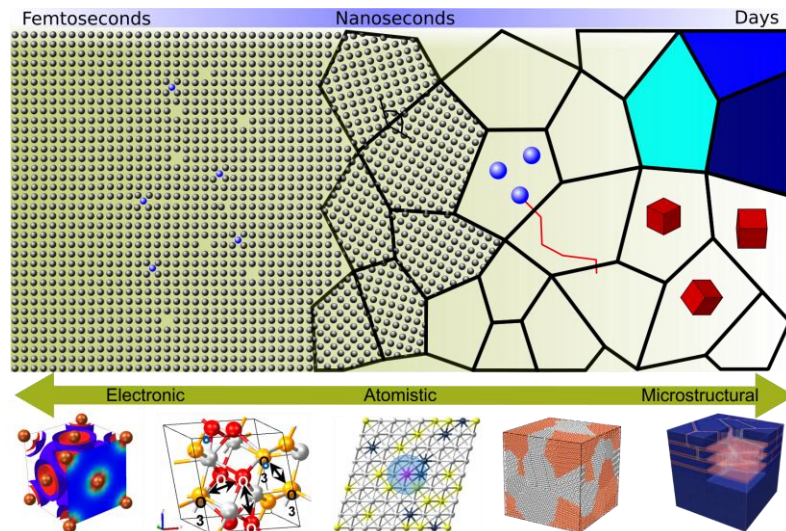
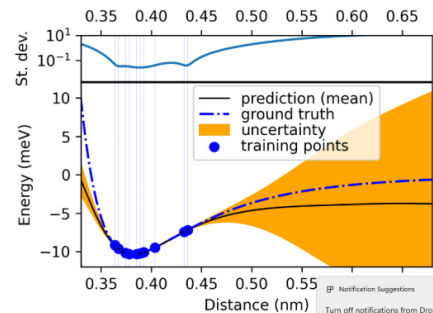
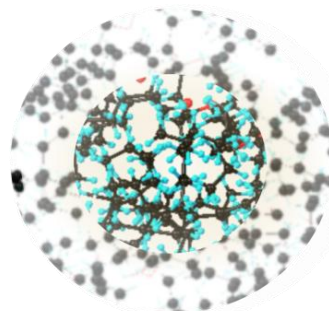


Machine and Human Learning of Materials

Leonid Kahle

Research Scientist, Materials Design S.A.R.L

17 Oct 2025



Human Learning of Materials

Human Learning Timeline



Materials Design

Legacy: Founded 1998, serving 700+ institutions worldwide.

Our Mission: Creating Engineering Value from Materials Simulations.

Offerings: MedeA® software, support, consulting, and contract research.

Presence: Headquartered in San Diego, USA, and Paris, Europe.

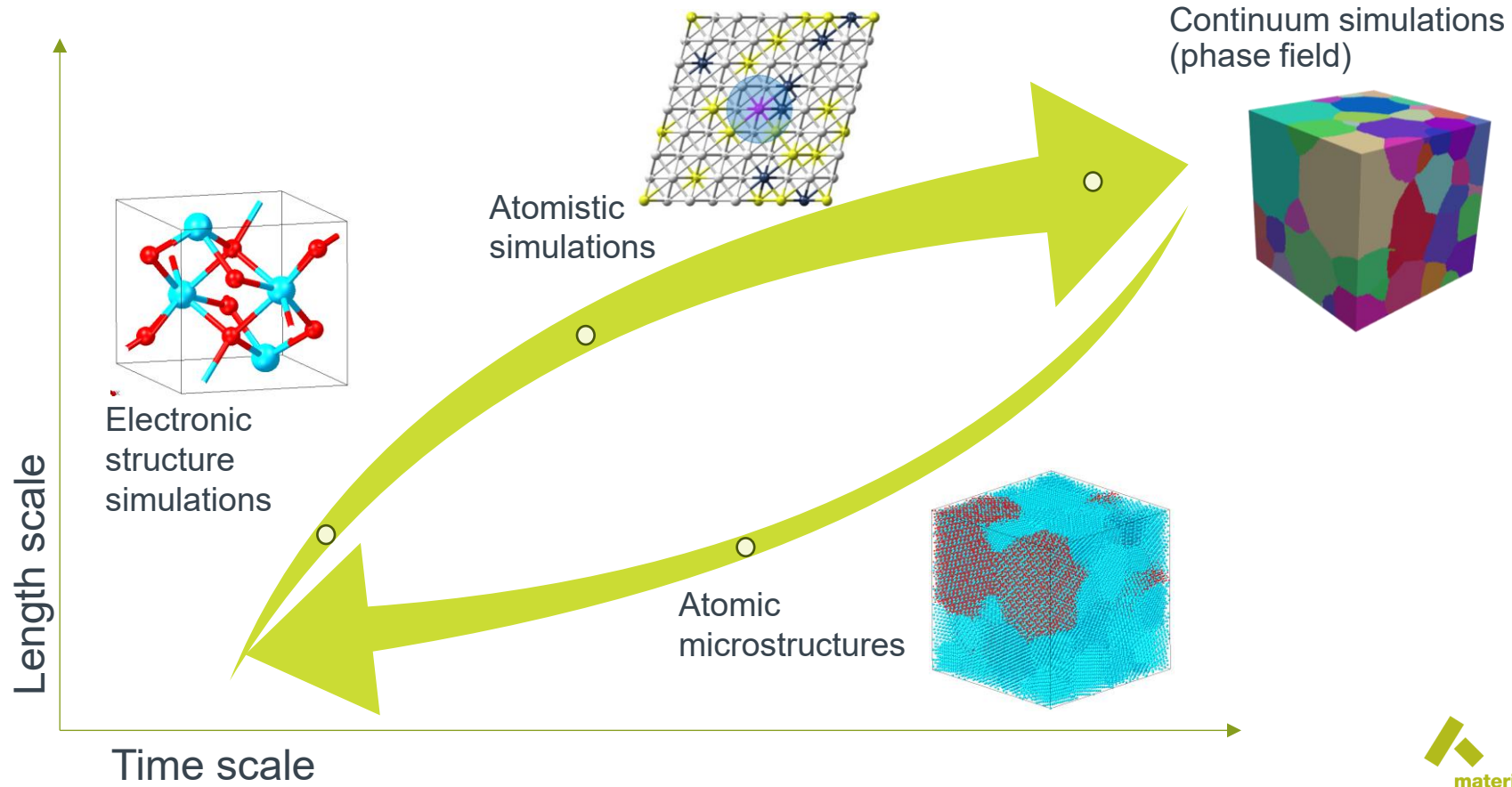
Partnerships: Collaborating with technology and business partners globally.

Expertise: Computational materials science, chemistry, chemical engineering.



www.materialsdesign.com

Accessing different time and length scales



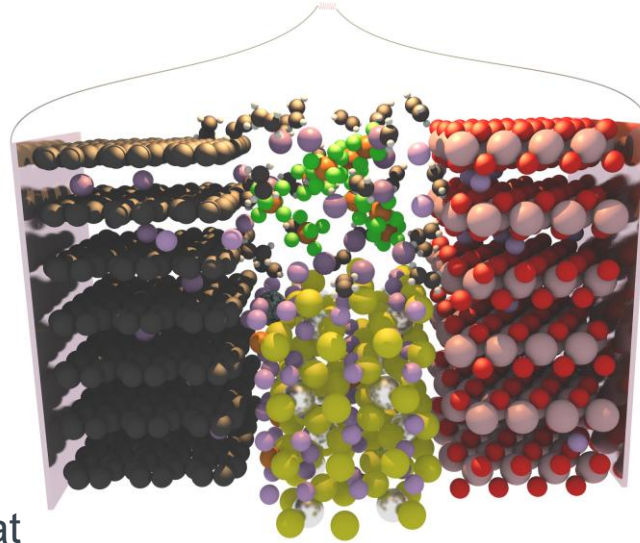
Machine Learning of Materials

Material Science at the atomic scale

Task: finding a better battery material!

How much Li fits into the anode?

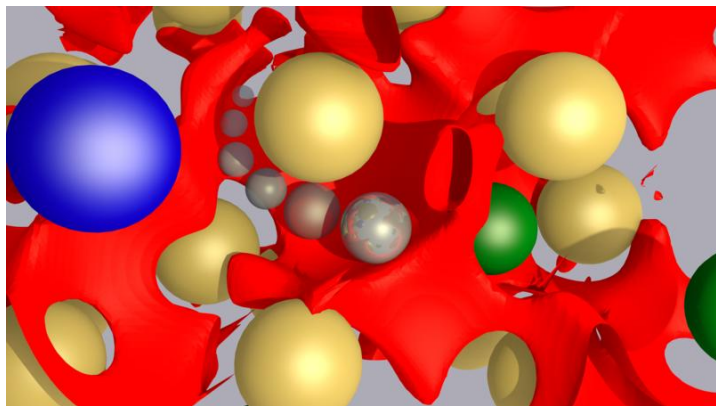
Which materials are stable at such voltages?



What is the voltage of an anode/cathode?

What are good materials for electrolytes with high electronic resistance and high ionic conductivity?

What is the ground truth?

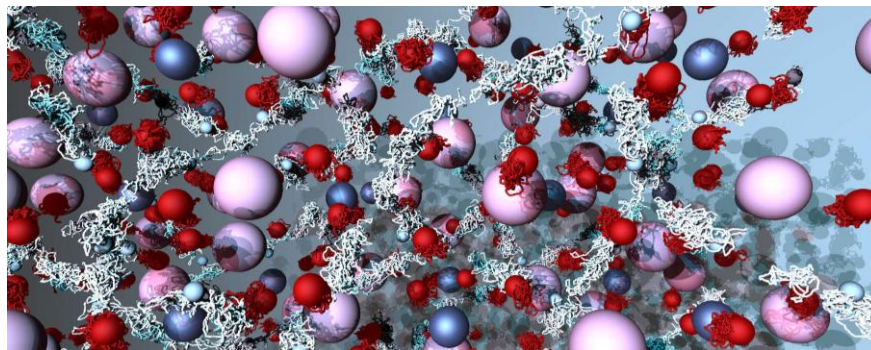


Electronic properties: $\hat{H}\Psi = E \Psi$

To calculate electronic properties from first principles, we need to solve (approximations to) the Schrödinger's equation.

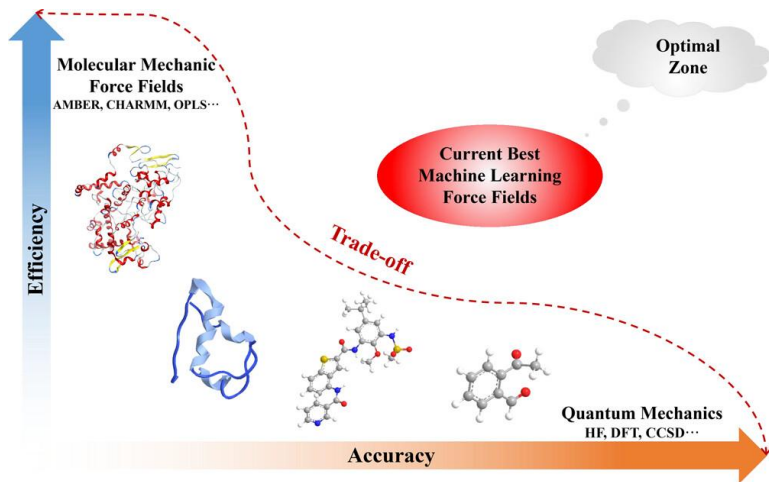
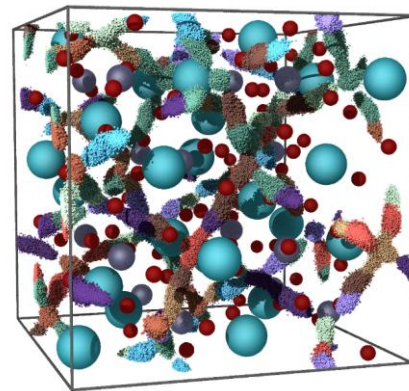
Dynamic properties: $F = ma$

To calculate dynamic properties, we integrate the forces on atoms over time.



Machine learning in dynamic simulations

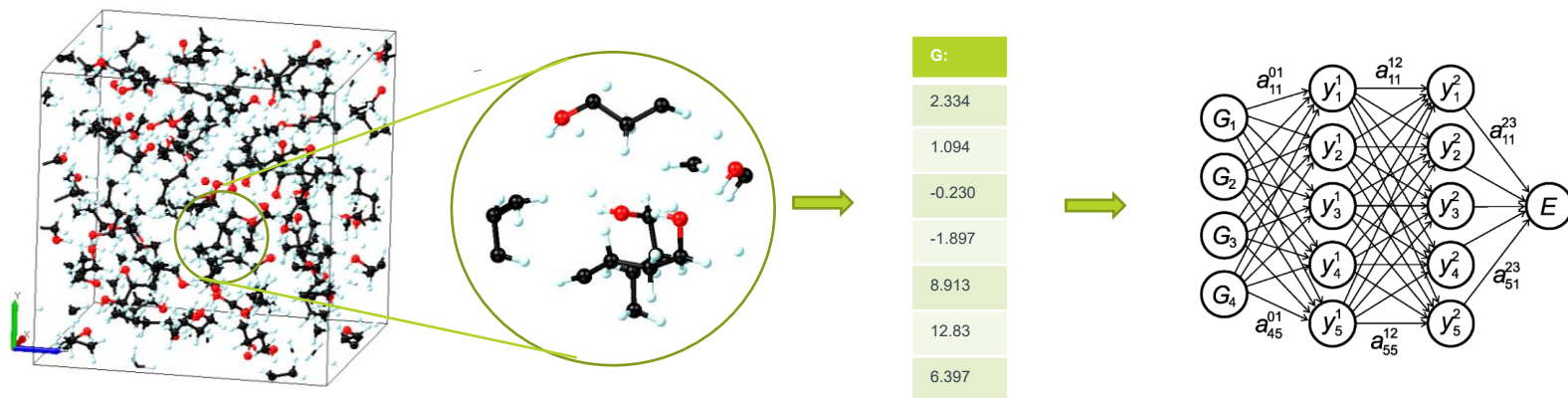
Machine learning can help analyze simulations with “classical” methods, e.g., using unsupervised learning of the atomic environment [1]



Machine learning interatomic potentials can drive dynamic simulation combining high accuracy of advanced QM methods at the cost of empirical potentials.

[1] Kahle et al., Unsupervised landmark analysis for jump detection in molecular dynamics simulations, Phys. Rev. Mat. 3, 055404 (2019)

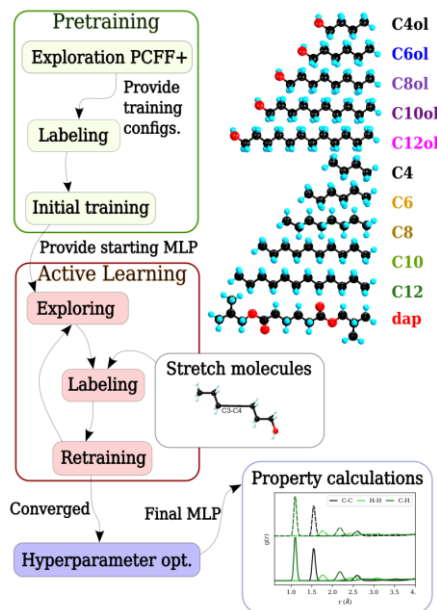
Obtaining an Atomistic ML Potential for Organics



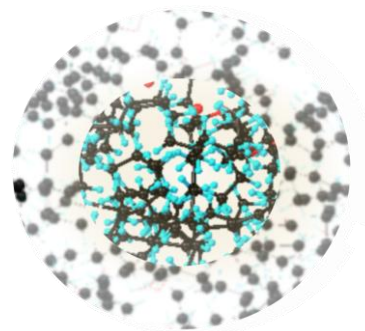
Challenge in organic liquids are the different contributions:

- Complex, strong, short-range covalent bonding.
- Simple, weaker, long-range vdW and electrostatic interactions.
- Very complex environments.

Learning about alcohol!

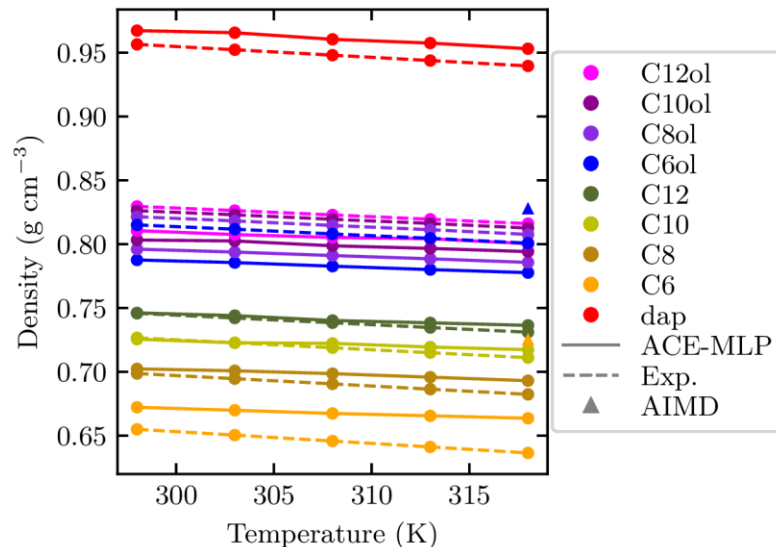
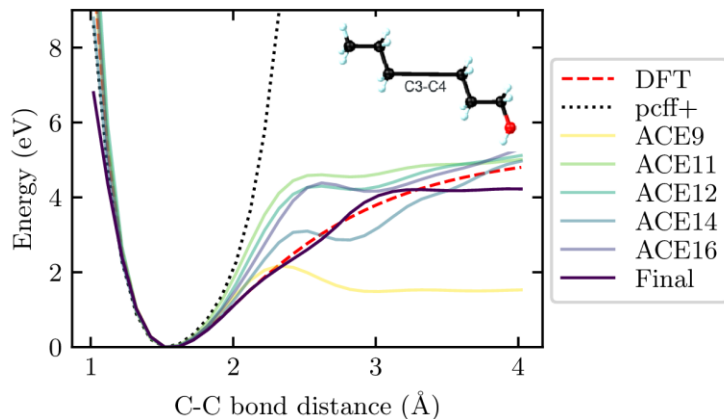


- An industry client wants to understand heat transfer between metals and alcohols
- No empirical model → Train a machine learning interatomic potential (MLIP)
- Complex short intramolecular and simple long-range intermolecular interactions -> A dual cutoff MLIP
- Active learning for training set generation.

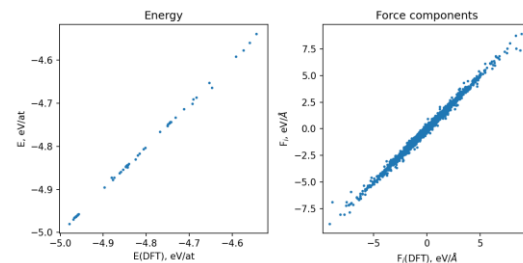


Kahle et al., A dual-cutoff machine-learned potential for condensed organic systems obtained via uncertainty-guided active learning, Phys. Chem. Chem. Phys., 2024, 26, 22665

Learning about alcohol!



- Subsequent iterations of active learning improve results.
- How to add training data is the key!
- Good final reproduction of density (not trained!)



Kahle et al., A dual-cutoff machine-learned potential for condensed organic systems obtained via uncertainty-guided active learning, Phys. Chem. Chem. Phys., 2024, 26, 22665

Predicting Uncertainty of Deep Models

- ML predictors extrapolate badly (hallucinations)

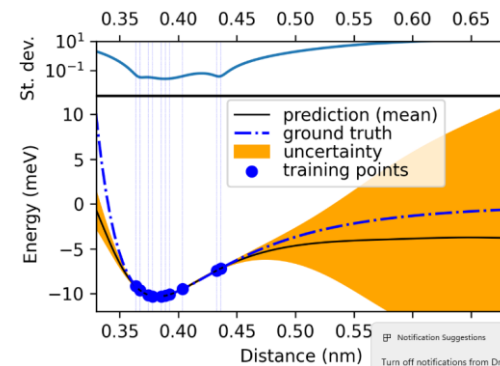
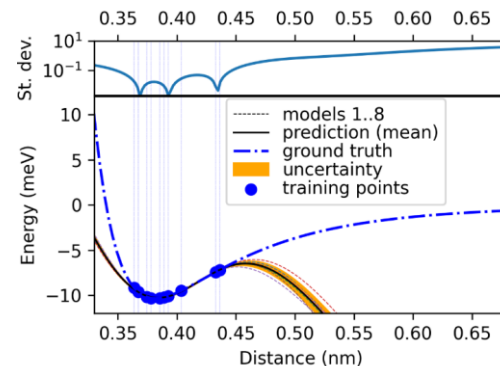
Name 5 fruits with 2 times the letter R

Here are **5 fruits** that each contain the letter **R** twice:

1. Cherry
2. Currant
3. Blackberry
4. Raspberry
5. Strawberry

- The same can happen with MLIPs, *out-of-distribution* predictions will be wrong – and we don't know!
- How can we estimate deep NN uncertainty?

Kahle and Zipoli., Quality of uncertainty estimates from neural network potential ensembles, PRE 105, 015311 (2022)



Making predictions...

Trends and observations, predictions and hallucinations



Trends and observations:

- A new ML model every few months.
- Foundational MLIPs are coming.
- Do we have the right performance metrics ?
- Data quality & quantity is a challenge.

Full Test Set Unique Prototypes 10k Most Stable

Click on column headers to sort table rows

☒ Compliant models ☒ Non-compliant models ☐ Energy-only models ☒ Heatmap Columns

Model	CPS	Acc	F1	DAF	Prec	MAE	R ²	K _{sim}	RMSD	Training Set	Params	Targets	Date Added	Links	r _{test}	Org
eSEN-30M-OAM	0.888	0.977	0.925	6.069	0.928	0.018	0.866	0.170	0.061	6.0M (113M) OMat24+MP1+Alex	30.2M	EF ₅₀	2025-03-17	🔗 📄 📊 📈	6 Å	🏠
Nequip-OAM-L	0.870	0.967	0.893	5.823	0.890	0.022	0.865	0.166	0.065	6.0M (113M) OMat24+Alex+MP1	9.6M	EF ₅₀	2025-09-08	🔗 📄 📊 📈	6 Å	🏠
ORB v3	0.861	0.971	0.905	5.912	0.904	0.024	0.821	0.210	0.075	6.47M (133M) MP1+Alex+OMat24	25.5M	EF ₅₀	2025-04-05	🔗 📄 📊 📈	6 Å	🏠
SevenNet-MF-ompa	0.845	0.969	0.901	5.825	0.890	0.021	0.867	0.317	0.064	6.0M (113M) OMat24+Alex+MP1	25.7M	EF ₅₀	2025-03-13	🔗 📄 📊 📈	6 Å	🏠
Allegro-OAM-L	0.840	0.966	0.895	5.674	0.867	0.022	0.868	0.319	0.065	6.0M (113M) OMat24+Alex+MP1	9.7M	EF ₅₀	2025-09-08	🔗 📄 📊 📈	6 Å	🏠
GRACE-2L-OAM	0.837	0.963	0.880	5.774	0.883	0.023	0.862	0.294	0.067	6.0M (113M) OMat24+Alex+MP1	12.6M	EF ₅₀	2025-02-06	🔗 📄 📊 📈	6 Å	🏠
DPA-3.1-3M-FT	0.802	0.963	0.884	5.667	0.866	0.023	0.869	0.469	0.069	163M OpenLAM	3.27M	EF ₅₀	2025-06-05	🔗 📄 📊 📈	6 Å	🏠
eSEN-30M-MP	0.797	0.946	0.831	5.260	0.804	0.033	0.822	0.340	0.075	146k (1.58M) MP1	30.1M	EF ₅₀	2025-03-17	🔗 📄 📊 📈	6 Å	🏠
MACE-MPA-0	0.795	0.954	0.852	5.582	0.853	0.028	0.842	0.412	0.073	3.37M (12M) MP1+Alex	9.06M	EF ₅₀	2024-12-09	🔗 📄 📊 📈	6 Å	🏠
AlphaNet-v1-OMA	0.769	0.968	0.901	5.747	0.879	0.024	0.831	0.644	0.079	6.0M (113M) OMat24+Alex+MP1	4.65M	EF ₅₀	2025-05-12	🔗 📄 📊 📈	5 Å	🏠
MatterSim v1 5M	0.767	0.959	0.862	5.852	0.895	0.024	0.863	0.574	0.073	17M MatterSim	4.55M	EF ₅₀	2024-12-16	🔗 📄 📊 📈	6 Å	🏠
GRACE-1L-OAM	0.761	0.944	0.824	5.255	0.803	0.031	0.842	0.516	0.072	6.0M (113M) OMat24+Alex+MP1	3.45M	EF ₅₀	2025-02-06	🔗 📄 📊 📈	6 Å	🏠
Egnorm MP1	0.756	0.929	0.786	4.844	0.741	0.040	0.799	0.408	0.084	146k (1.58M) MP1	1.31M	EF ₅₀	2025-05-26	🔗 📄 📊 📈	6 Å	🏠
Nequip-MP-L	0.733	0.921	0.761	4.704	0.719	0.043	0.791	0.452	0.086	146k (1.58M) MP1	9.6M	EF ₅₀	2025-09-08	🔗 📄 📊 📈	6 Å	🏠
Nequip MP	0.729	0.914	0.751	4.455	0.681	0.044	0.782	0.446	0.085	146k (1.58M) MP1	708k	EF ₅₀	2025-08-17	🔗 📄 📊 📈	6 Å	🏠
Allegro-MP-L	0.720	0.915	0.751	4.516	0.690	0.044	0.778	0.504	0.082	146k (1.58M) MP1	18.7M	EF ₅₀	2025-09-08	🔗 📄 📊 📈	6 Å	🏠
DPA-3.1-MP1	0.718	0.936	0.803	5.024	0.768	0.037	0.812	0.450	0.080	146k (1.58M) MP1	4.81M	EF ₅₀	2025-06-05	🔗 📄 📊 📈	6 Å	🏠

<https://matbench-discovery.materialsproject.org/>

3 predictions for ML in (materials) science:

- AI will not replace creative thinkers with domain knowledge.
- Understanding underlying theory (of ML and materials science) is key to progress.
- Bright future for machine-learning experts in materials science.