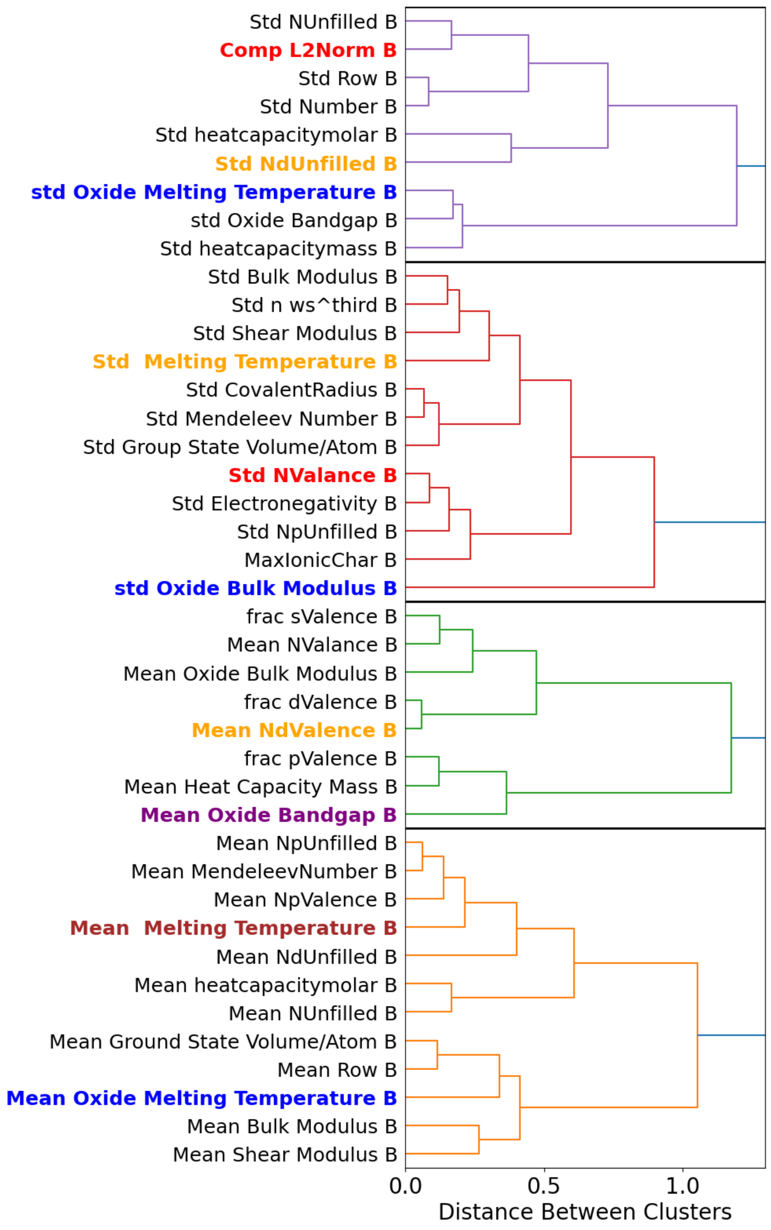


- 139 total features from MAGPIE [1] – combine elemental properties
  - Oxygen not included
  - Split A and B site
  - One-Hot Encoding of the space group number
  - Oxygen partial pressure: diffusion mechanism
- Oxide properties – band gap, bulk modulus, melting temperature
- Strong correlations between features

[1] Ward, Logan, *et al.* (2016)

Feature Reduction



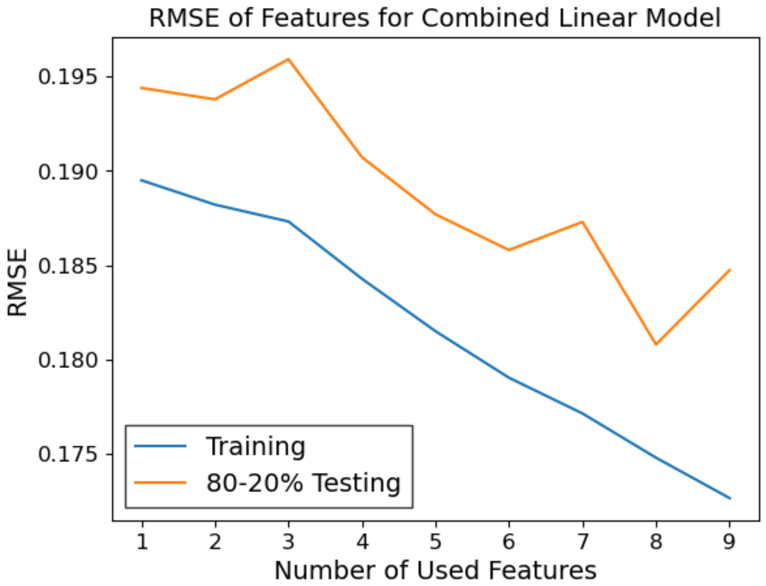


- Feature reduction occurs first before model fitting
  - Ward hierarchical clustering
  - Groups most correlated features at each step
- Group A/B features separately using combined database
  - Choose included feature using a greedy algorithm to maximize correlation with residual
  - Same grouping used, but different features

Feature Used in Model
- Perovskite
- Only Single Vacancy Pyrochlore
- Only All Pyrochlore
- Both Pyrochlore
- Perovskite and Single Vacancy Pyrochlore

Machine Learning Models

- 6 ML models: small datasets
- Linear Model – number of features minimize 80-20% test error

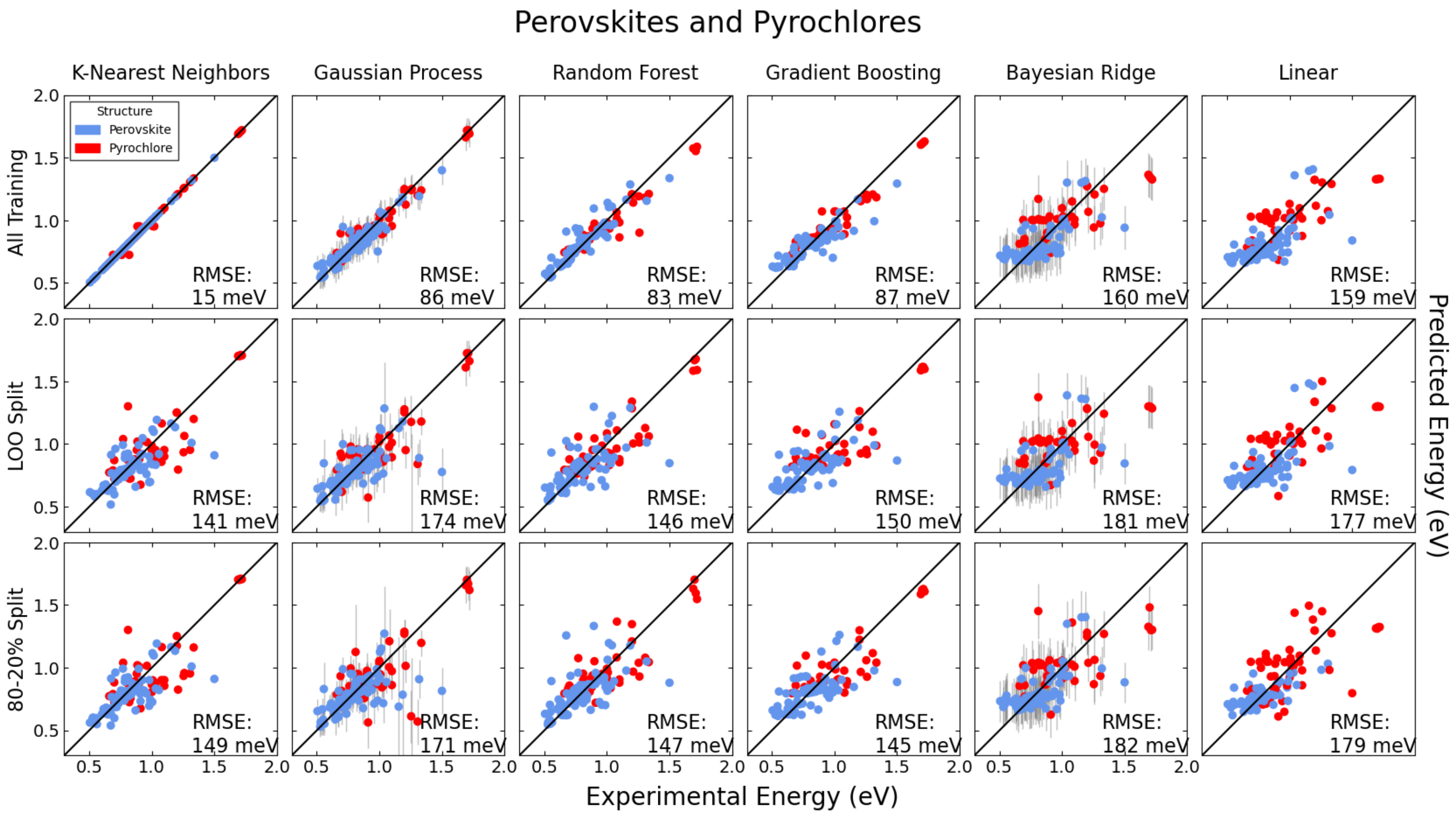


- 3 test-train splits
  - All training
  - Leave-One-Out
  - Random (80-20%)

ML Models	Kernel Based	Decision Tree	Linear	Error Prediction
K-Nearest Neighbors	■			
Gaussian Process	■			■
Random Forest		■		
Gradient Boosting Tree		■		
Bayesian Ridge			■	■
Linear			■	

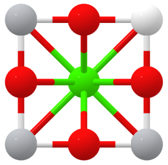
Combined Machine Learning Model

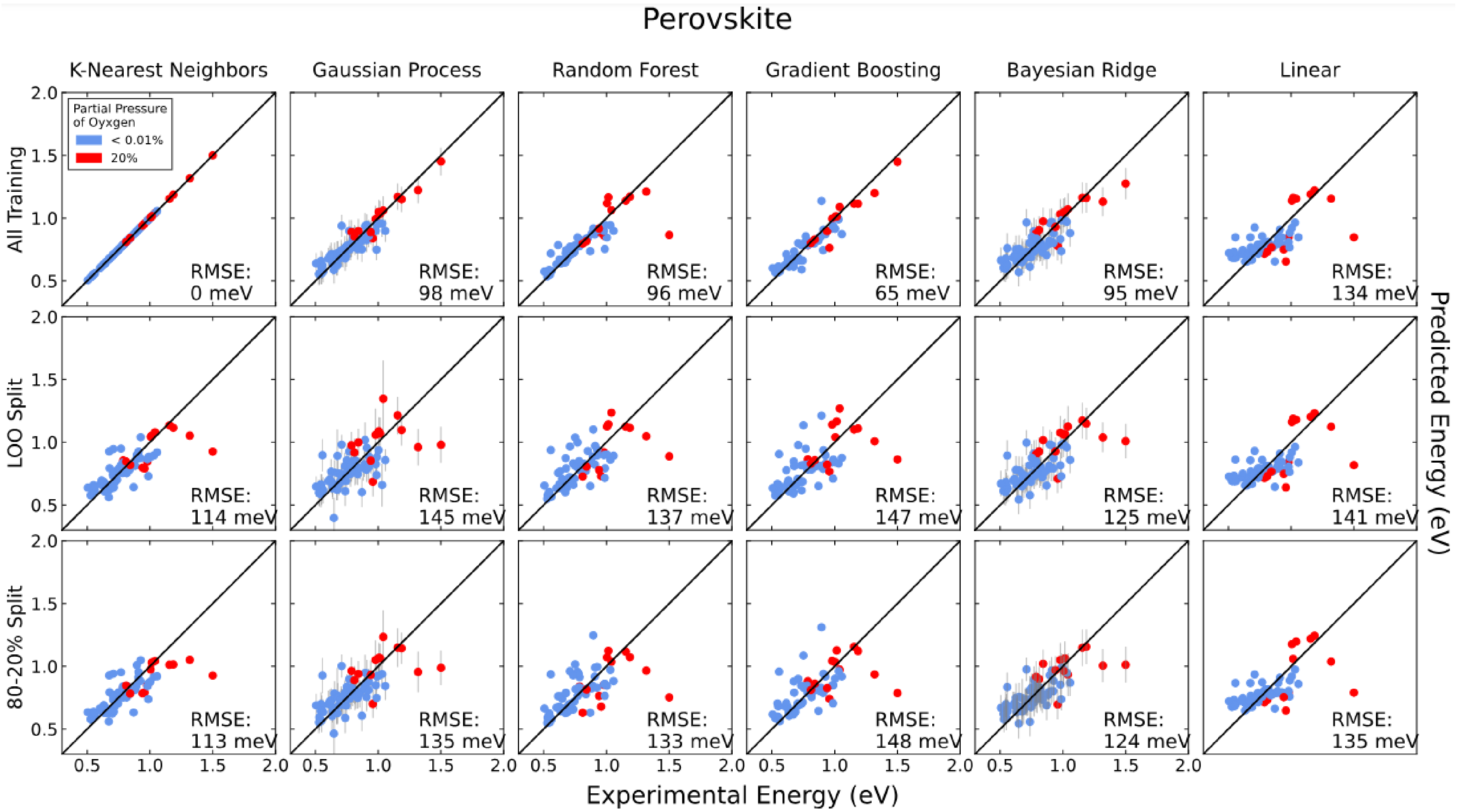
- Larger standard deviation for pyrochlore activation energies (282 meV) vs. perovskite (185 meV)
- Linear models have poor training errors – cannot grasp difference between pyrochlores and perovskites



Perovskite Machine Learning Model

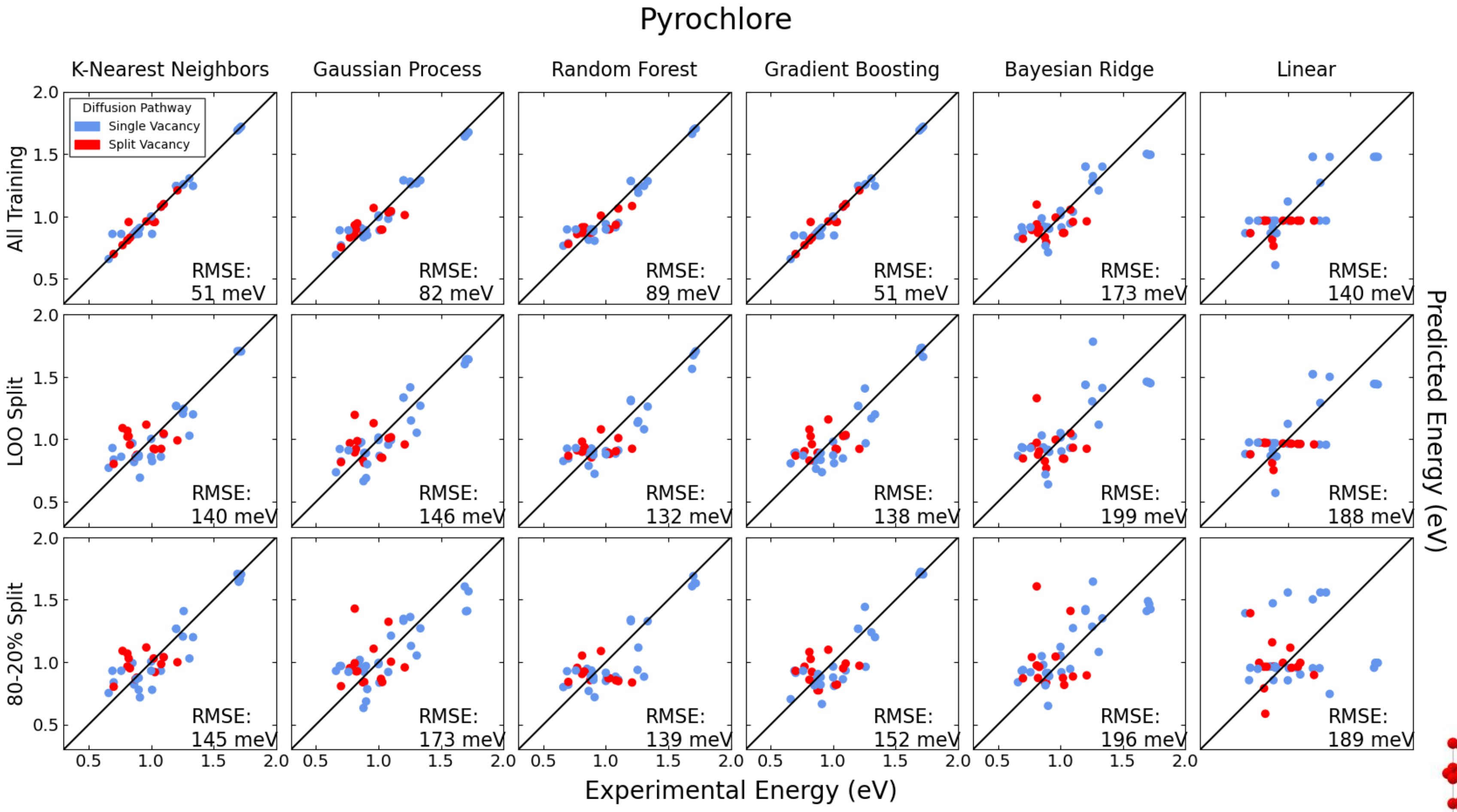
- Largest experimental activation energy ($LaCaO_3$) also has the largest error (618 meV) almost twice as large as the second largest (372 meV)
- High Oxygen partial pressure



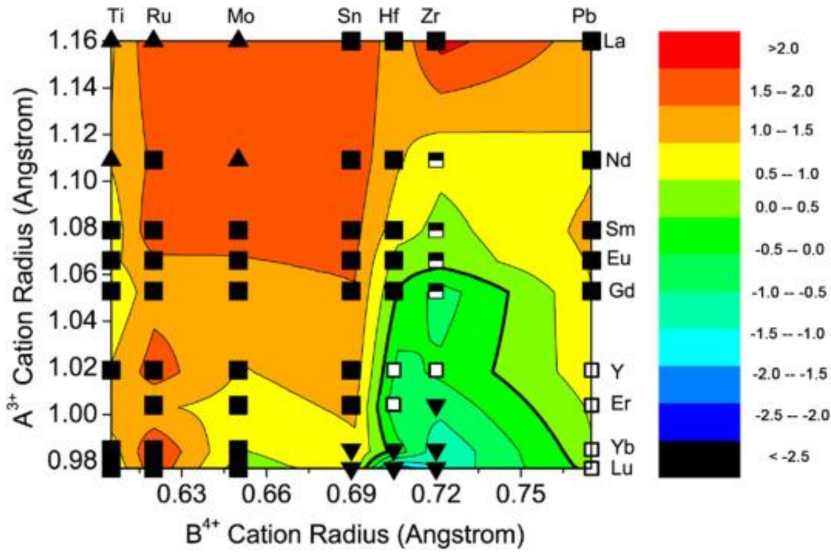


Pyrochlore Machine Learning Model

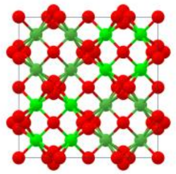
- Multiple diffusion pathways:
  - DFT-based Single vacancy (26) vs. Split vacancy (14) based on cation radii
- Models, especially linear, cannot differentiate



DFT Energy Difference Between Single and Split Vacancy

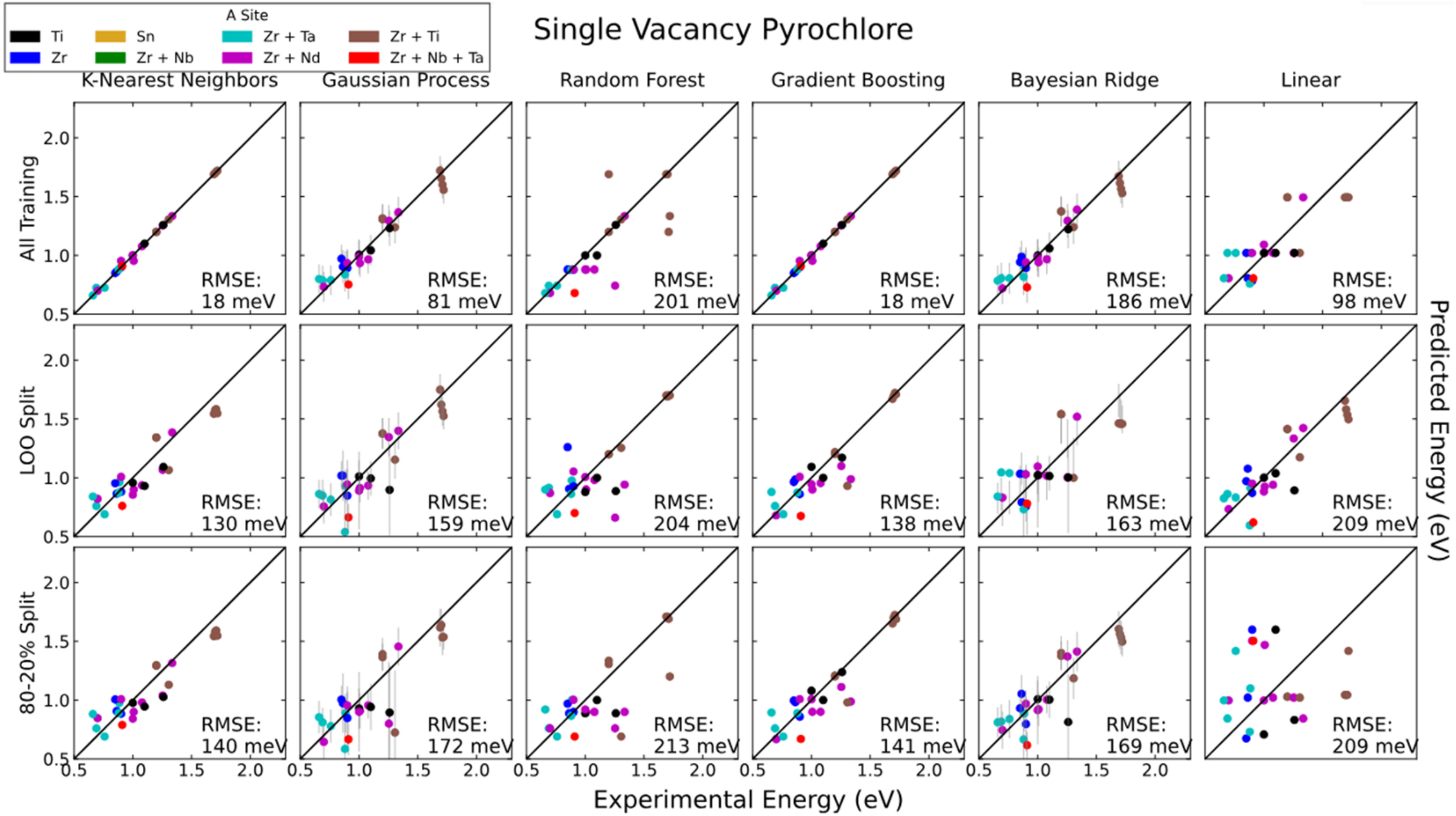


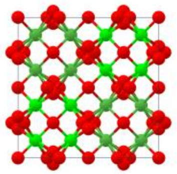
Li, Y. and Kowalski, P. M. (2017)



Pyrochlore Single Vacancy

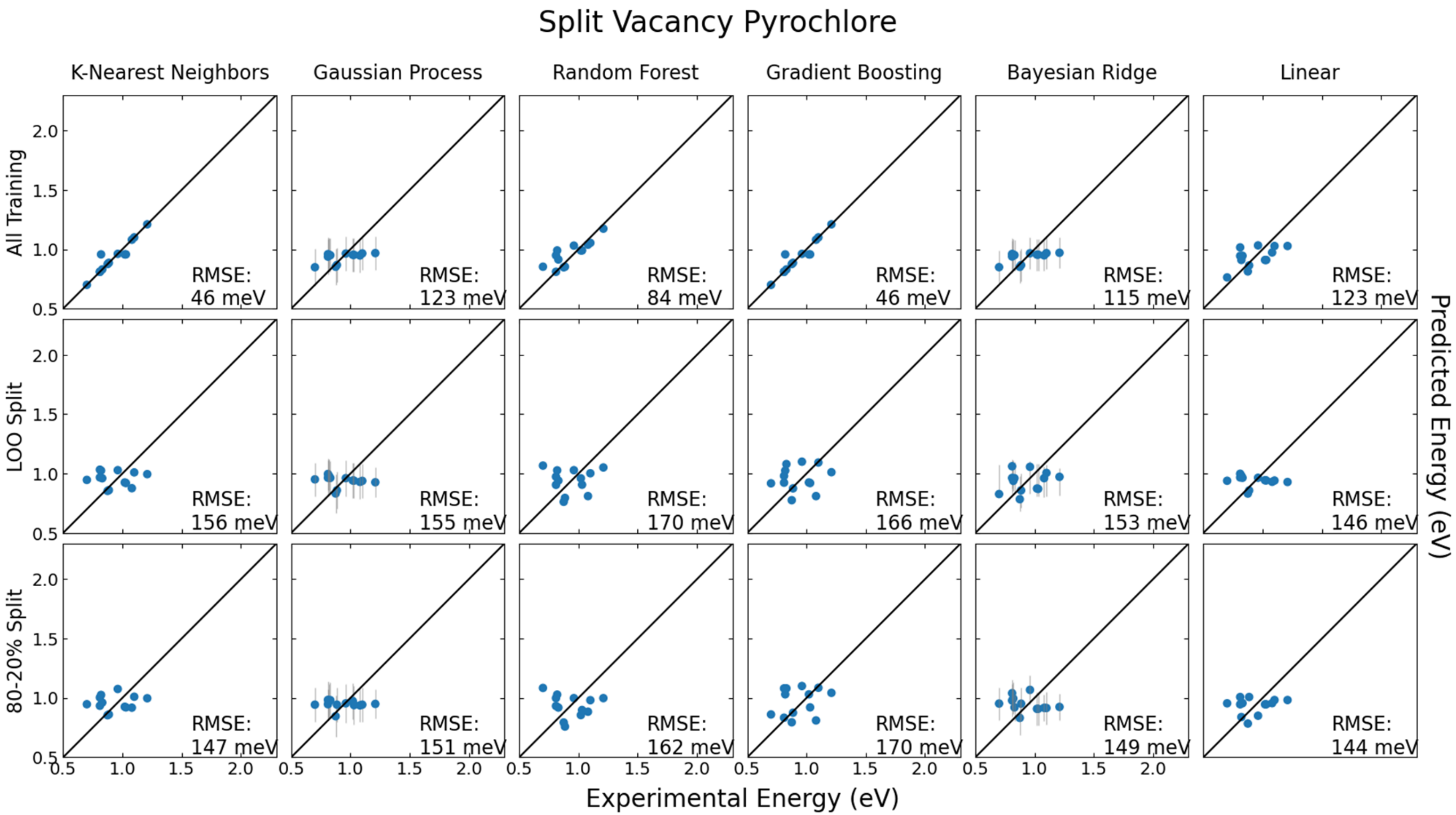
- Similar performance to full pyrochlore models
- Linear model still lacks the ability to fit activation energy

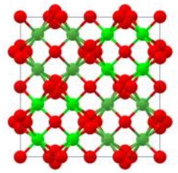


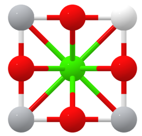


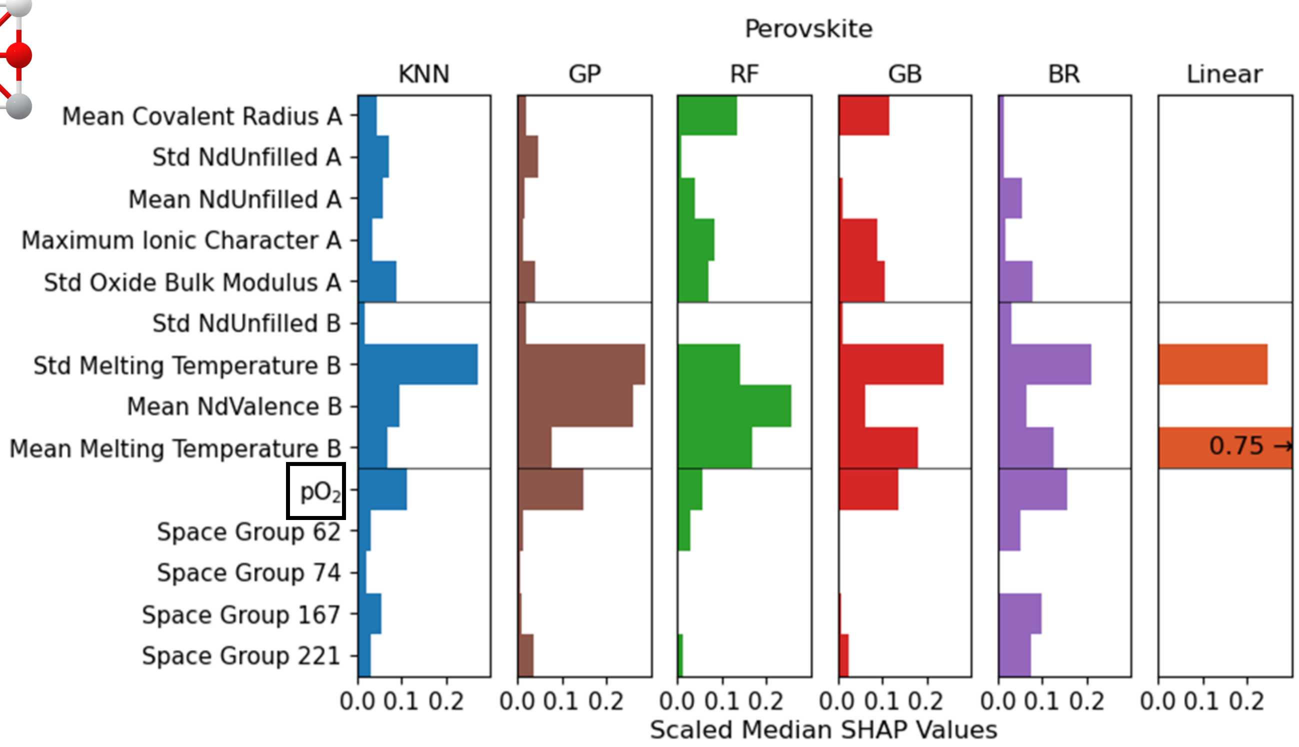
Pyrochlore Split Vacancy Woes

- Large RMSEs for split vacancy models
  - Larger than the standard deviation (141 meV) for any test data
- Low correlation between activation energies and features: highest correlation = 0.49



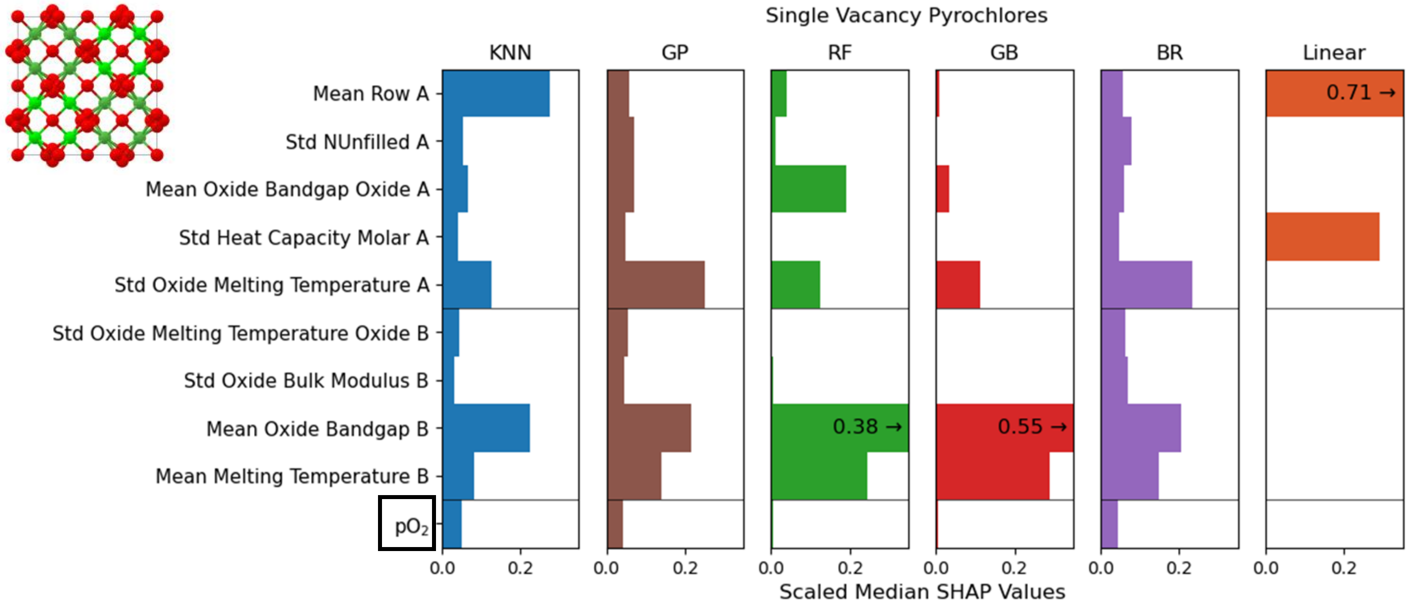






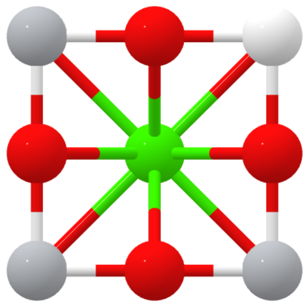
Feature Importances

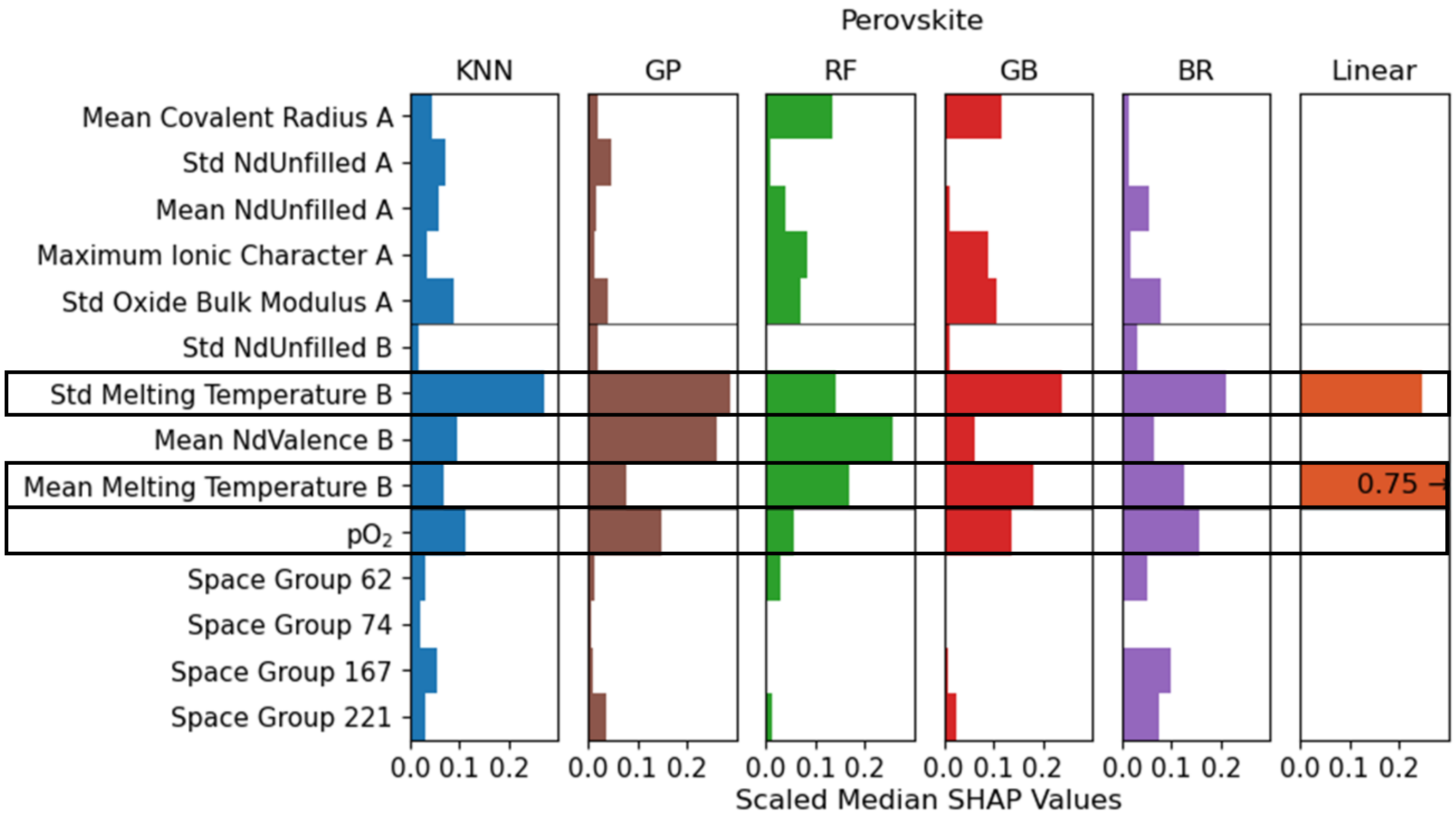
- Median Shapley additive explanation (SHAP) Importances: measure how the feature impacts each prediction
- Site B more important than Site A
  - Site A features are more important for pyrochlores than perovskites
- Agreement between models
- pO_2: differences arise from experimental conditions



Perovskite Feature Importances

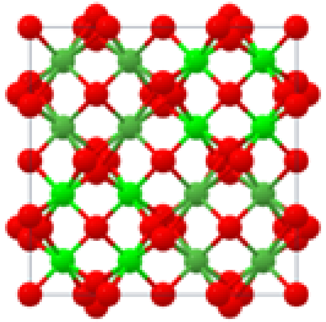
- Most important features: Std Melting Temperature B, Mean Melting Temperature B, pO_2
- Agreement with linear model – features also have high correlation with the activation energy

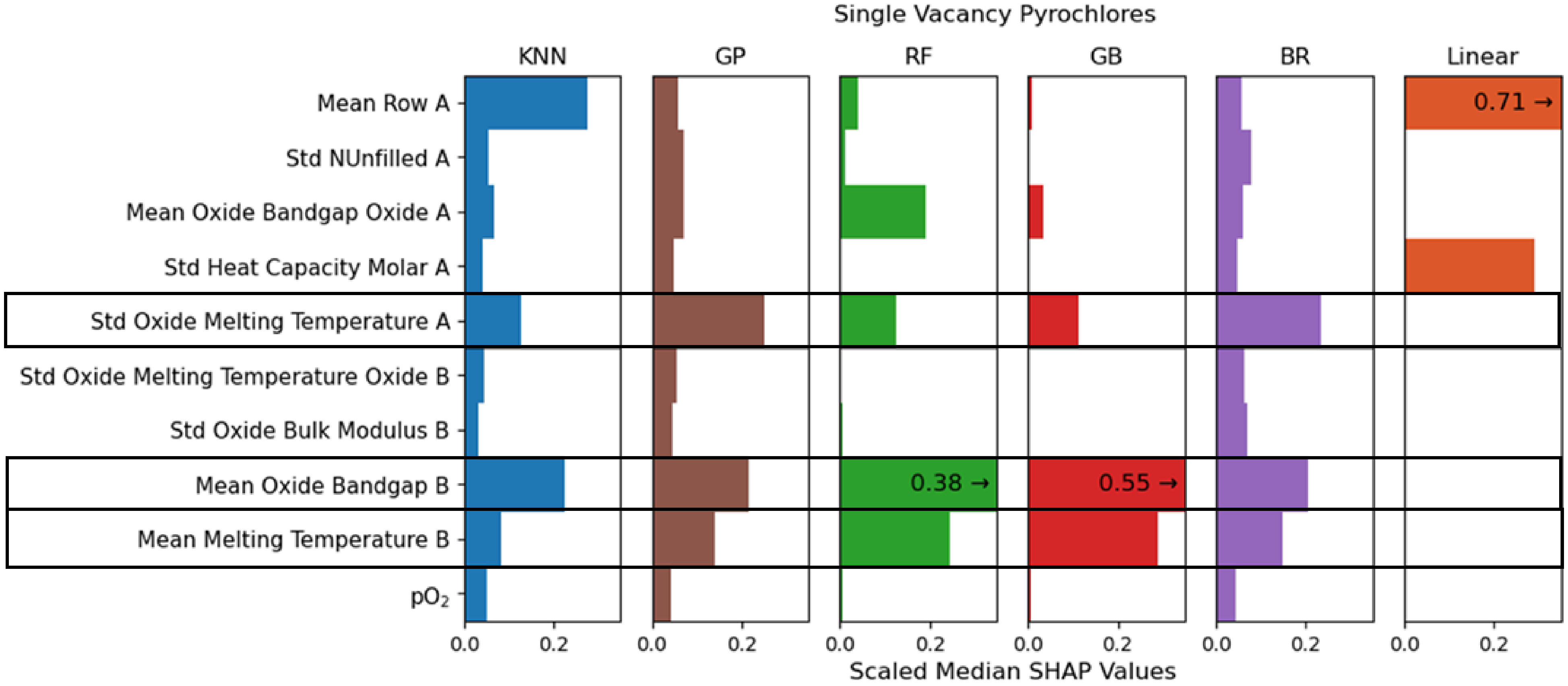




Pyrochlore Feature Importances

- Most important features: Mean Oxide Bandgap B, Mean Melting Temperature B, Std Oxide Melting Temperature A
  - Oxide features are a better representation of pyrochlore features
- Disagreement with linear model

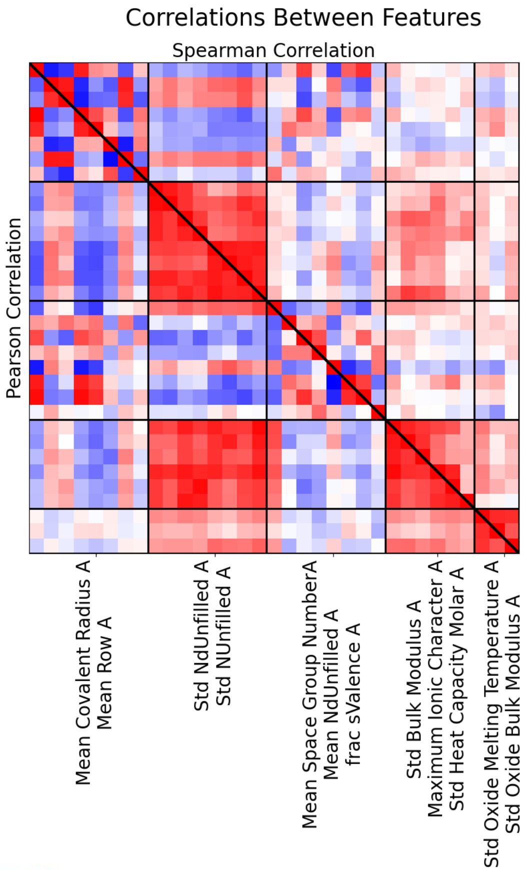


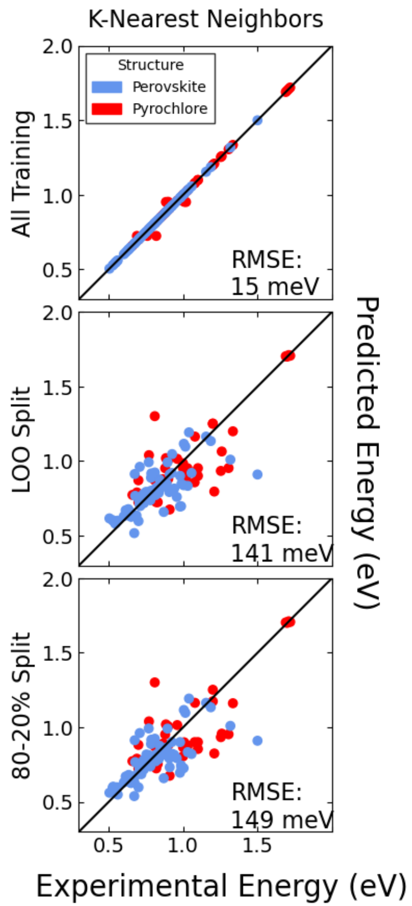


Explainable Machine Learning Models for Predicting Oxygen Activation Energies in Perovskites and Pyrochlores

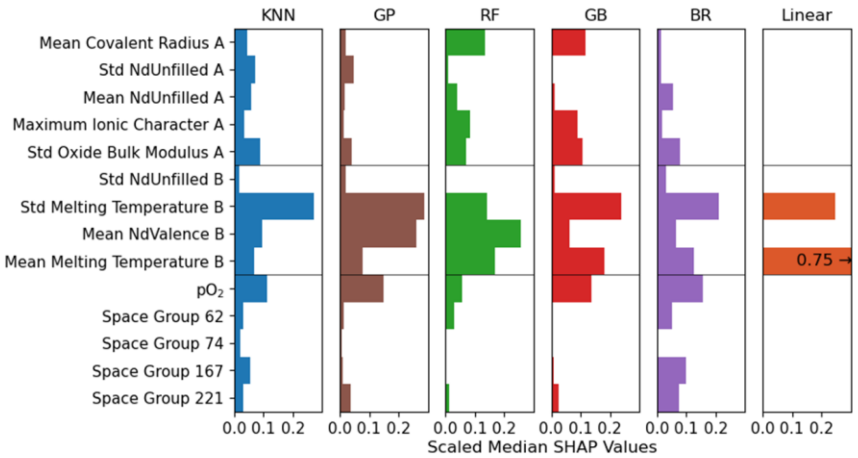
Grace M. Lu [1], and Dallas R. Trinkle [1]

[1] Department of Materials Science and Engineering, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, US





Perovskite



Single Vacancy Pyrochlore