Introduction

Sai Gautam Gopalakrishnan,[1] Keith Tobias Butler[2]

Hands-on sessions[1]: Dereje Bekele Tekliye, Sougat Purohit, Debsundar Dey, Tejus Rohatgi, Reshma Devi, and Pritam Ghosh

[1]Materials Engineering, Indian Institute of Science

[2]Chemistry, University College London

saigautamg@iisc.ac.in[1]; k.t.butler@ucl.ac.uk[2]; https://sai-mat-group.github.io

Jan 7, 2025

Acknowledgments

Group picture in Jun 2024

Prof. Keith Butler (UCL)

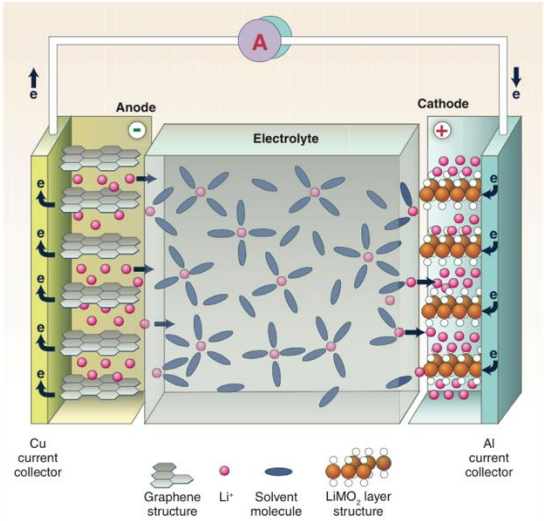
Debolina

Ankur

Javeed

Why bother about materials science?

Key performance bottlenecks in key applications: governed by materials used

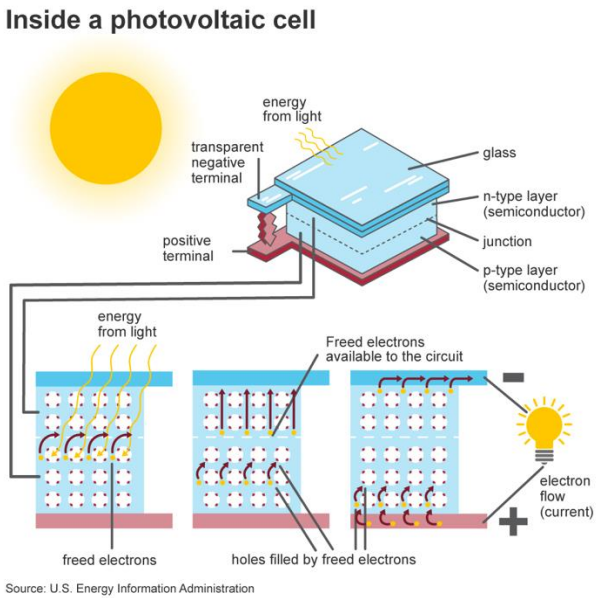


B. Dunn et al., Science 2011

Energy and power density of a battery: limited by materials used as electrodes (and at times, electrolytes)

Key material properties: stability, ionic mobility, reaction energies

Usage of better materials → better performance

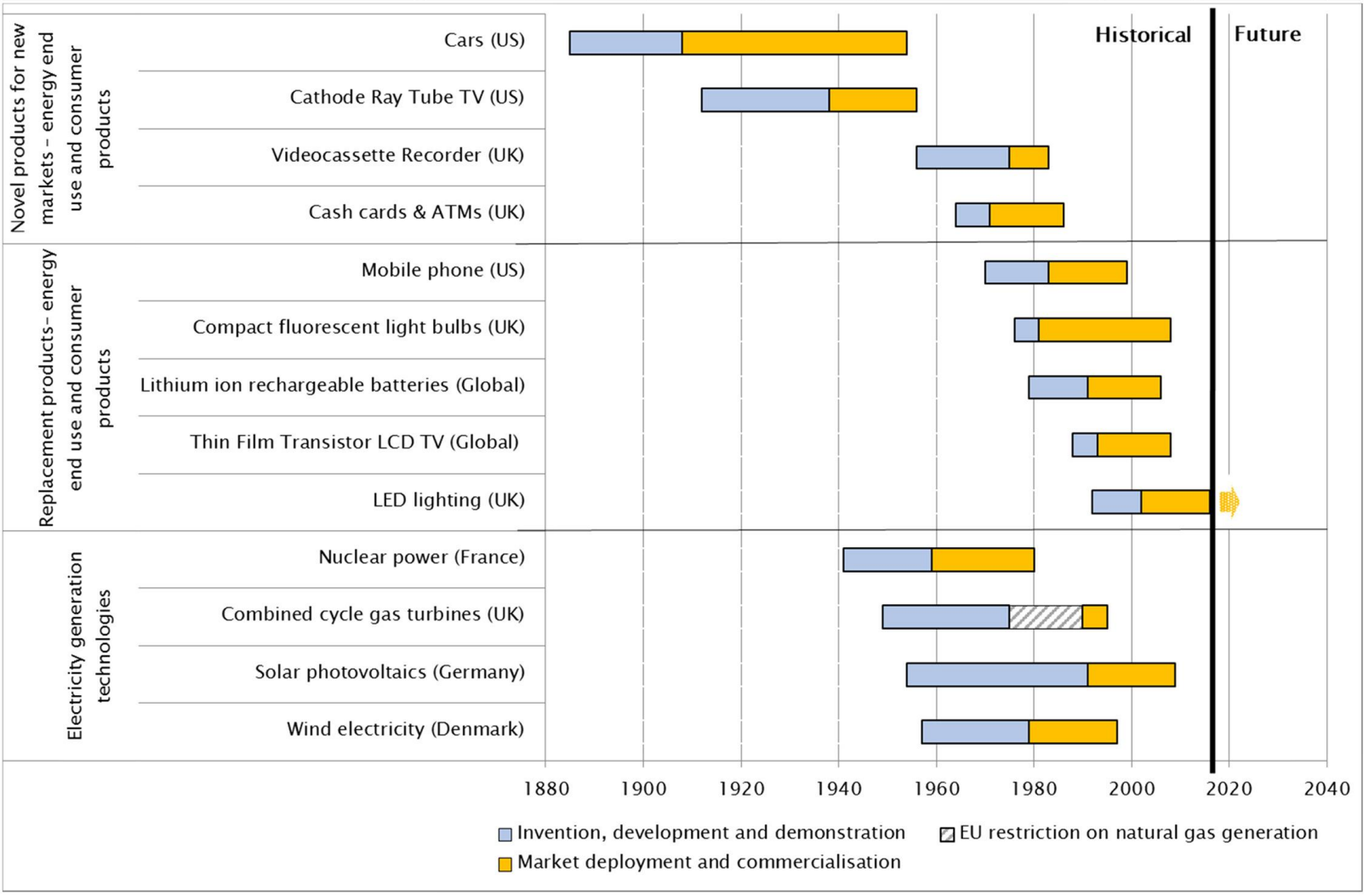


Efficiency of a photovoltaic: choice of semiconductor used as the light absorber

Key material properties: band gap, stability, resistance to point defects

Why use machine learning (ML) in materials science?

Technological innovation and deployment is a ‘slow’ process: often limited by materials

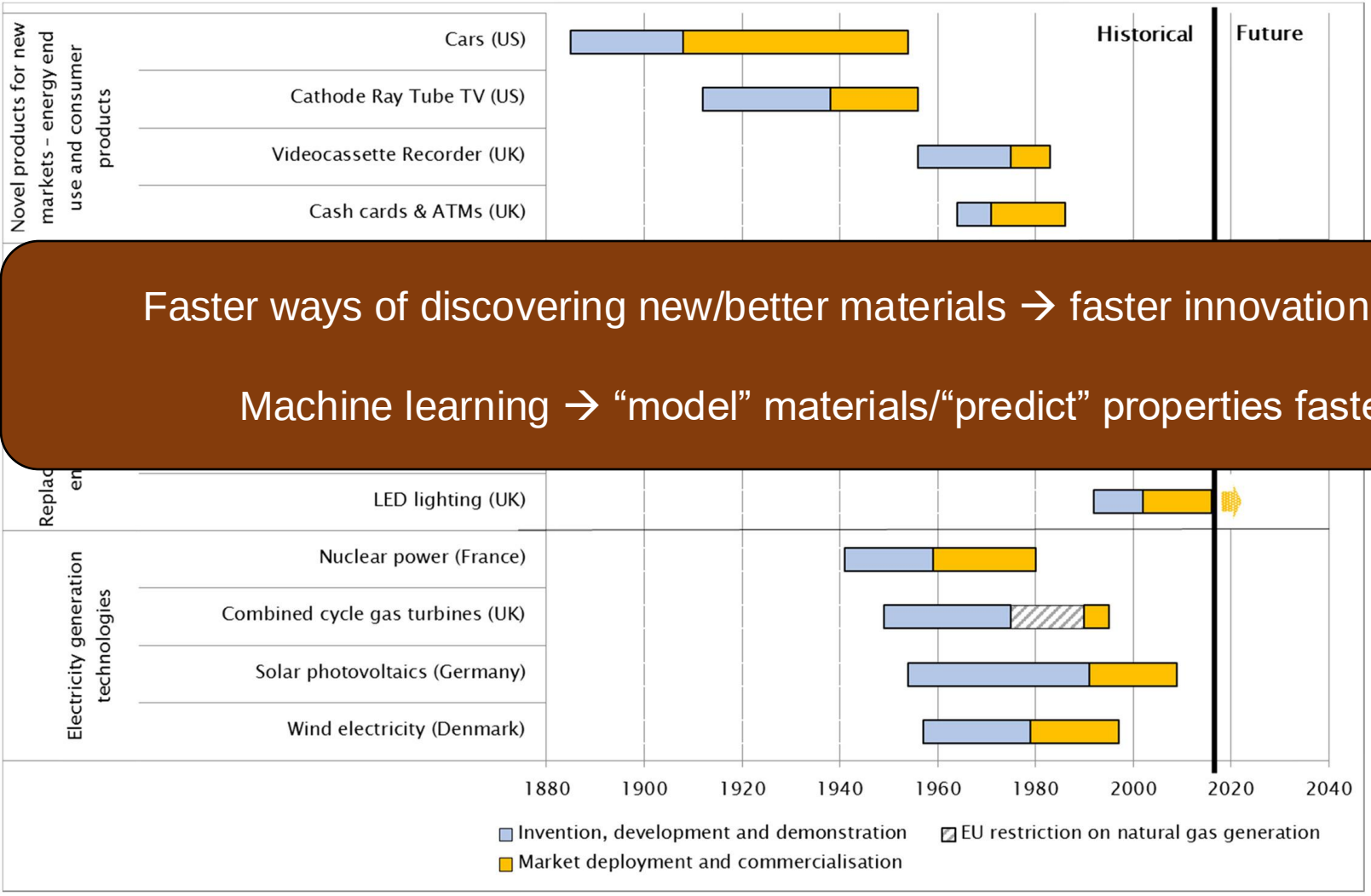


Innovation is particularly slow in energy generation sector!

Gross et al., Energy Policy 123, 682-699 (2018)

Why use machine learning (ML) in materials science?

Technological innovation and deployment is a ‘slow’ process: often limited by materials



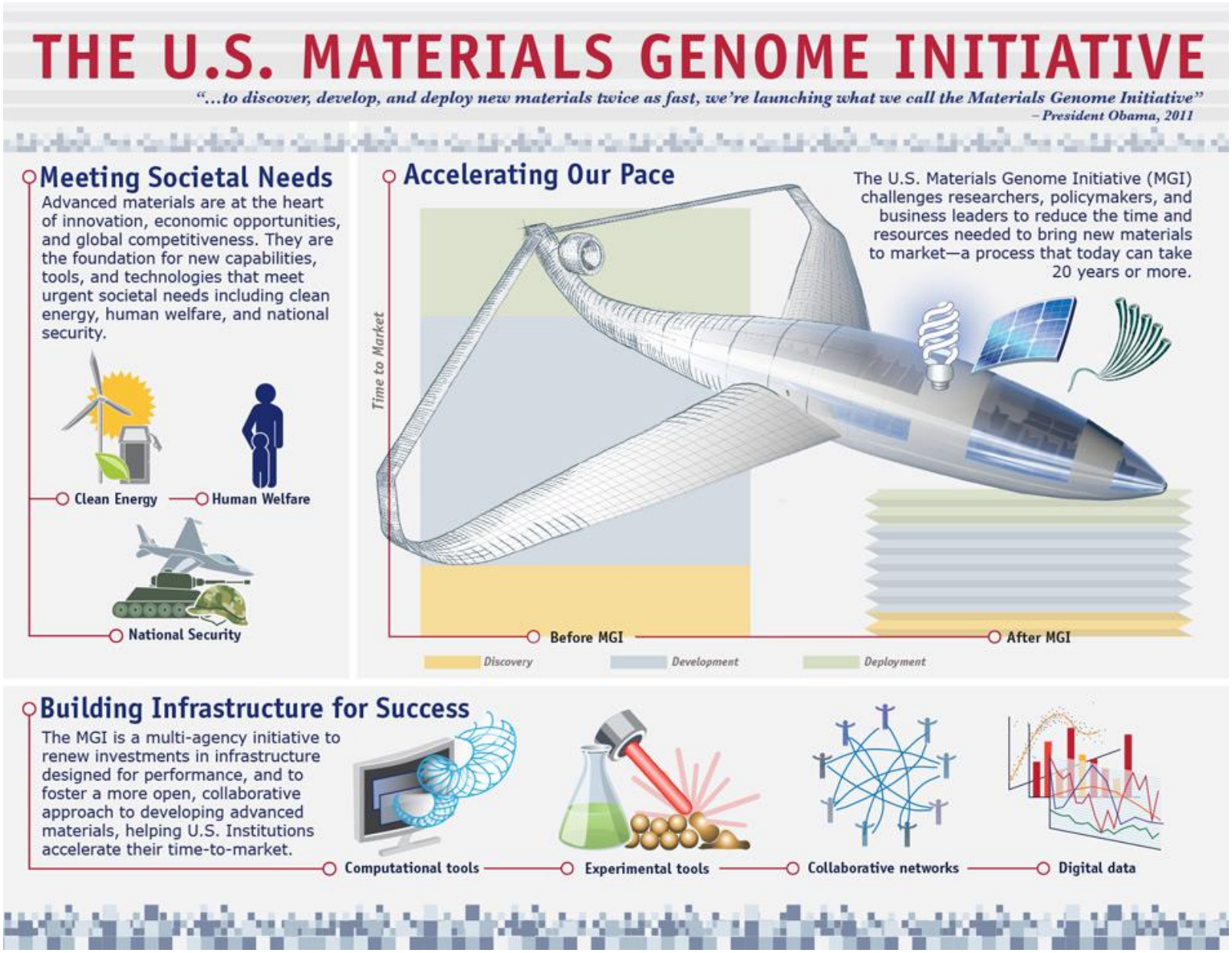
Faster ways of discovering new/better materials → faster innovation cycles

Machine learning → “model” materials/“predict” properties faster

Innovation is particularly slow in energy generation sector!

Gross et al., Energy Policy 123, 682-699 (2018)

Materials Genome (2011-present)



Evolution of ‘modelling’ in materials science

On the determination of molecular fields. —II. From the equation of state of a gas

J. E. Jones

Published: **01 October 1924** | https://doi.org/10.1098/rspa.1924.0082

1924

1957

RESEARCH ARTICLE | AUGUST 13 2004

Phase Transition for a Hard Sphere System

Special Collection: JCP 90 for 90 Anniversary Collection

B. J. Alder; T. E. Wainwright

Check for updates

J. Chem. Phys. 27, 1208–1209 (1957)

https://doi.org/10.1063/1.1743957 **Article history**

Inhomogeneous Electron Gas

P. Hohenberg and W. Kohn

Phys. Rev. **136**, B864 – Published 9 November 1964

1964

1975

Clustering and ordering in solid solutions

D. de Fontaine

Computer simulation of local order in condensed phases of silicon

Frank H. Stillinger and Thomas A. Weber

Phys. Rev. B **31**, 5262 – Published 15 April 1985; Erratum Phys. Rev. B **33**, 1451 (1986)

1986

1996

Generalized Gradient Approximation Made Simple

John P. Perdew, Kieron Burke, and Matthias Ernzerhof

Phys. Rev. Lett. **77**, 3865 – Published 28 October 1996; Erratum

From ultrasoft pseudopotentials to the projector augmented-wave method

G. Kresse and D. Joubert

Phys. Rev. B **59**, 1758 – Published 15 January 1999

1999

2003

Shock Waves in High-Energy Materials: The Initial Chemical Events in Nitramine RDX

Alejandro Strachan, Adri C. T. van Duin, Debashis Chakraborty, Siddharth Dasgupta, and William A. Goddard, III

Phys. Rev. Lett. **91**, 098301 – Published 28 August 2003

Generalized Neural-Network Representation of High-Dimensional Potential-Energy Surfaces

Jörg Behler and Michele Parrinello

Phys. Rev. Lett. **98**, 146401 – Published 2 April 2007

2007

Predicting Crystal Structures with Data Mining of Quantum Calculations

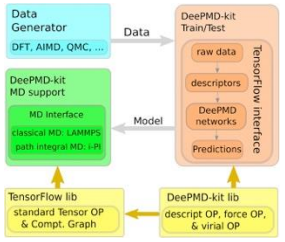
Stefano Curtarolo, Dane Morgan, Kristin Persson, John Rodgers, and Gerbrand Ceder

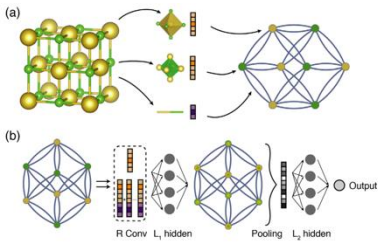
Phys. Rev. Lett. **91**, 135503 – Published 24 September 2003

THE U.S. MATERIALS GENOME INITIATIVE

2011-2018

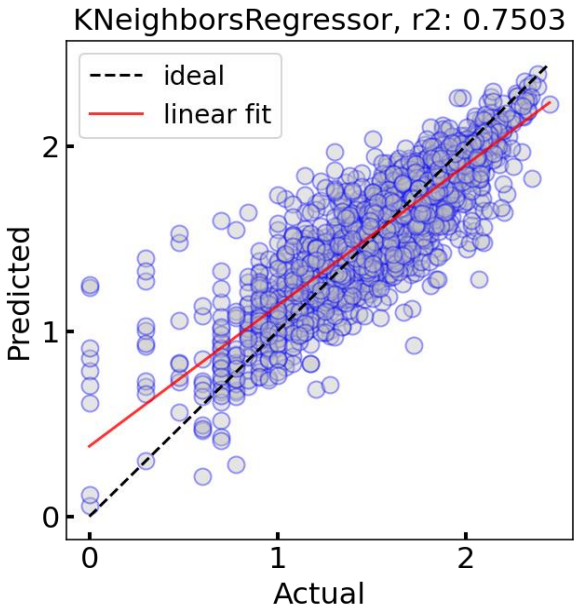
2018-present



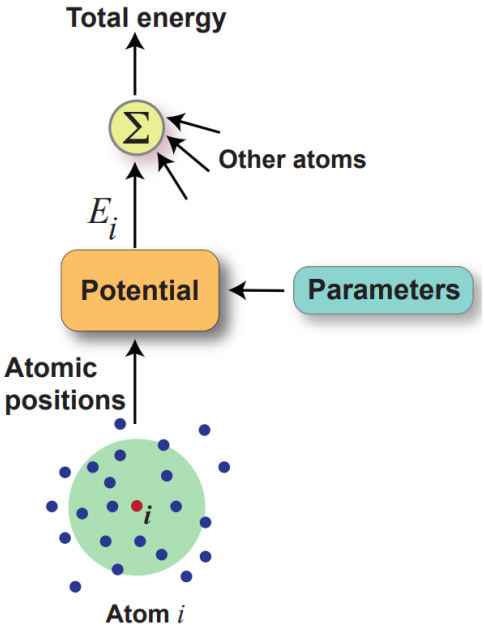


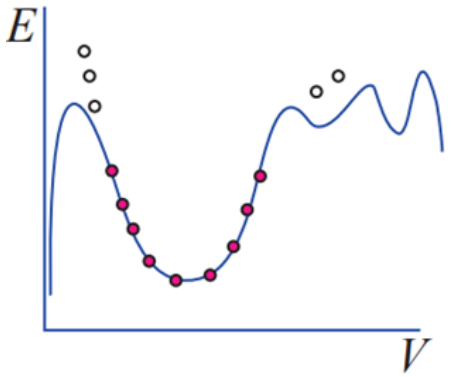
Types of ML in materials science

Regressions: make property predictions better with ‘simple’ inputs (also classifications)

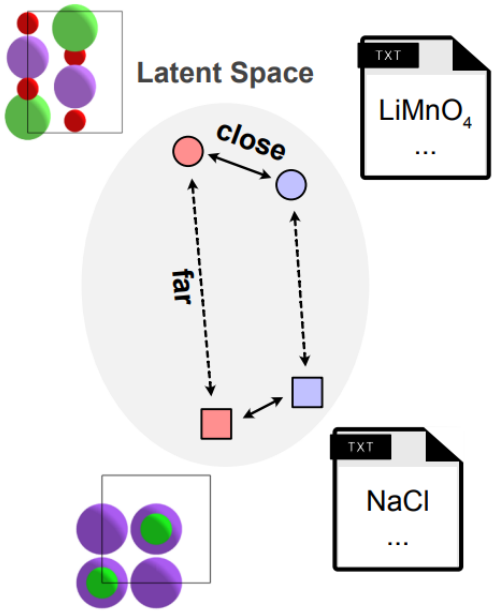


Interatomic potentials: describe potential energy surface accurately



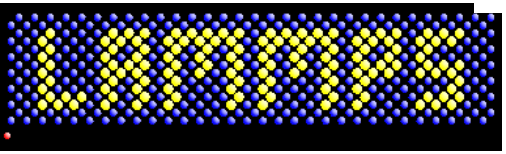


Advanced topics: Diffusion (generative) models, language models, transfer learning



Where does the data come from?

Optimization design and scale-up



Home

Search

Home | Benchmark Info | Full Benchmark Data | How To Use | Leaderboards Per Task | Reference

Home

Leaderboard-Property: General Purpose Algorithms on `matbench_v0.1`

Find more information about this benchmark on the benchmark info page

Task name	Samples	Algorithm	Verified MAE (unit) or ROCAUC	Notes
matbench_steels	312	MODNet (v0.1.12)	87.7627 (MPa)	
matbench_jdft2d	636	MODNet (v0.1.12)	33.1918 (meV/atom)	
matbench_phonons	1,265	MegNet (kgcnn v2.1.0)	28.7606 (cm^-1)	structure required
matbench_expt_gap	4,604	MODNet (v0.1.12)	0.3327 (eV)	
matbench_dielectric	4,764	MODNet (v0.1.12)	0.2711 (unitless)	
matbench_expt_is_metal	4,921	AMMExpress v2020	0.9209	
matbench_glass	5,680	MODNet (v0.1.12)	0.9603	
matbench_log_gvrh	10,987	coNGN	0.0670 (log10(GPa))	structure required
matbench_log_kvrh	10,987	coNGN	0.0491 (log10(GPa))	structure required
matbench_perovskites	18,928	coGN	0.0269 (eV/unit cell)	structure required
matbench_mp_gap	106,113	coGN	0.1559 (eV)	structure required
matbench_mp_is_metal	106,113	CGCNN v2019	0.9520	structure required
matbench_mp_e_form	132,752	coGN	0.0170 (eV/atom)	structure required

2019)

https://matbench.materialsproject.org/

Data organization: python/API

ML: python

ICSD

NIST National Institute of Standards and Technology

HTEM

Schedule

07 JAN 2025 Tuesday

09:00 a.m. REGISTRATION
BRIEF INAUGURATION

09:30 a.m. SESSION 1:
BASICS OF SUPERVISED ML-
REGRESSION &
CLASSIFICATION MODELS

11:00 a.m. COFFEE BREAK

11:30 a.m. TUTORIAL FOR SESSION 1

01:00 p.m. LUNCH

02:00 p.m. SESSION 2A:
UNSUPERVISED ML-
DEEP LEARNING

TUTORIAL FOR SESSION 2A

03:30 p.m. COFFEE BREAK

04:00 p.m. SESSION 2B:
UNSUPERVISED ML-
GRAPH NEURAL NETWORKS

TUTORIAL FOR SESSION 2B

05:30 p.m. COFFEE BREAK

06:00 p.m. POSTER SESSION

07:00 p.m. DINNER

08 JAN 2025 Wednesday

09:30 a.m. SESSION 3:
MACHINE LEARNED
INTERATOMIC POTENTIALS

11:00 a.m. COFFEE BREAK

11:30 a.m. TUTORIAL FOR SESSION 3

01:00 p.m. LUNCH

02:00 p.m. SESSION 4A:
ADVANCED AI/ML-
TRANSFER LEARNING

TUTORIAL FOR SESSION 4A

03:00 p.m. COFFEE BREAK

03:15 p.m. SESSION 4B:
ADVANCED AI/ML-
LANGUAGE MODELS

TUTORIAL FOR SESSION 4B

04:15 p.m. COFFEE BREAK

04:30 p.m. SESSION 4C:
ADVANCED AI/ML-
DIFFUSION MODELS

TUTORIAL FOR SESSION 4C

05:30 p.m. CLOSING CEREMONY

05:45 p.m. HIGH TEA