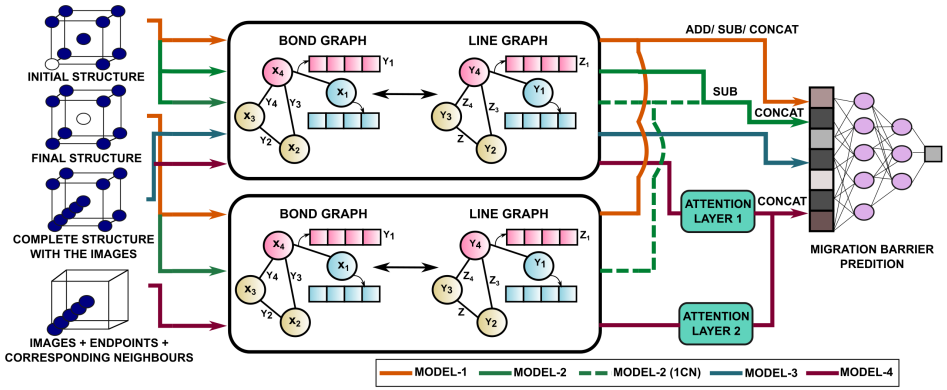


Optimal use of transfer learning for predicting ionic migration barriers in solids

[1]Reshma Devi, [1]Avaneesh Balasubramanian, [2]Keith T. Butler, [1]Sai Gautam Gopalakrishnan

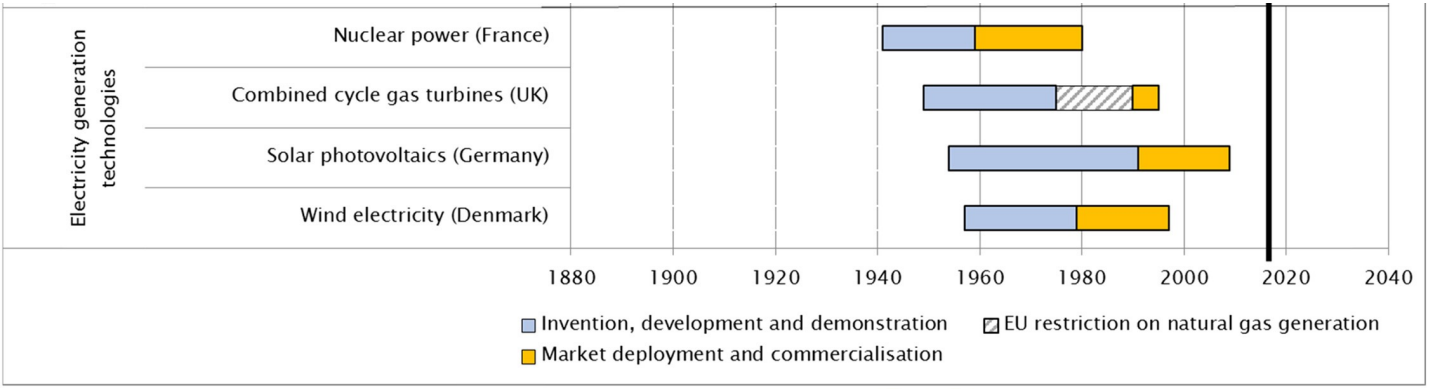
[1]Materials Engineering, Indian Institute of Science

[2]Chemistry, University College London

saigautamg@iisc.ac.in; https://sai-mat-group.github.io

Institut für Nachhaltige Wasserstoffwirtschaft (INW-1), Forschungszentrum Jülich
Sep 15, 2025

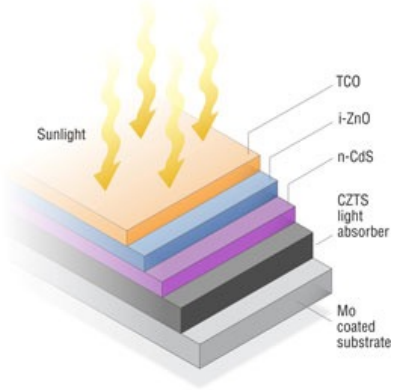
We work broadly on energy materials



Technological innovation cycle → slow in energy sector

Need rapid identification of novel materials for improving performance

Design better electrodes and solid electrolytes



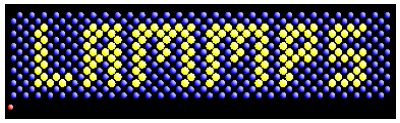
Develop better light-absorbing semiconductors

Identify better materials for H_2 generation and sensing

Identify novel materials for applications

Understand underlying materials phenomena better

Make theory better/build models



Acknowledgments

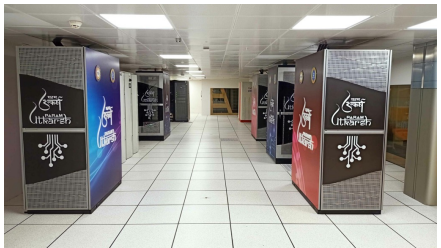
Group picture in May 2025

Dr. Keith Butler

Reshma

Avaneesh

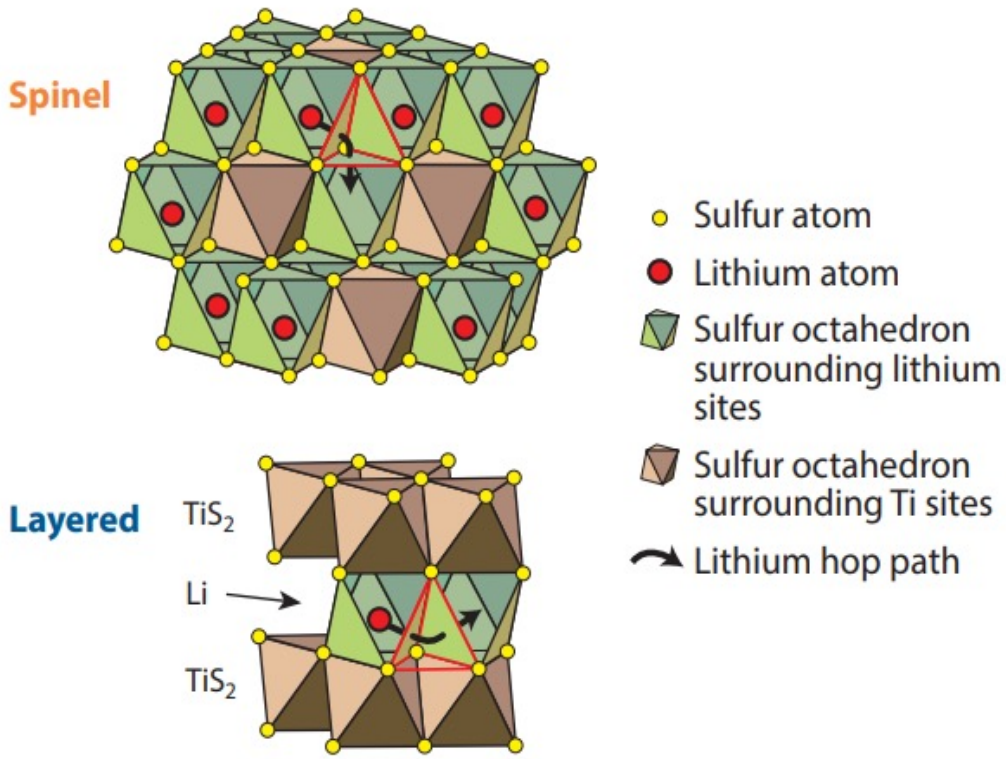
SERC (IISc)



Param Utkarsh (CDAC)

Archer (UK)

Migration barriers govern rate performance in batteries



Intercalation electrodes: ionic diffusivity (D) within the bulk a major factor in rate performance

$$D = D_o \exp(-\frac{E_m}{k_B T})$$

D_o : Diffusivity pre-factor (carrier concentration, correlations, etc.)
k_B: Boltzmann constant
T: Temperature
E_m: Migration barrier

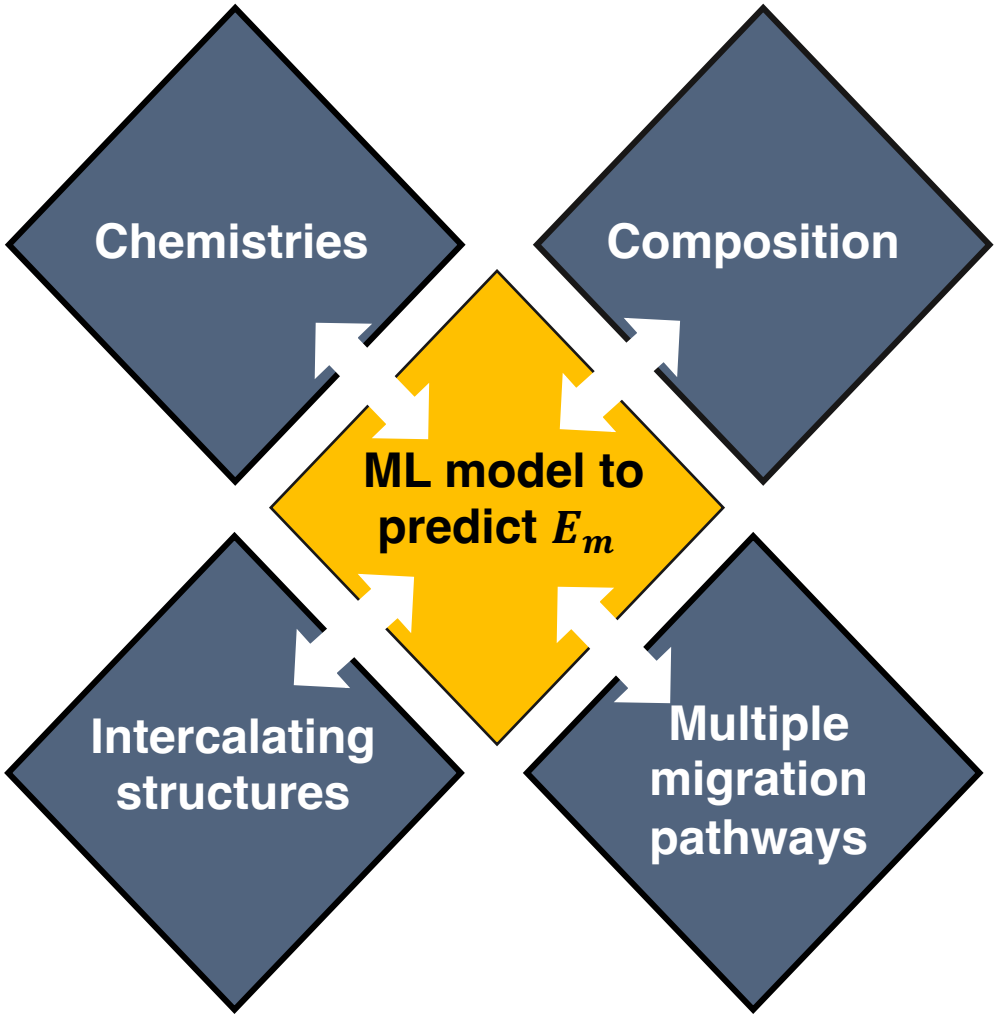
Migration barrier: dominant factor determining diffusivity

Experimental measurements of E_m: variable-temperature impedance spectroscopy, variable temperature nuclear magnetic resonance, etc.
Computational predictions of E_m: ab initio molecular dynamics, nudged elastic band (NEB) with density functional theory (DFT)

How to accurately and swiftly predict migration barriers? Use machine learning?

van der Ven et al. **Ann. Rev. Mater. Res.** *48*, 27-55 (2018)

What we need for fast and accurate E_m predictions?



How accurate are first-principles calculations in predicting E_m?

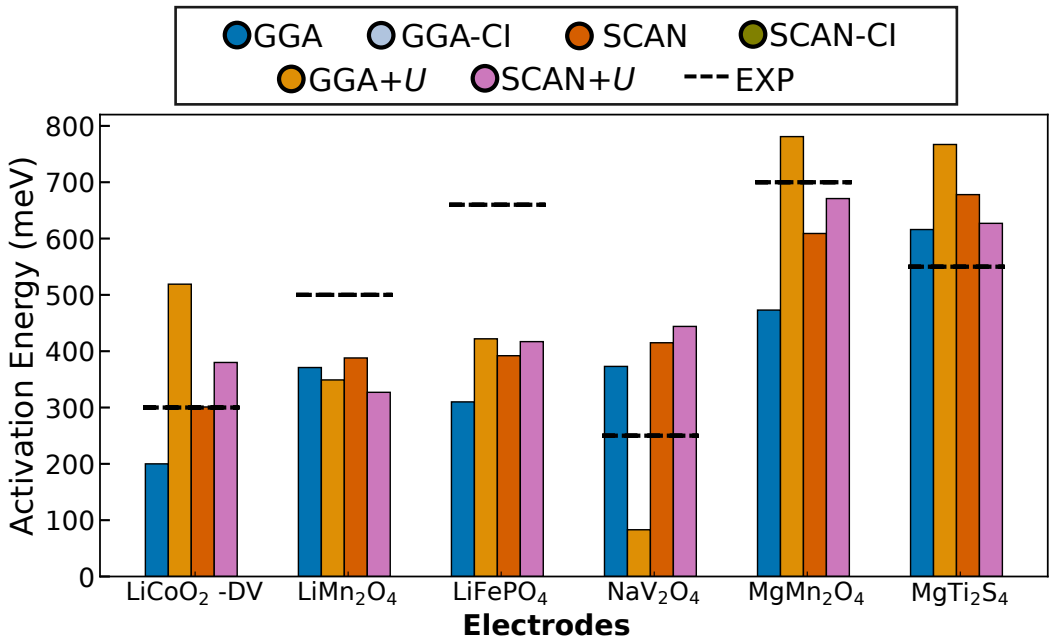
Is there a dataset of E_m values belonging to a diverse set of structures, chemistries, compositions, and migration pathways?

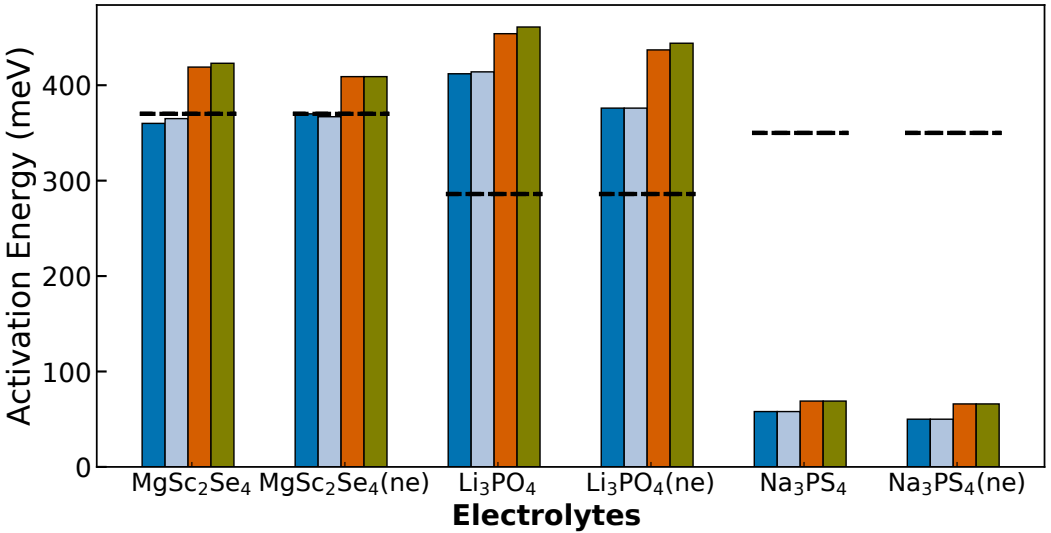
If there is a dataset, how do we construct a generalizable model that accurately predicts E_m?

- Using transfer learning

Accuracy

On average accuracy is in the 140-180 meV range





MAE of SCAN (140 meV): lowest

MAE of GGA: 178 meV

Qualitative trends: reliable with GGA

SCAN: computational cost+ convergence

60 meV in E_m: ~1 order of magnitude D at 300 K for micron-sized particles

125 meV in E_m: same D at 300 K, one order change in particle size

SCAN: strongly constrained and appropriately normed

GGA: Generalized gradient approximation

U: Hubbard *U* correction

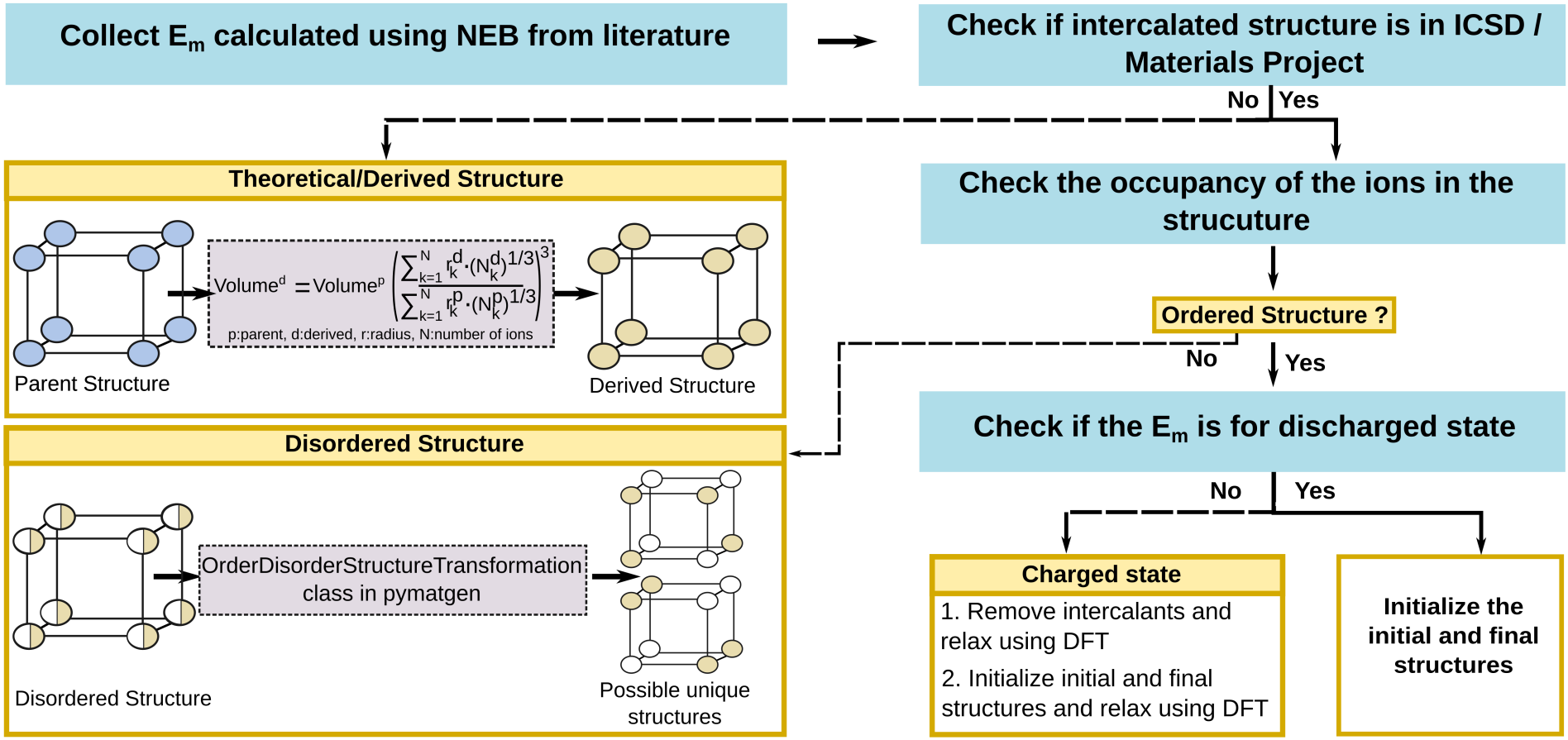
CI: Climbing image; EXP: Experimental

ne: Background charge; DV: divacancy

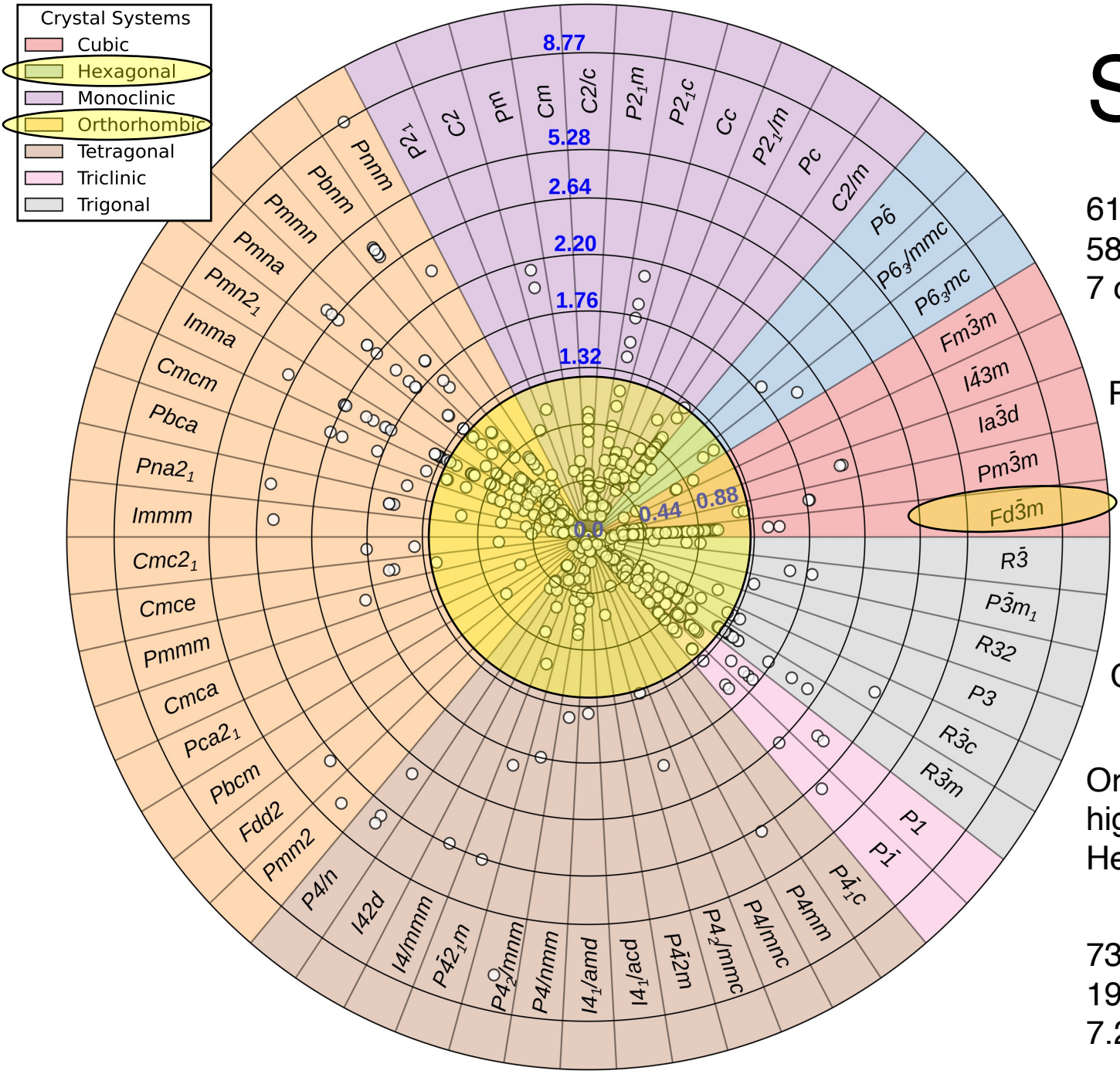
Devi et al., **npj Comput. Mater.** *8*, 160 (2022)

Dataset

Workflow to generate a curated dataset of E_m



- Consider papers that reported unambiguously the methods and the structure used
- Generalized gradient approximation or its Hubbard *U* preferred as the functional
- Well labelled images of E_m and/or the minimum energy pathway (MEP)



Statistics

619 datapoints spanning
58 space groups across
7 crystal systems

Range: 0.03 to 8.77 eV

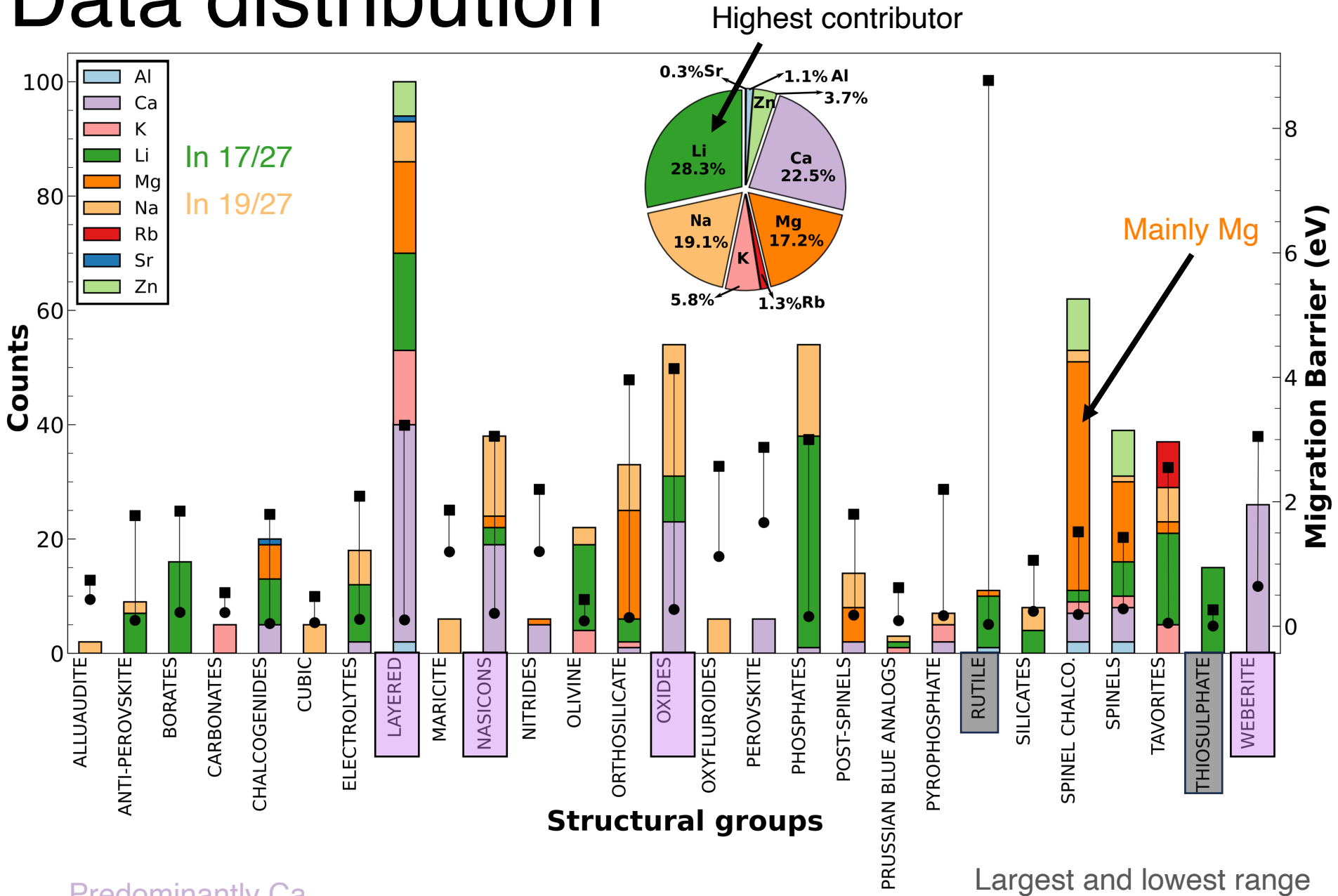
528 electrodes
91 electrolytes

Cubic spinels: 94 entries

Orthorhombic (206): highest
Hexagonal (6): lowest

73.4%: $E_m < 1$ eV
19.4%: 1 eV $< E_m <$ 2 eV
7.2%: $E_m > 2$ eV

Data distribution

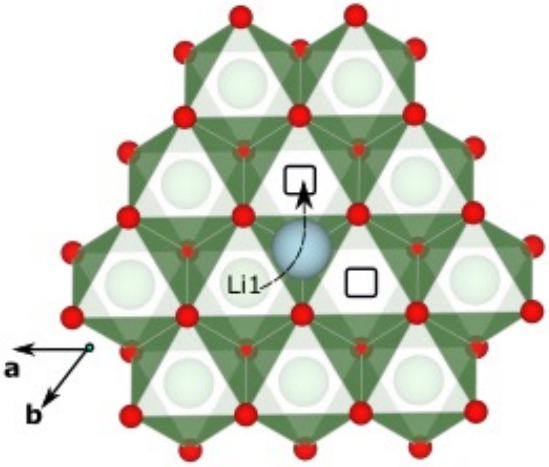


How to create generalizable models?
Transfer learning

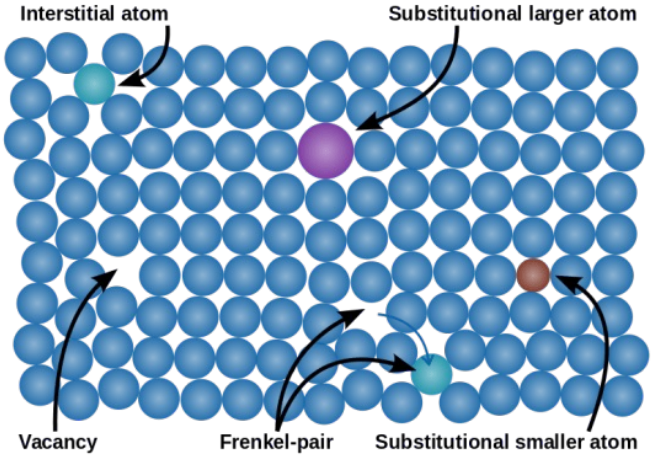
Materials science is data limited

Several key material properties that govern performance in applications have limited data

- 'Small' datasets (< 10^4 datapoints)
  - Ionic mobilities, defect formation energies, adsorption energies,…
- Limits application of deep learning (DL) frameworks

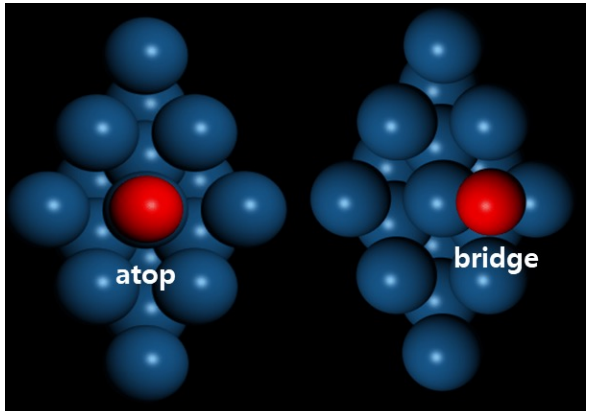


Devi et al., **npj Comput. Mater.** 2022



https://www.differencebetween.com/difference-between-point-defect-and-line-defect/

https://sites.psu.edu/



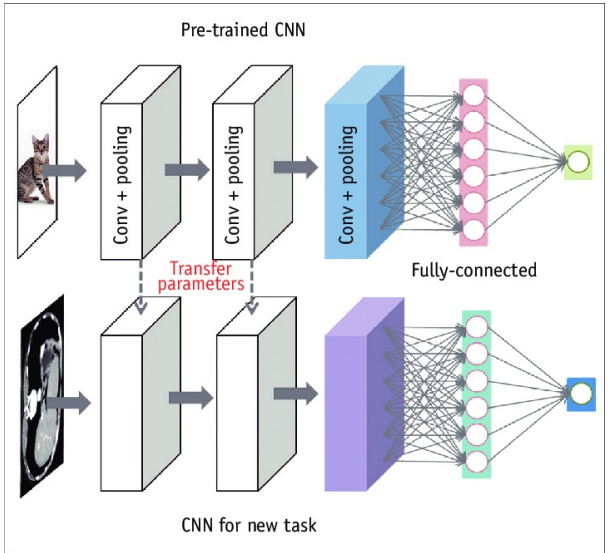
Materials science is data limited

Several key material properties that govern performance in applications have limited data

- 'Small' datasets (< 10^4 datapoints)
  - Ionic mobilities, defect formation energies, adsorption energies,…
- Limits application of deep learning (DL) frameworks

Transfer learning: efficiently use DL frameworks on small datasets

- Pre-train (**PT**) on 'large' dataset, fine-tune (**FT**) on 'small' dataset



Do et al., **Korean J. Radiol.** 2020

Materials science is data limited

Several key material properties that govern performance in applications have limited data

- 'Small' datasets ($< 10^4$ datapoints)
  - Ionic mobilities, defect formation energies, adsorption energies,…
- Limits application of deep learning (DL) frameworks

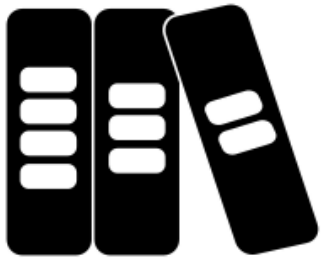
Transfer learning: efficiently use DL frameworks on small datasets

- Pre-train (**PT**) on 'large' dataset, fine-tune (**FT**) on 'small' dataset

How useful is transfer learning in materials science?

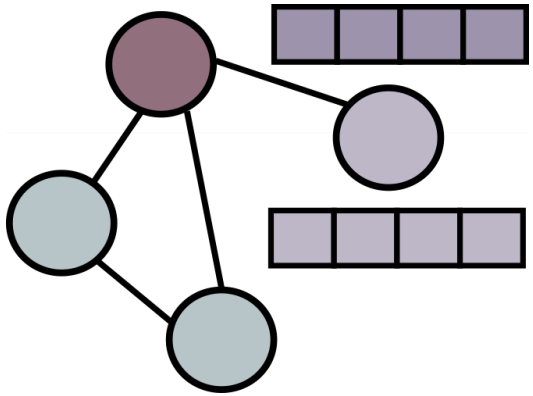
- Optimal ways to use?
- Ways to generate 'generalizable' models?

Handles to consider



Dataset(s)

- What, how, how many?



Architecture

- Graph neural network

Frozen

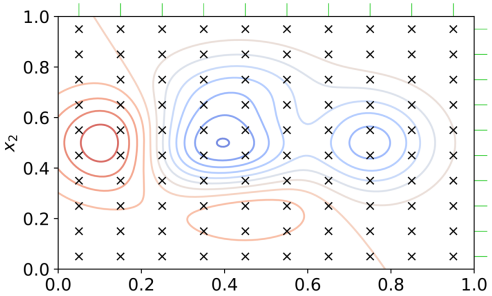
Unfrozen

Strategy

- FT techniques in pair-wise PT/FT models
- Multi-property PT (**MPT**) models

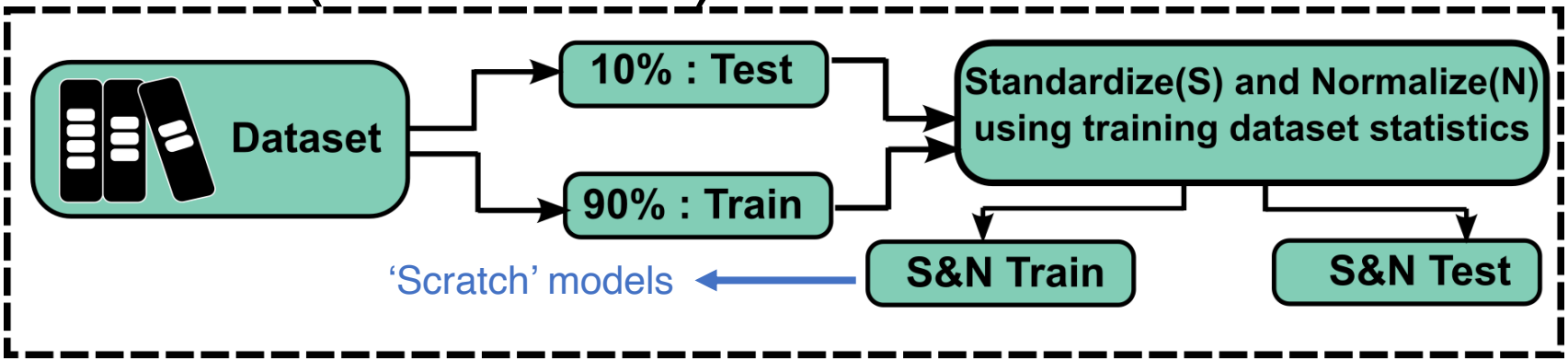
(Learning) Hyperparameters

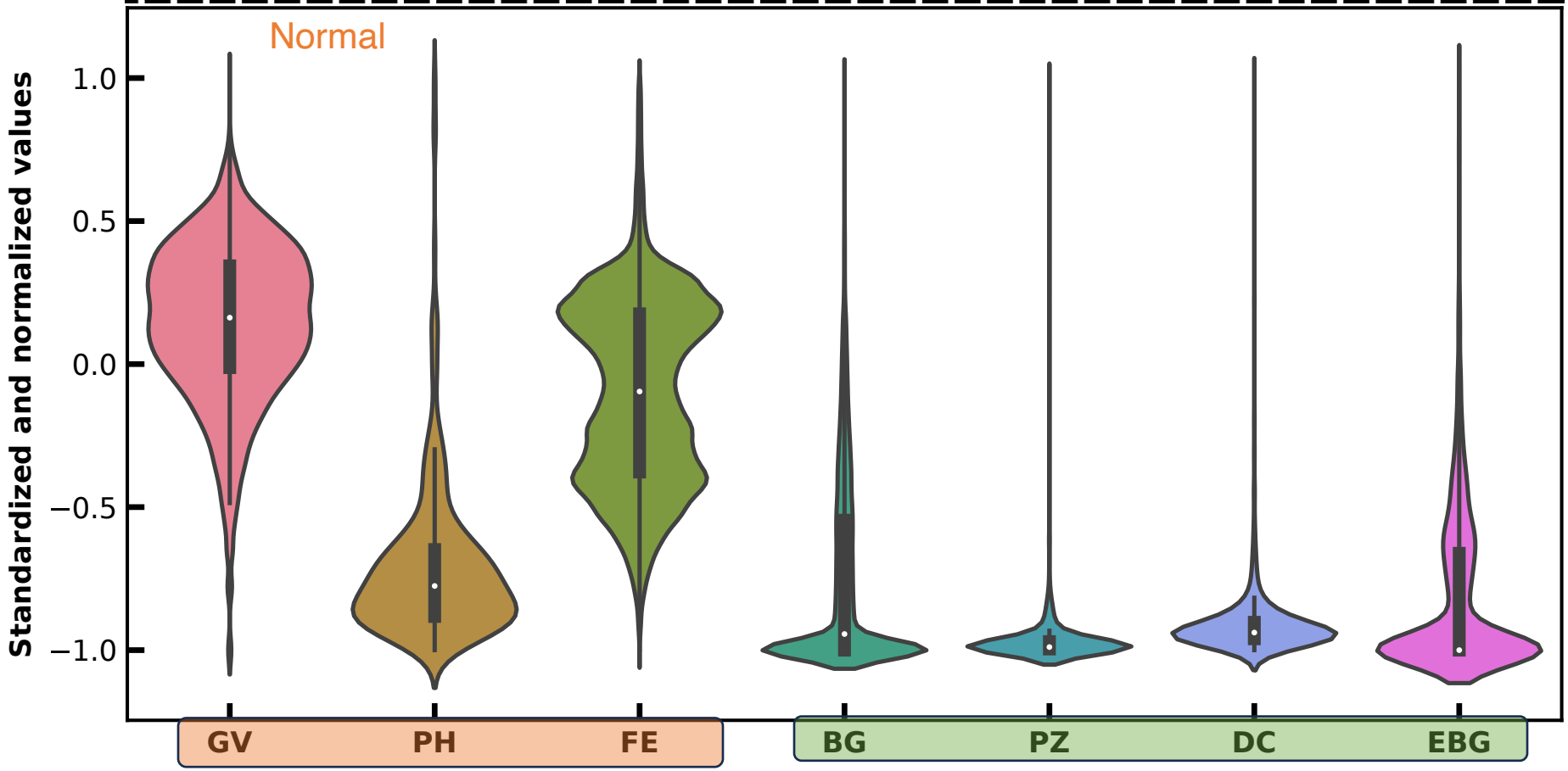
- Data sampling
- Learning rate
- Number of datapoints during PT, FT



https://en.wikipedia.org/

7 datasets (Matminer)

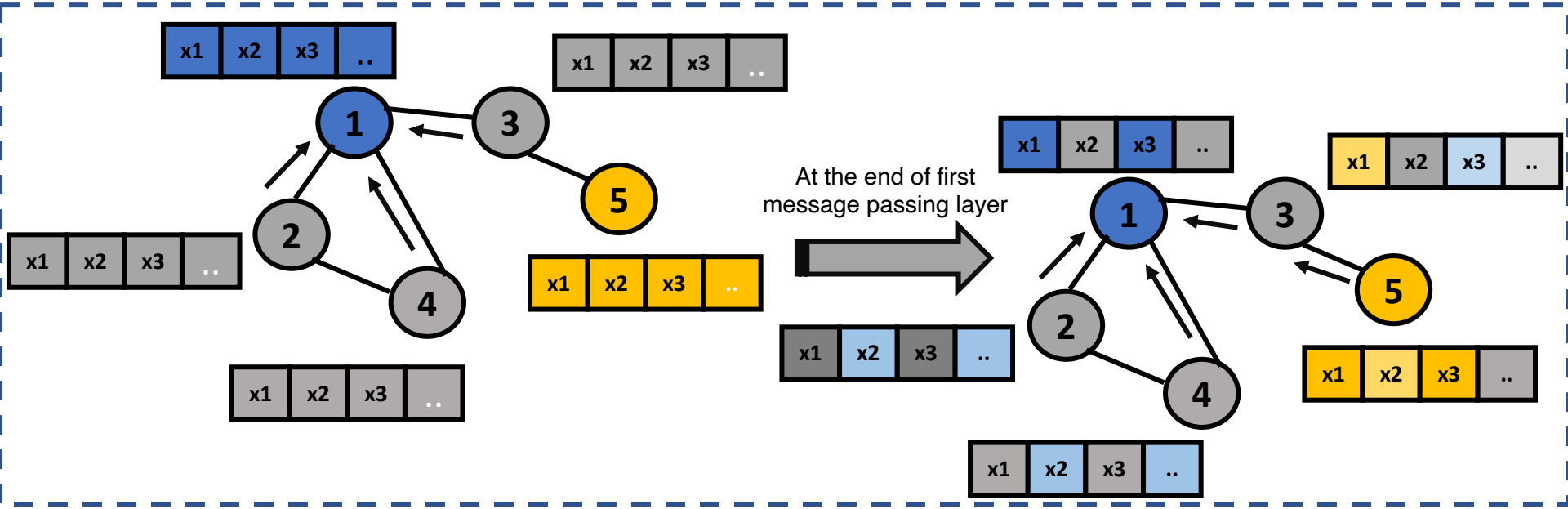


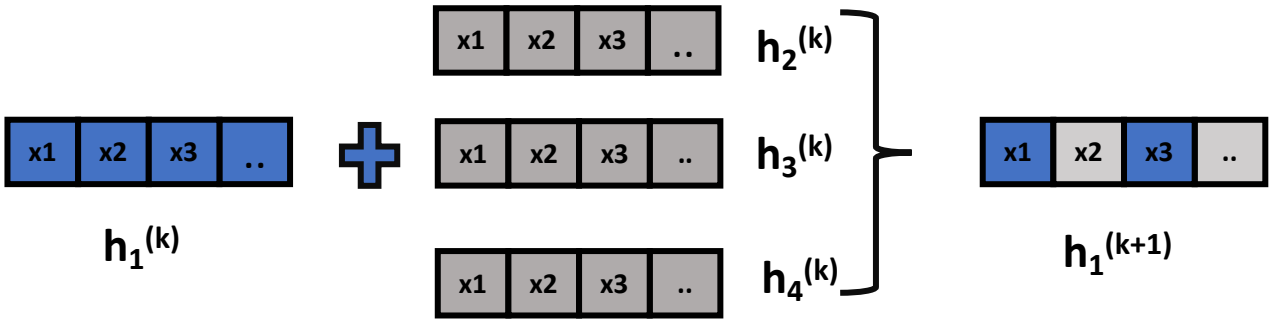


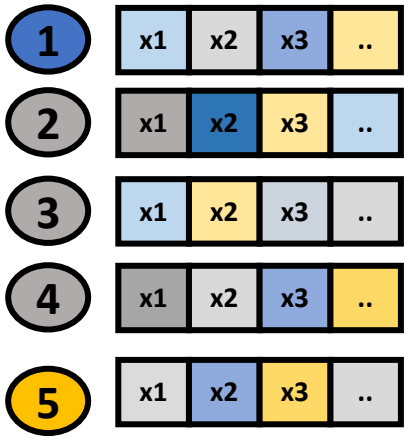
1. Ward *et al.*, **Comput. Mater. Sci.** *152*, 60–69 (2018).

Base model architecture: graph neural networks

Specific architecture: ‘ALIGNN’



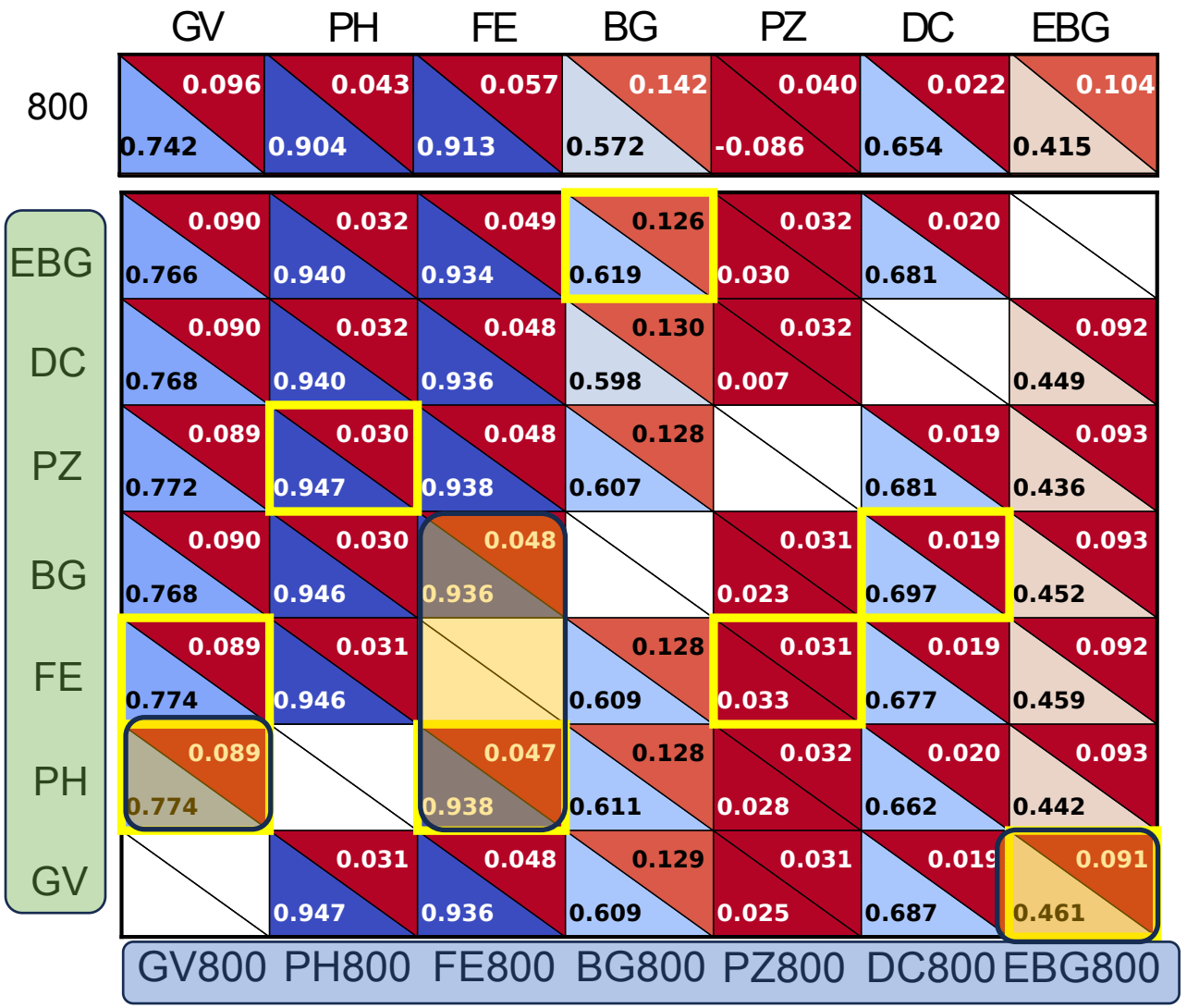




After multiple message passing layers

ALIGNN: Choudhary and DeCost, **npj Comput. Mater.** *7*, 185 (2021)

7×6 combinations of pair-wise models



Pair-wise models: better than scratch

- Average increase in R^2: 25%
- Average decrease in MAE: 16%

Best models: GV, PH, FE ($R^2 > 0.75$)

Average models: BG, DC, EBG

Specific PT property: little influence on FT

No symmetry

FT dataset+size

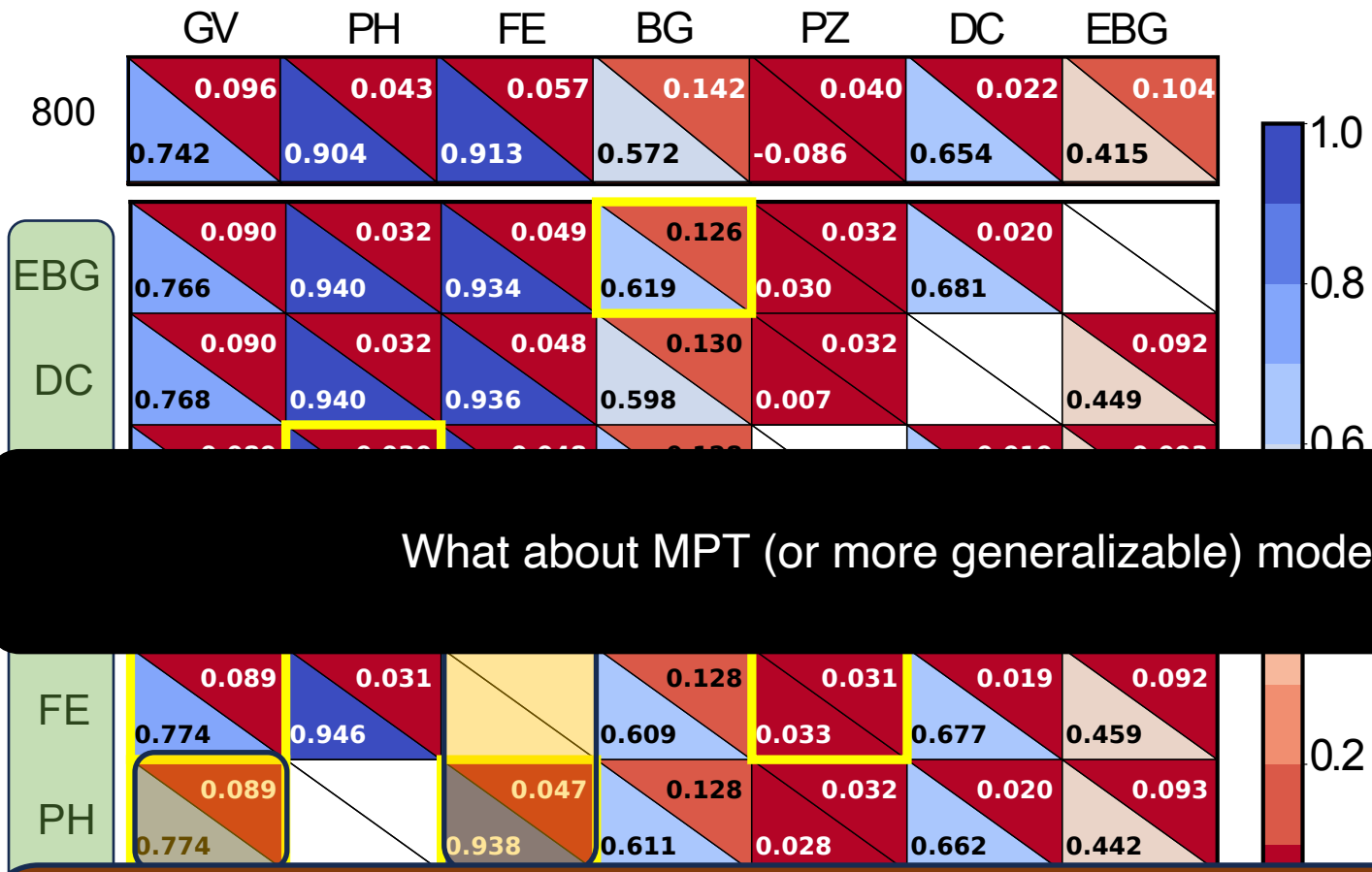
PT dataset (941)

Best model

Test scores

GV: Shear modulus; PH: Phonons: FE: Formation energy; BG: Band gap
PZ: Piezoelectric modulus; DC: Dielectric constant; EBG: Experimental band gap

7×6 combinations of pair-wise models



Pair-wise models: better than scratch

- Average increase in R^2: 25%
- Average decrease in MAE: 16%

Best models: GV, PH, FE ($R^2 > 0.75$)

What about MPT (or more generalizable) models?

property: little influence on FT

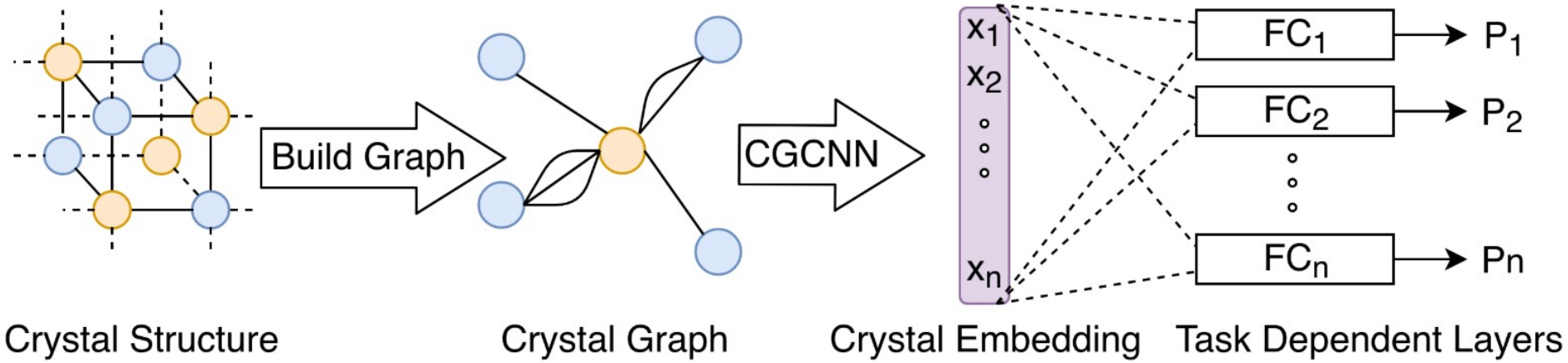
No symmetry

At capped dataset size, specific PT property is a weak handle; Normal distribution is better

Pair-wise transfer learning has significant utility

MPT: (Beta) Generalized models

Inspiration from literature: multi-task crystal graph convolutional neural network[1]



MPT models can generalize dependence of several properties on the structure

- Build cumulative dataset: 132,270 points
  - Remove overlaps
- Add task-dependent prediction heads with a one-hot encoded vector
  - Presence/absence of property
- Modify loss function

$$\mathcal{L} = \frac{1}{N}\sum_{i=1}^{N} |y_p^i - y_t^i| \delta^i$$

- PT on all (but one) property, FT on one property

1. Sanyal et al., **arXiv** 1811.05660 (2018)

MPT: better out-of-domain than PT-FT

Band gap of 2D materials (1,103 datapoints) from JARVIS-DFT[1]

Model	Test R^2	Test MAE
Scratch	0.635	0.148
MPT (all seven datasets)	**0.671**	**0.125**
FE(100K)	0.670	0.127
BG(50K)	0.617	0.138
PH(1256)	0.628	0.145
GV(10,987)	0.626	0.143
EBG(2,481)	0.619	0.143

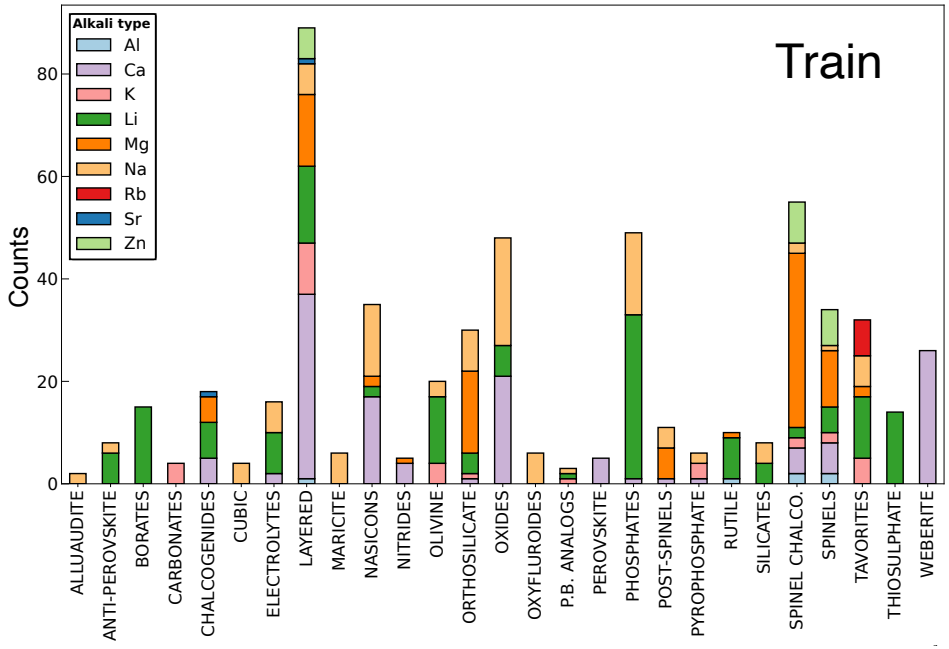
On average, MPT is 6% and 10% better on R^2 and MAE than PT-FT
Closest performer to MPT is FE: largest dataset within MPT
MPT models: may generalize quite well with more properties

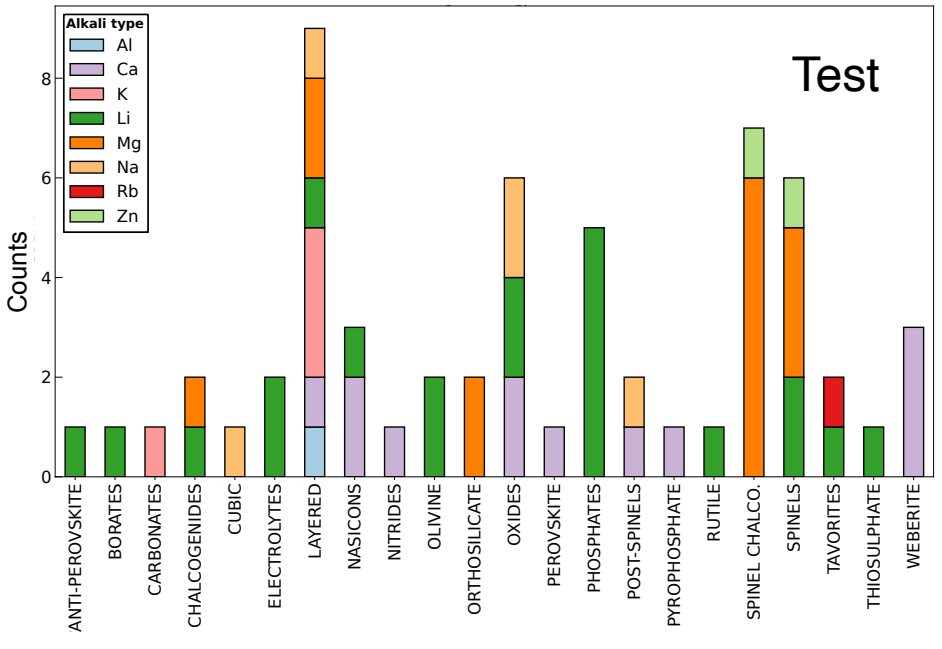
FE: Formation energy; BG: Band gap; GV: Shear modulus;
EBG: Experimental band gap

1. Choudhary et al. **npj Comput. Mater.** *6*, 173 (2020)

How to predict migration barriers?
Using the MPT model

Careful split of dataset



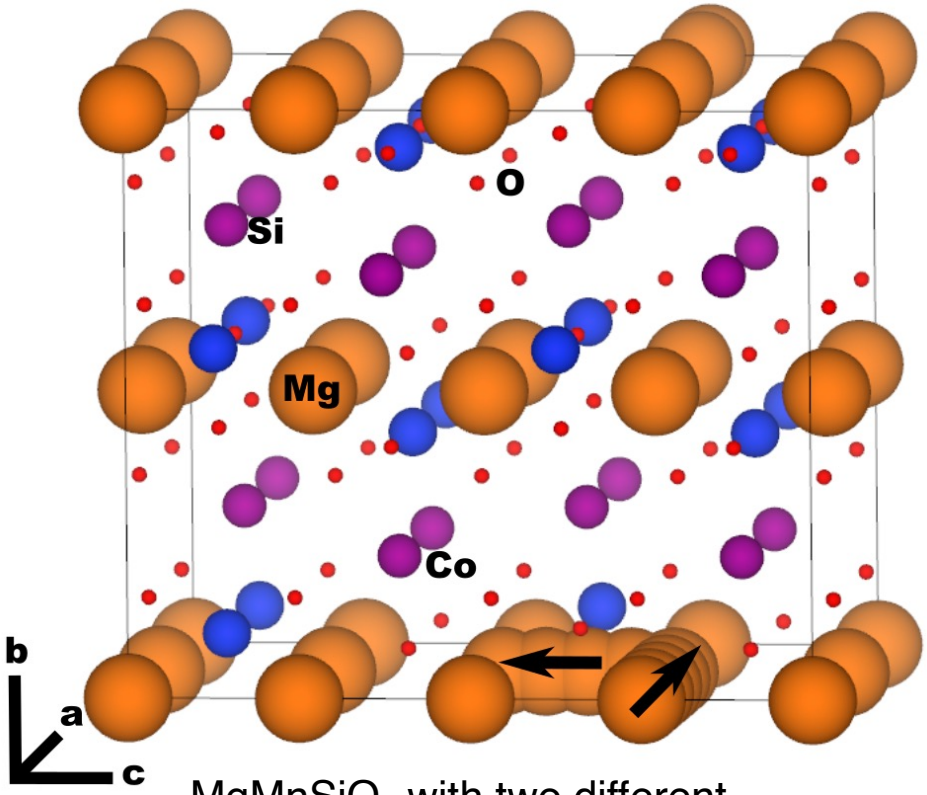


Structural groups

559 datapoints in train; 60 in test

- Similar distribution in train and test
- Prevent unfair penalization: 1 test datapoint for groups contributing 1-2%
- Random sampling within each structural group

Distinguish multiple paths in a structure: use modified model architectures

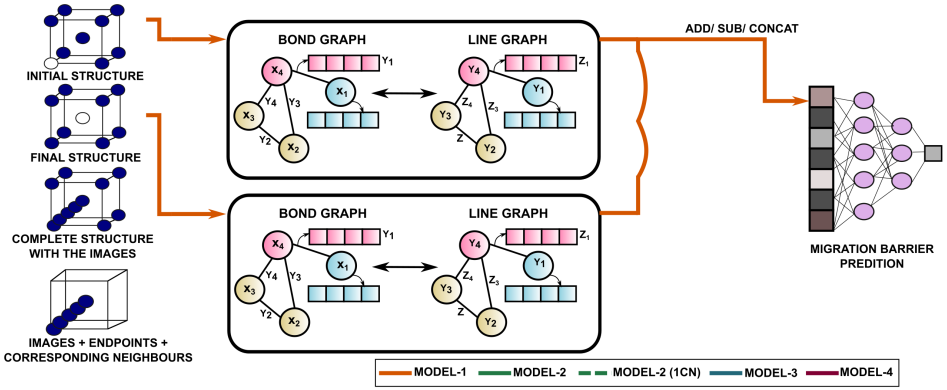


$MgMnSiO_4$ with two different pathways in the same structure

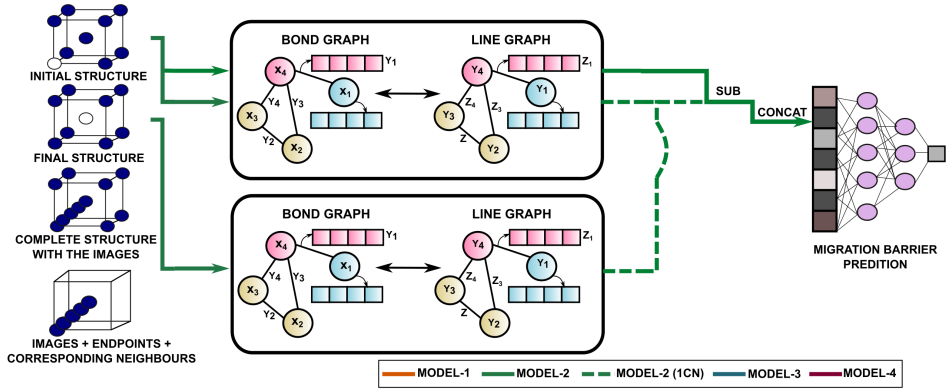
PT: MPT, FT: E_m

Distinguish multiple paths in a structure: use modified model architectures

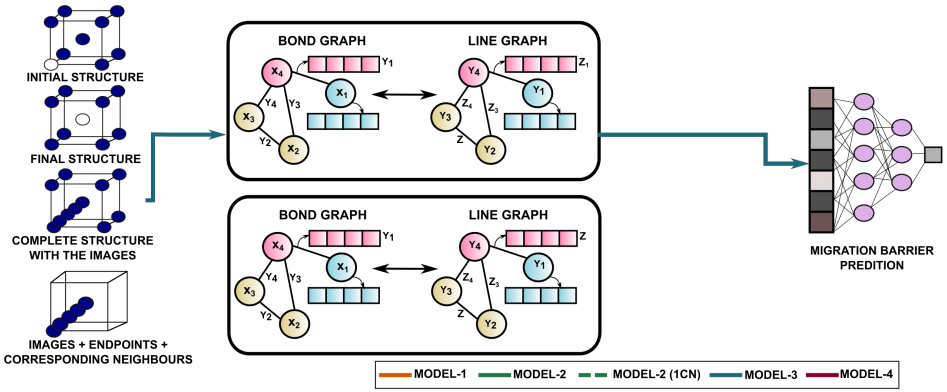
Model 1: Take initial and final as input



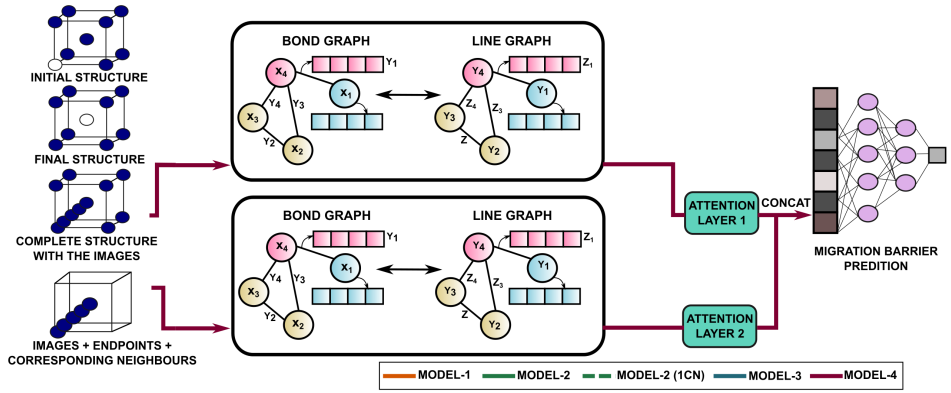
Model 2: Take initial and delta as input



Model 3: Band as input

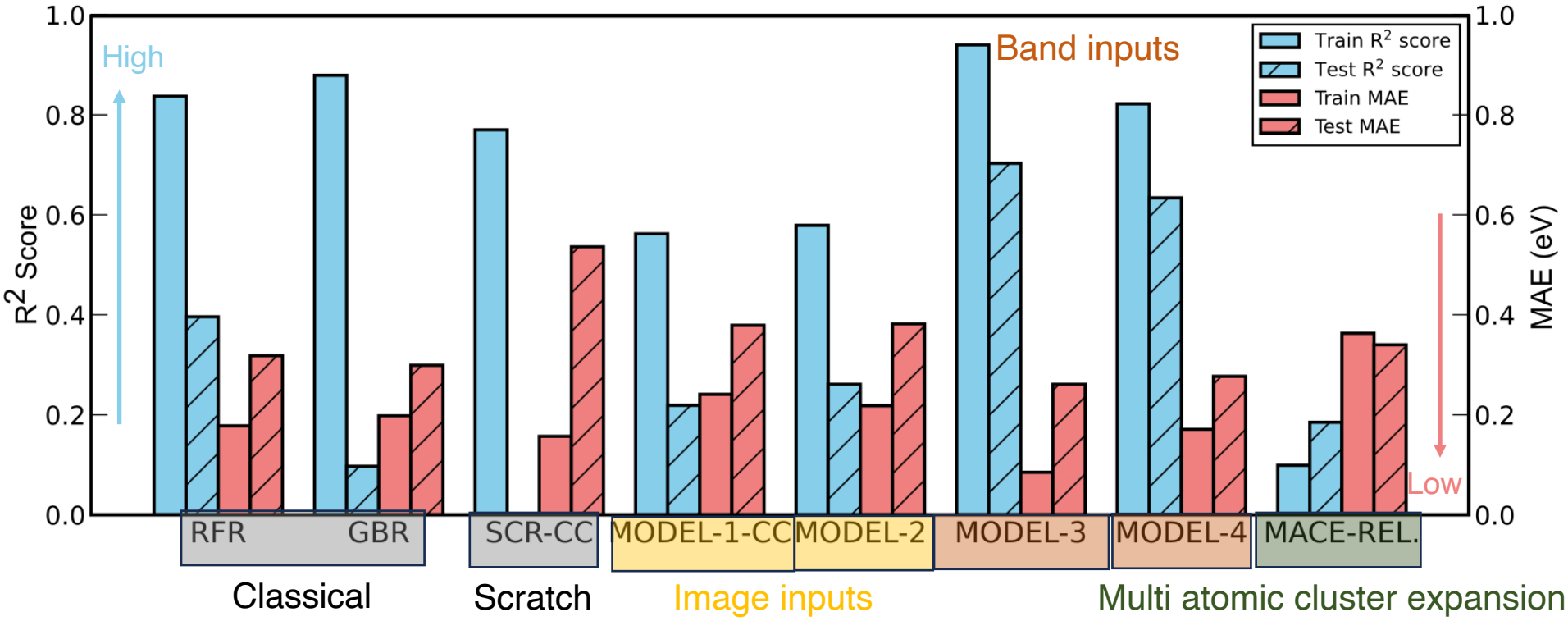


Model 4: Band with attention



PT: MPT, FT: E_m

Model-3: best performing model

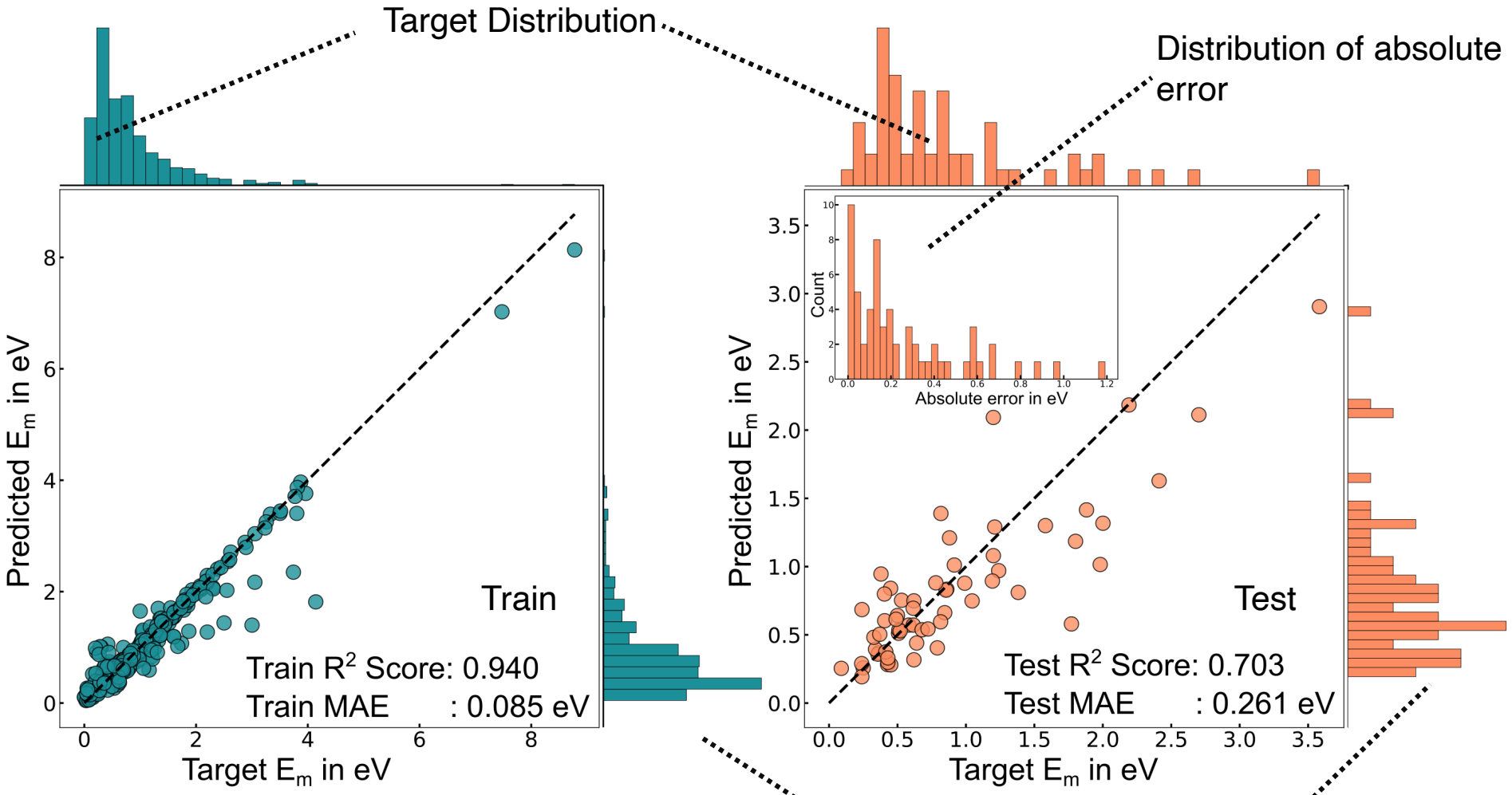


Model-3 performance (test set statistics, hashed bars):

- R^2: 0.703, MAE: 0.261 eV (~1.4-1.8× DFT error)
- Better scores than classical and scratch models
- Versus MACE: slightly better MAEs, significantly better R^2
- Identifies multiple migration pathways in a structure (band inputs > image inputs)
- Model-4 is closest performer to Model-3

RFR: Random forest regressor; GBR: Gradient boosted regressor; SCR: Scratch; CC: Concatenate; REL: Relax

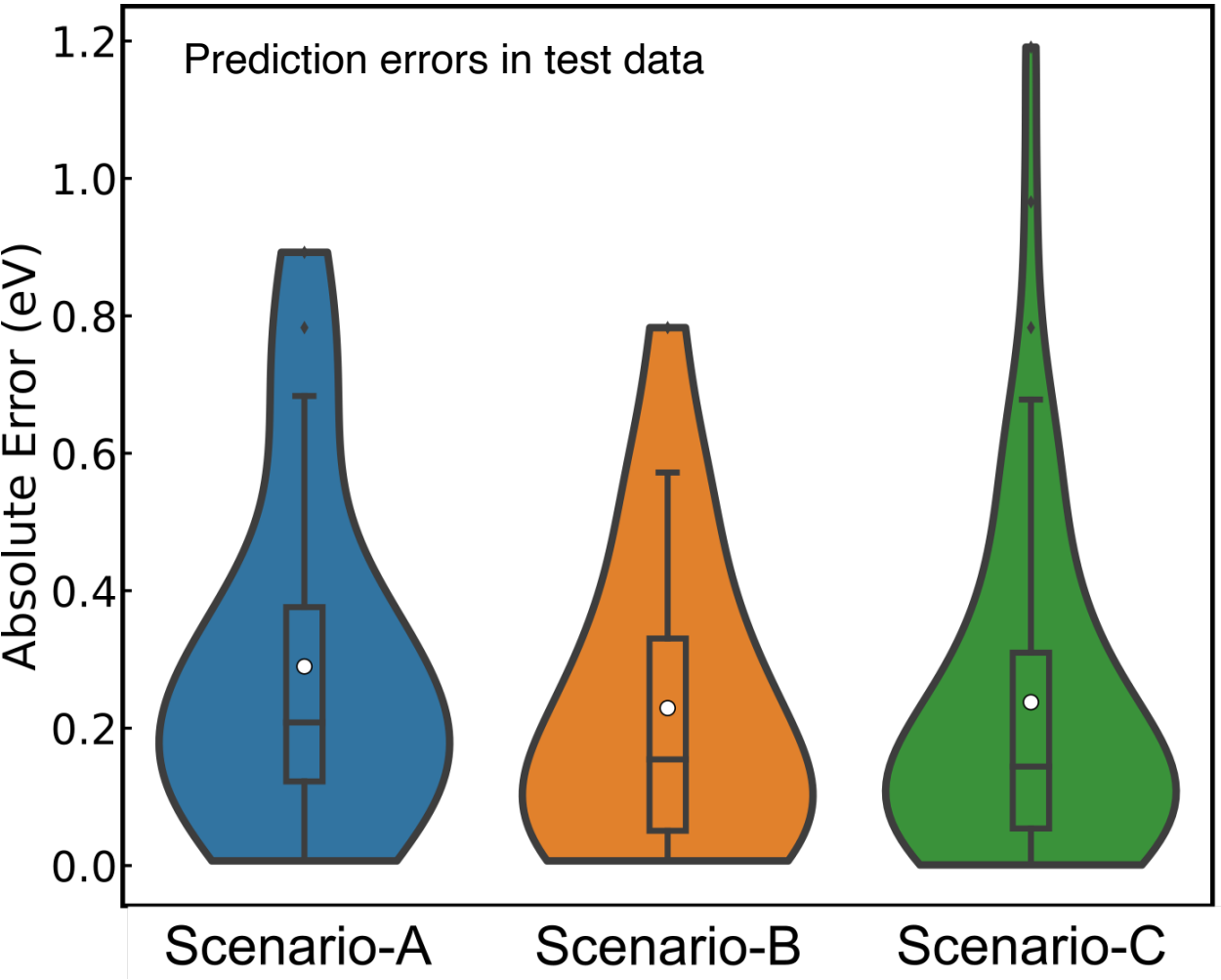
Model-3 performance



70% test dataset has prediction errors < 0.3 eV

- 32% high accuracy (< 0.1 eV error)
- 38% moderate accuracy (0.1-0.3 eV error)

How does Model-3 generalize?



Generalization across migration pathways in a structure

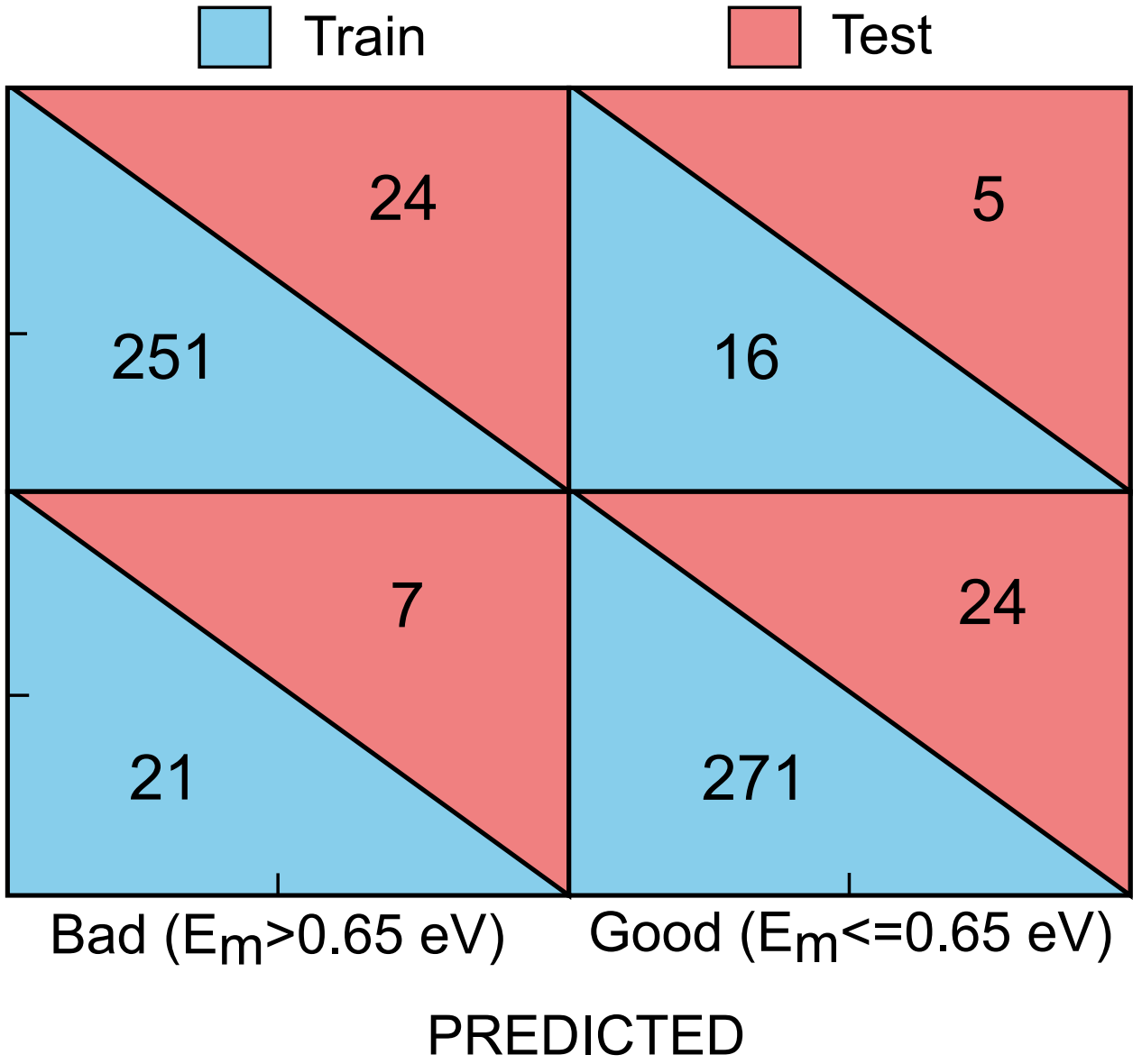
Generalization across intercalant composition (charged vs. discharged)

Generalization across chemistries (different intercalant/transition metal/anion, same structure)

Lower mean/median errors for Model-3 in scenarios B & C: better generalization across composition and chemistry compared to migration pathways!

- Can be used as a screening tool

Model 3 as a classifier



Threshold: 0.65 eV
- Nano-sized cathode
- C/2 rate, 300 K

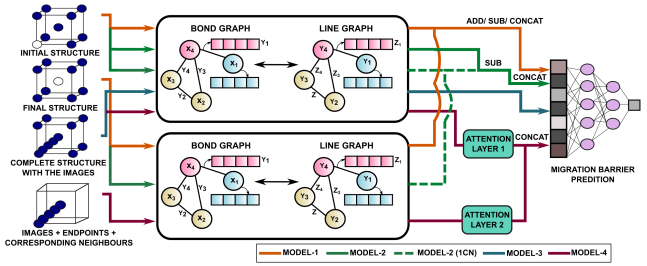
Accuracy: 80%

Precision:
82.8% (good conductors)
77.4% (bad conductors)

Model-3 can be good classifier!
- Especially for identifying good conductors

Conclusions

- Predicting E_m: critical for discovery of new materials with 'high' intercalant mobilities
- Accuracy: DFT gets qualitative trends right, reasonably close to experiments
- Dataset: curated 619 datapoints of calculated E_m
- Generalizable model to predict E_m
  - Leverage transfer learning, modified graph networks



"A literature-derived dataset of migration barriers for quantifying ionic transport in battery materials", R. Devi, A. Balasubramanian, K.T. Butler, and G. Sai Gautam, **arXiv**, 2508.06459 (2025) [under review].

"Optimal pre-train/fine-tune strategies for accurate material property predictions", R. Devi, K.T. Butler, and G. Sai Gautam, npj Comput. Mater. 10, 300 (2024).

"Leveraging transfer learning for accurate estimation of ionic migration barriers in battery materials ", R. Devi, K.T. Butler, and G. Sai Gautam, arXiv, 2508.06436 (2025). [under review]

saigautamg@iisc.ac.in; https://sai-mat-group.github.io; https://github.com/sai-mat-group