Machine Learning and High-Performance Computing for Material Science

Murat Keçeli
Computational Science Division
Argonne National Laboratory

Argonne NATIONAL LABORATORY

November 26, 2021 @ Magnetic Properties from First Principles Workshop

Motivation

- The rational design of molecules and materials is the key to solve the challenging problems we face from pandemics to global warming.
- Coupling formula-driven simulations with data-driven methods is required to accelerate material discovery.

Motivation

- The rational design of molecules and materials is the key to solve the challenging problems we face from pandemics to global warming.
- Coupling formula-driven simulations with data-driven methods is required to accelerate material discovery.

BARRON'S

TECHNOLOGY | OTHER VOICES

3 Technologies That Could Create Trillion-Dollar Markets Over the Next Decade

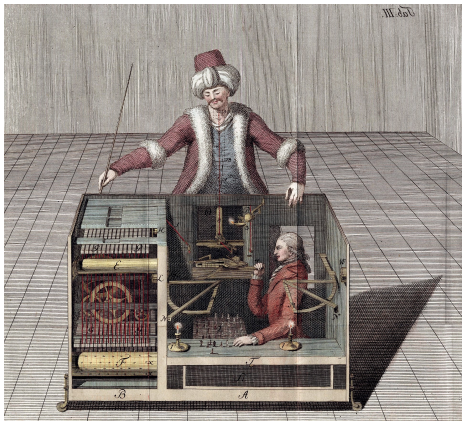
COMMENTARY By Greg Satell Updated April 21, 2019 / Original February 17, 2019

1. CRISPR
2. Quantum and neuromorphic computing
3. **Materials simulation**

*Traditionally, developing new materials has been a slow, painstaking process… Yet today, we're in the midst of a **materials revolution**. Powerful simulation techniques, combined with **increased computing power and machine learning**, are enabling researchers to automate much of the discovery process, vastly accelerating the development of new materials, in some cases by a factor of more than a hundred.*

What is Machine Learning?

- AI is defined broadly as the study of systems that can perceive their environment and take actions to achieve their goals.
  - Formal design of Turing-complete artificial neurons by McCullouch and Pitts in 1943 is considered as the first AI work.
- ML is the study of models that can learn patterns from data and make predictions without being explicitly programmed.
  - Original definition by Arthur Samuel in 1959.



The Turk, 1770

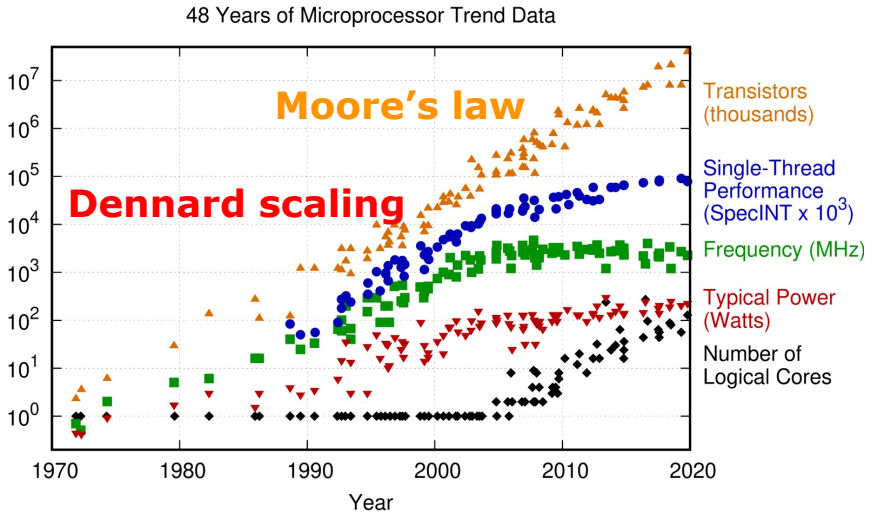
Deep Blue, 1996

DeepMind

AlphaZero, 2017

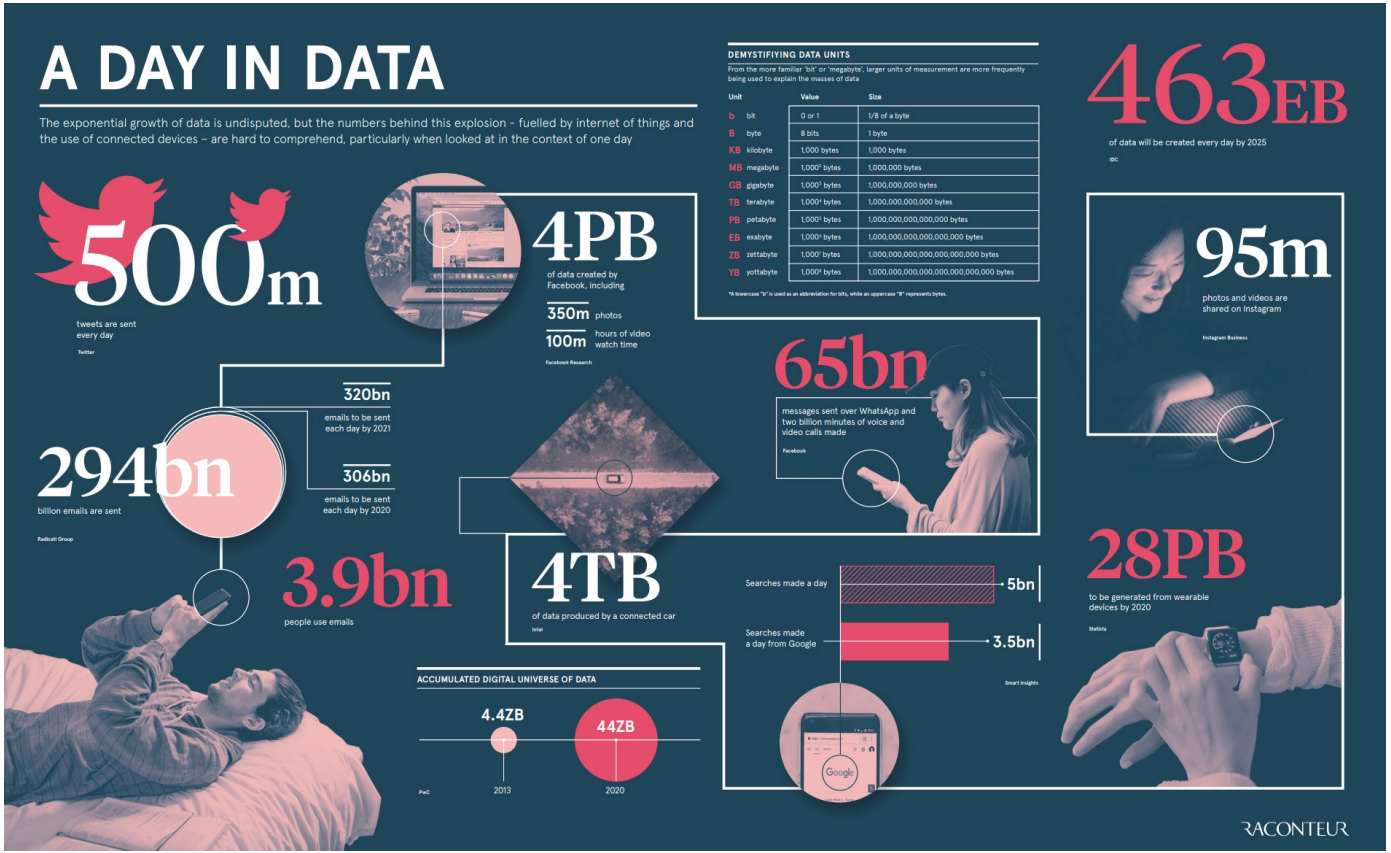
Faster & Cheaper Computing Power

https://en.wikipedia.org/wiki/FLOPS



https://github.com/karlrupp/microprocessor-trend-data

Big Data



THE 2020 ONLINE BIG DATA FACTS

4.6bn. people online

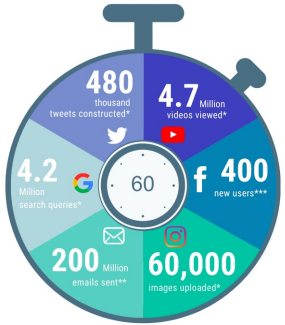
5.1bn. mobile phone owners

2bn. online shoppers

3.7bn. social media users

HOW MUCH DATA IS OUT THERE?
World data is predicted to reach **175ZB** by 2025. That much data would take one person 1.8 billion years to download at current internet speeds!

WHAT HAPPENS ONLINE EVERY MINUTE?



Learn more about Data Quality at
nodegraph.se

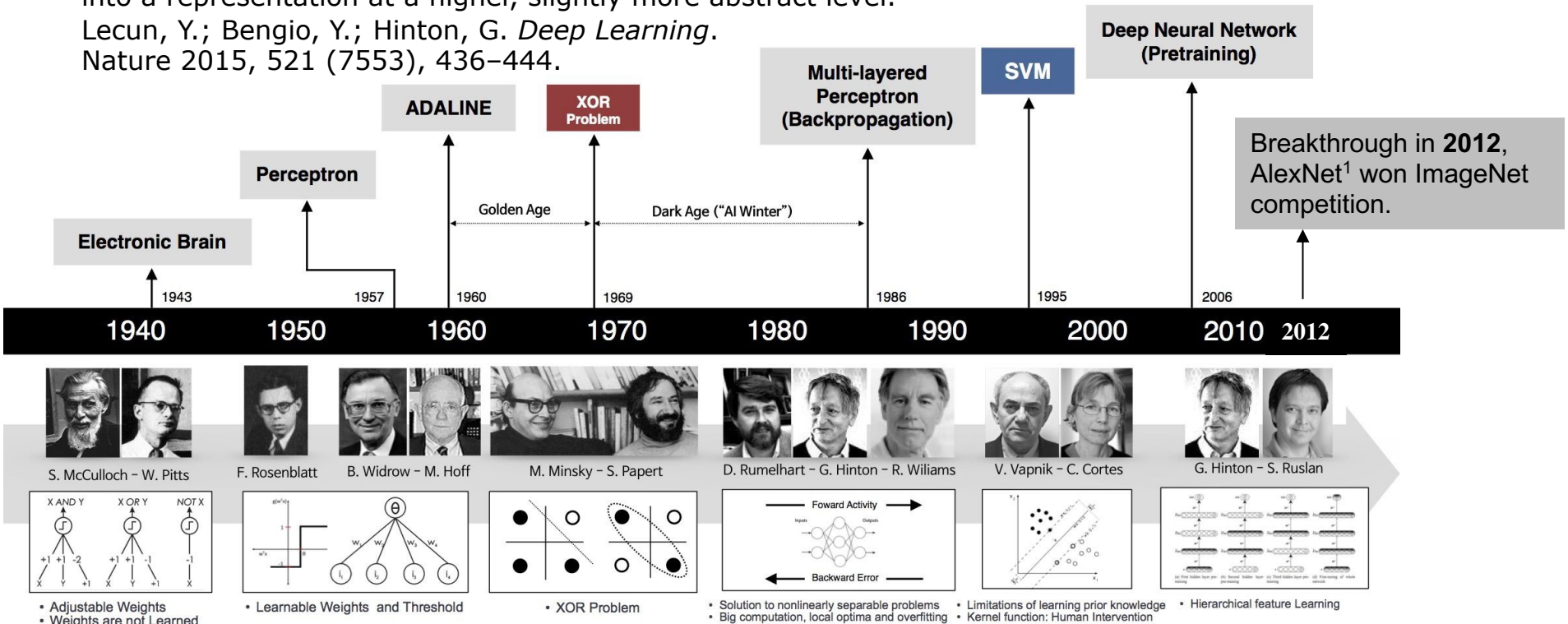
Sources: nodegraph.se, *internetlivestats.com, **lifewire.com, ***oneclickagency.com

https://www.visualcapitalist.com/how-much-data-is-generated-each-day/

https://www.nodegraph.se/how-much-data-is-on-the-internet/

Big data & compute enabled deep learning

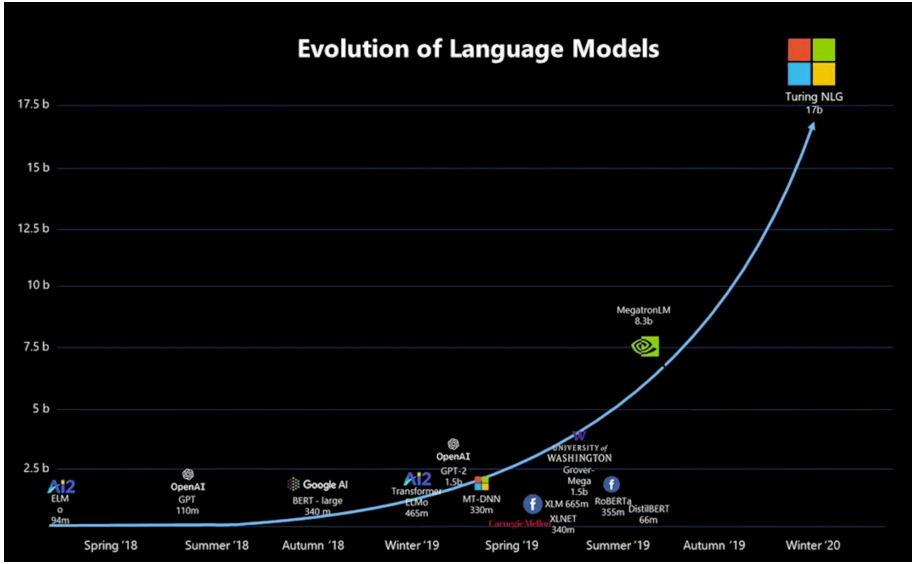
DL methods are "representation-learning methods with multiple levels of representation, obtained by composing simple but non-linear modules that each transform the representation at one level (starting with the raw input) into a representation at a higher, slightly more abstract level."
Lecun, Y.; Bengio, Y.; Hinton, G. *Deep Learning*. Nature 2015, 521 (7553), 436–444.

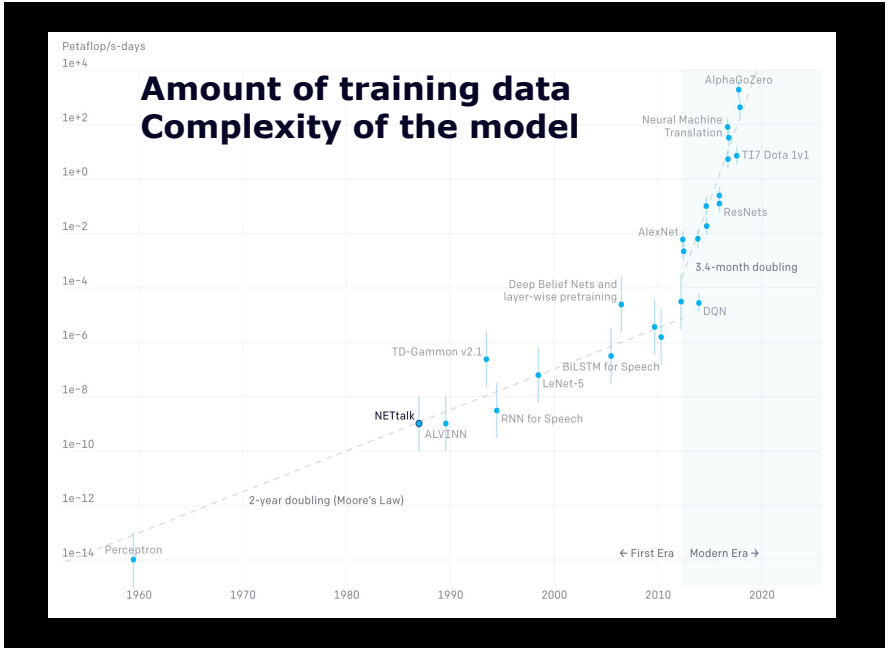


https://slides.com/beamandrew/deep-learning-101

[1]Krizhevsky, A., Sutskever, I. & Hinton, G. Proc. Advances in NeurIPS 25 1090 (2012)

Deeper Learning More Computing

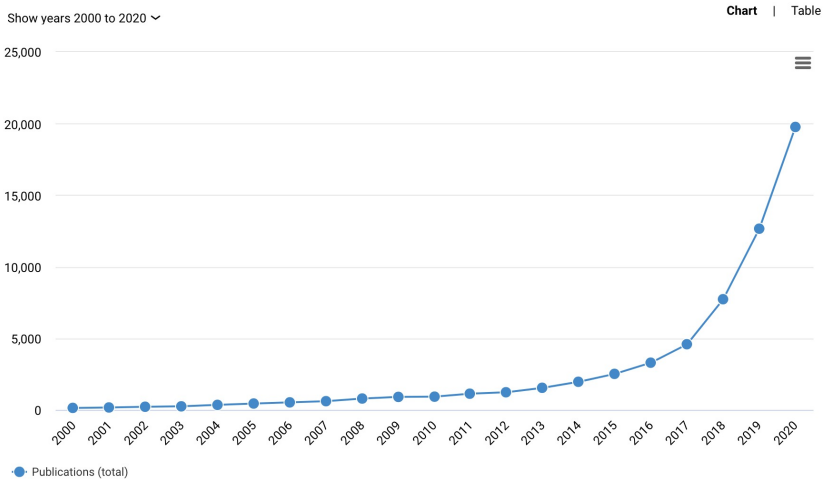




https://openai.com/blog/ai-and-compute

Machine Learning for Science

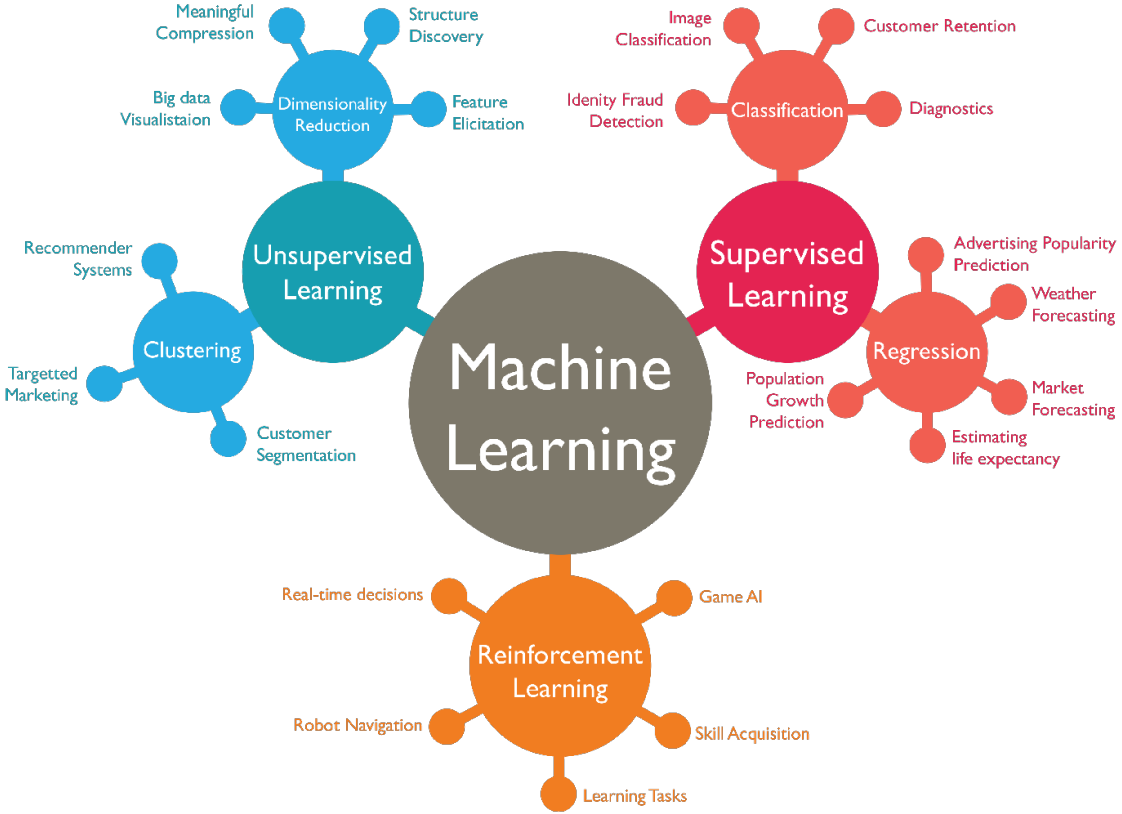
- Industry needs more data scientists and machine learning experts.
- More computing ➔ more data ➔ more computing ➔ …
- Barrier for entry is low.
- Free courses, open-source codes.
- ML for science grows exponentially. Accuracy and applicability is limited by the training set.



"Machine learning" publications in physical or chemical sciences journals

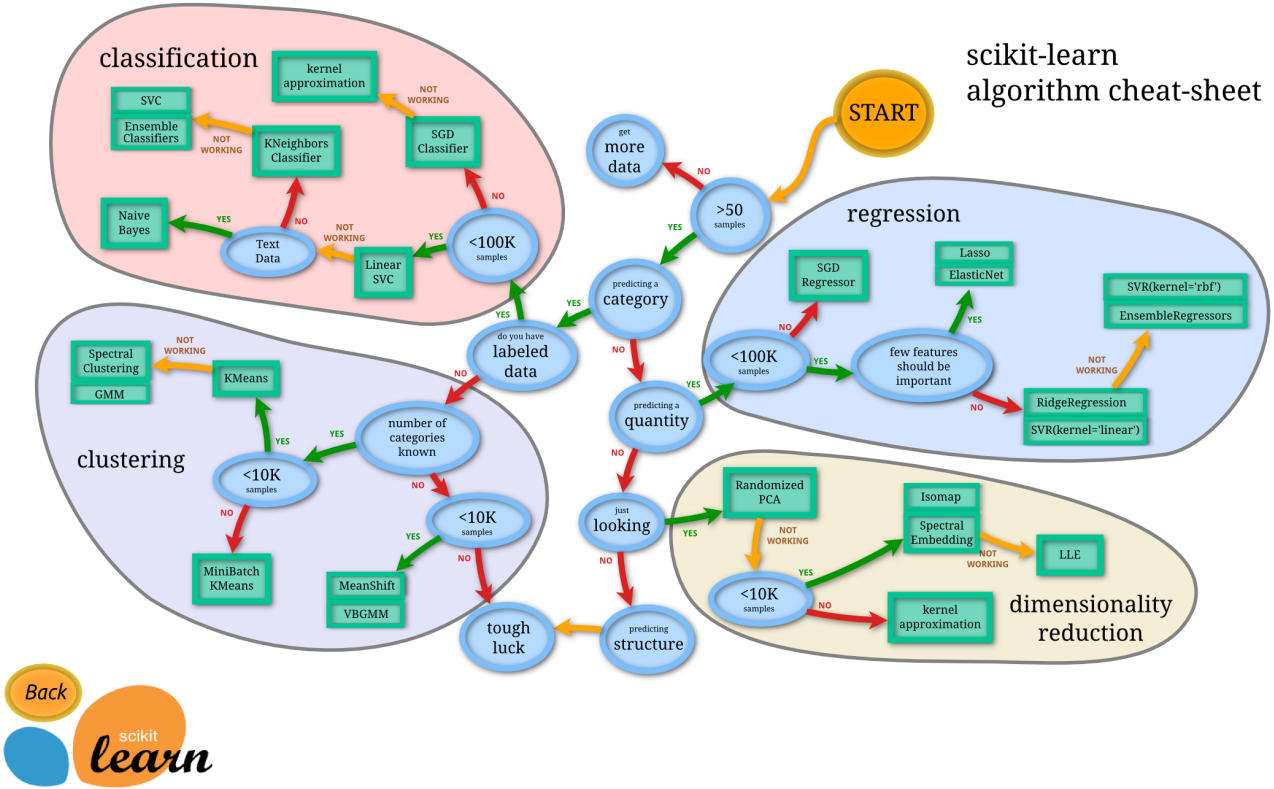
https://app.dimensions.ai/

Types of Machine Learning



https://towardsdatascience.com/machine-learning-types-2-c1291d4f04b1

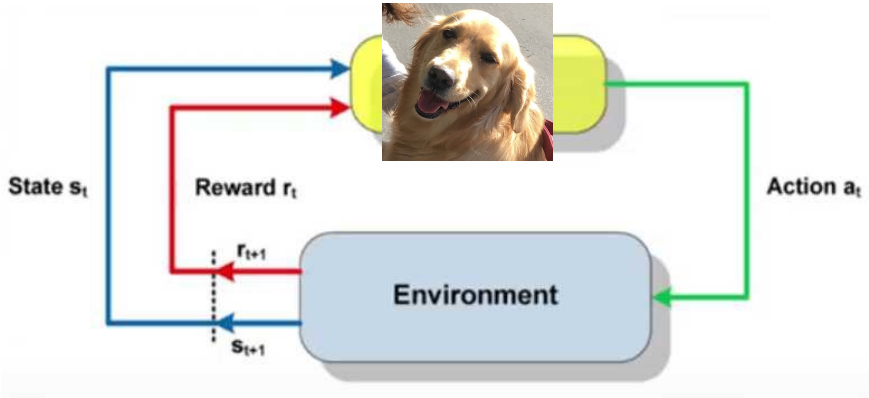
How to select the ML method?

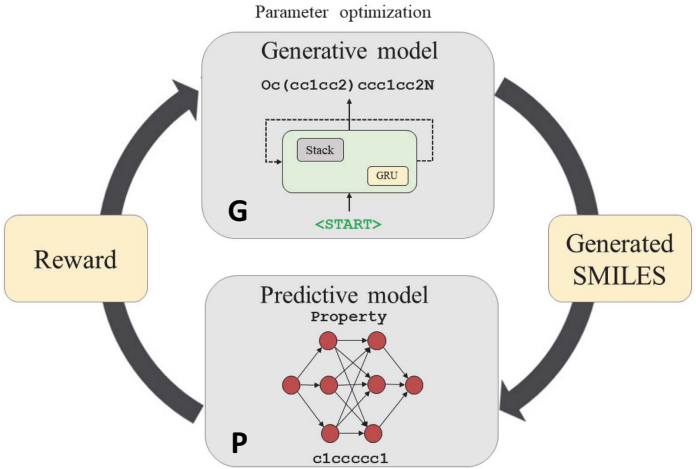


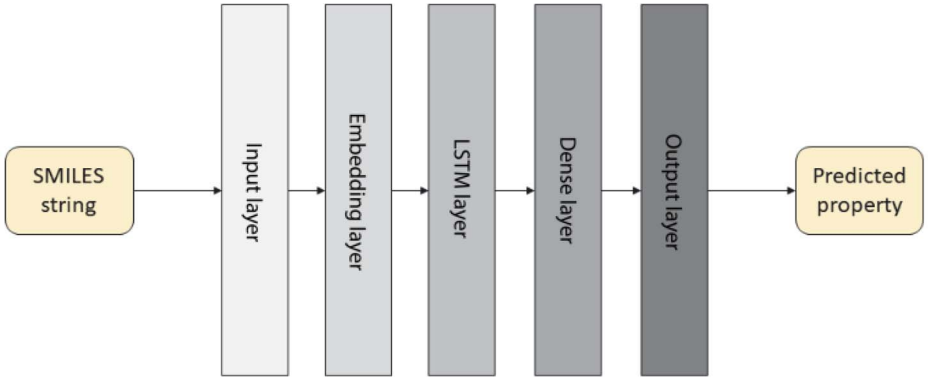
https://scikit-learn.org/stable/tutorial/machine_learning_map/index.html

Reinforcement Learning

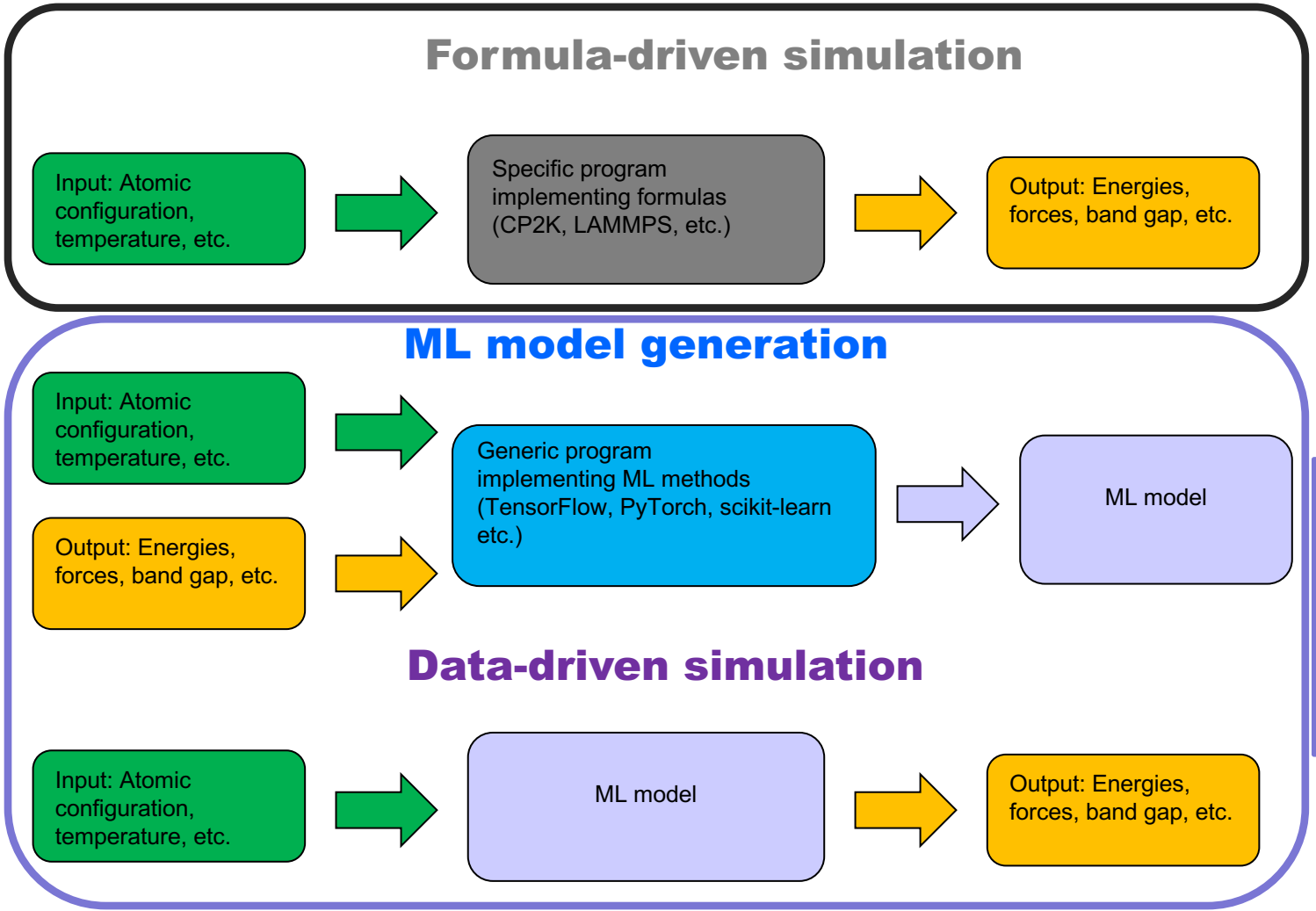
- Agents take actions to maximize their rewards.
- Does not require labelled input/output pairs.





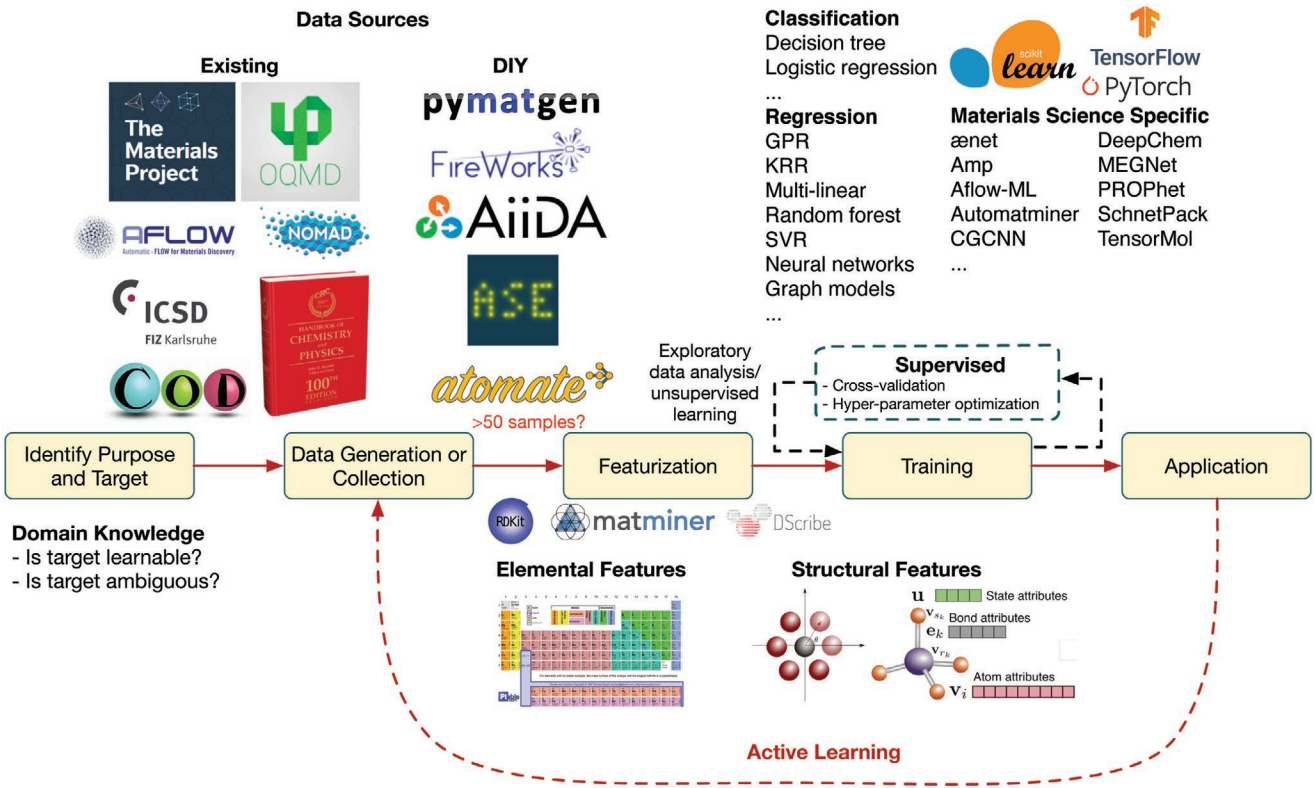


Popova, M.; Isayev, O.; Tropsha, A. Deep Reinforcement Learning for de Novo Drug Design. *Sci. Adv.* **2018**, *4* (7), 1–15.
Sui, F.; Guo, R.; Zhang, Z.; Gu, G. X.; Lin, L. Deep Reinforcement Learning for Digital Materials Design. ACS Mater. Lett. **2021**, 3 (10), 1433–1439.



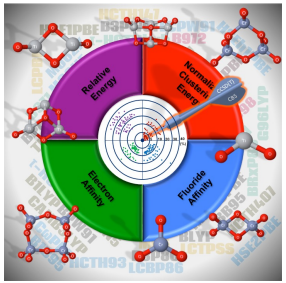
- Orders of magnitude faster with better scalability
- Accuracy depends on the training set and ML model

Machine Learning Workflow



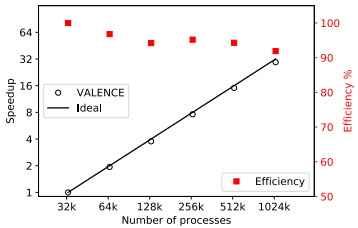
Chen, C.; Zuo, Y.; Ye, W.; Li, X.; Deng, Z.; Ong, S. P. A Critical Review of Machine Learning of Energy Materials. Adv. Energy Mater. **2020**, 10 (8), 1903242.

What Do We Want?



More accurate

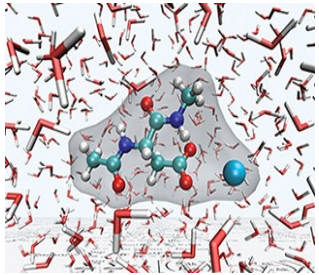
Chemical accuracy (1 kcal/mol) for thermochemistry and kinetics.



Faster simulations

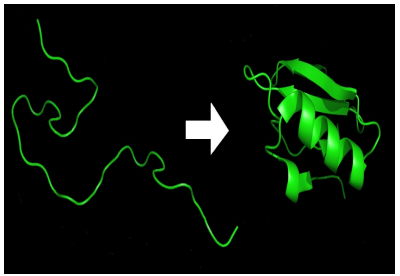
Push for strong scaling.

Profile, tune, benchmark.



Larger size

Push for weak scaling.
Disordered materials, biomolecules,
liquid phase chemistry,



Longer time

Protein folding (beyond ms, 10^{12} time steps)

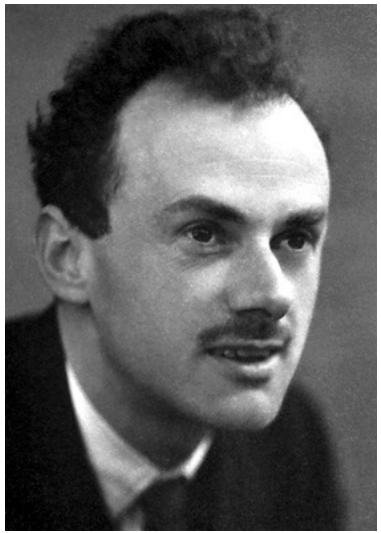
More calculations

High-throughput calculations, i.e. screening of candidates for a photovoltaic

The total number of:
atoms on the earth ~10^{50}
small organic molecules ~10^{60}

Our goal is to extend the applicability domain of predictive computational chemistry & material science.

Chemistry & Material Science: Quantum Mechanics



PAM Dirac
(Nobel prize in 1933)

"The underlying physical laws necessary for the mathematical theory of a large part of physics and the whole of chemistry are thus completely known, and the difficulty is only that the exact application of these laws leads to equations much too complicated to be soluble."*

Dirac equation

$$\left(\beta mc^2 + c\sum_{n=1}^{3}\alpha_n p_n\right)\psi(x,t) = i\hbar\frac{\partial\psi(x,t)}{\partial t}$$

Schrödinger equation

$$i\hbar\frac{\partial}{\partial t}\Psi(\mathbf{r},t) = \left[\frac{-\hbar^2}{2m}\nabla^2 + V(\mathbf{r},t)\right]\Psi(\mathbf{r},t)$$

Time-independent Schrödinger equation

$$\left[\frac{-\hbar^2}{2m}\nabla^2 + V(\mathbf{r})\right]\Psi(\mathbf{r}) = E\Psi(\mathbf{r})$$

*Proceedings of the Royal Society of London. Series A, Vol. 123, No. 792, 1929

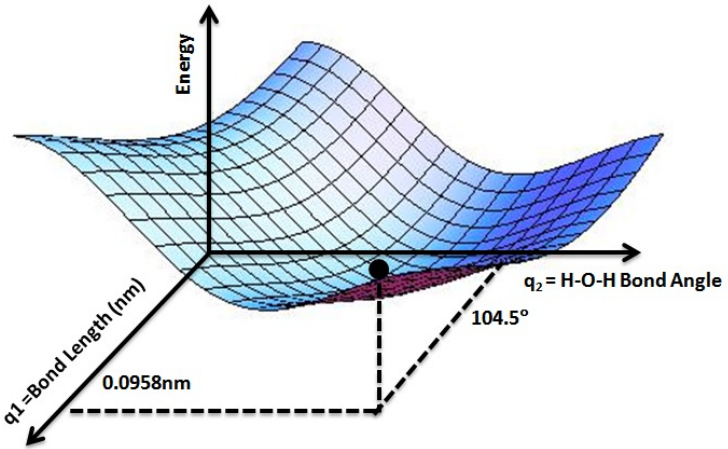
The Born–Oppenheimer Approximation*

- Approximate separation of electronic and nuclear degrees of freedom in the molecular Hamiltonian, since nuclei are much heavier than electrons.

$$\hat{H}_e(\mathbf{r};\mathbf{R})\Psi_e(\mathbf{r};\mathbf{R}) = E_e(\mathbf{R})\Psi_e(\mathbf{r};\mathbf{R})$$

$$\hat{H}_n(\mathbf{R})\Psi_n(\mathbf{R}) = \{E - E_e(\mathbf{R})\}\Psi_n(\mathbf{R})$$

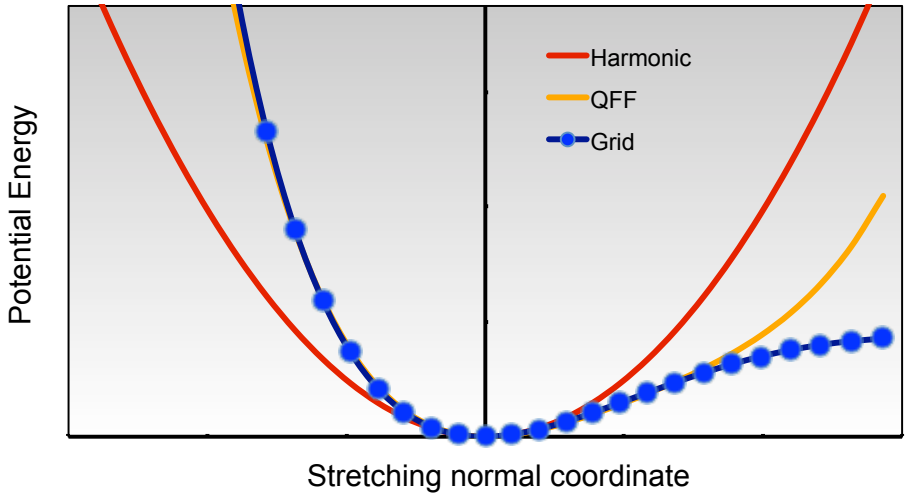
- Introduces the concept of **Potential Energy Surface**, which gives the energy of a molecule as a function of the nuclear coordinates.



*M. Born and R. Oppenheimer. Ann. Phys. 84, 457–484 (1927).

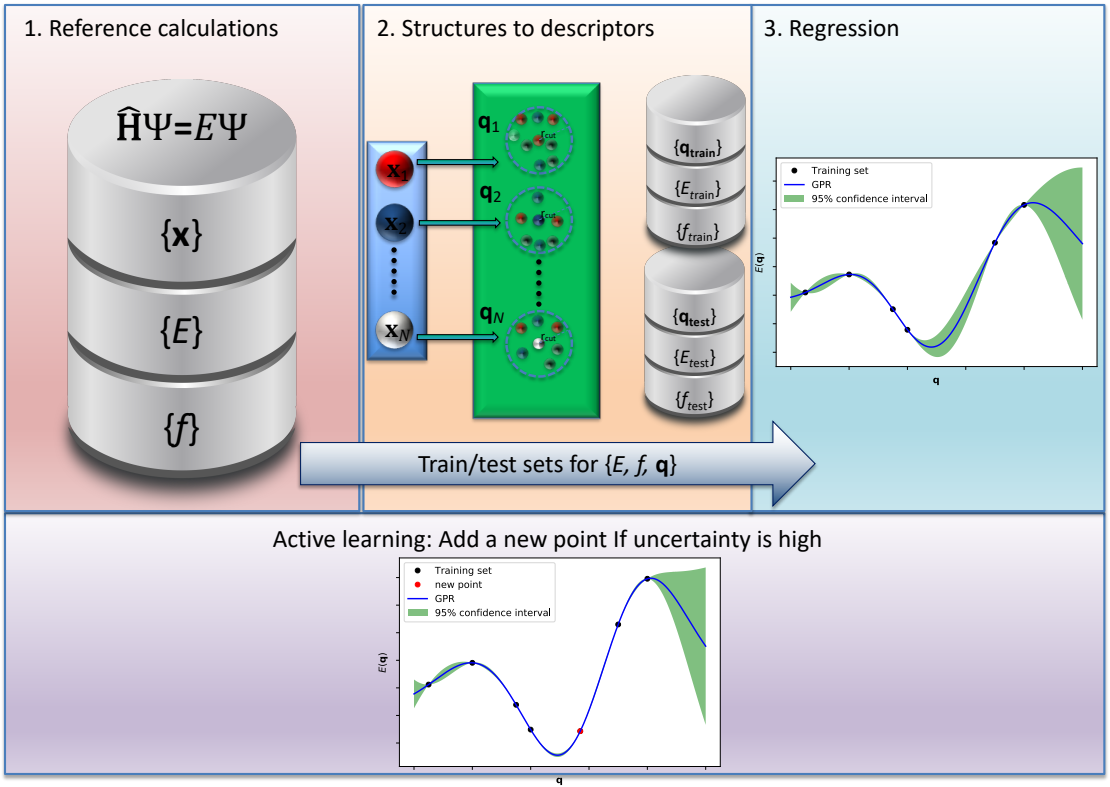
Potential Energy Surface / Force Field / Interatomic Potential

- Central concept for computational chemistry: kinetics, dynamics, spectroscopy
- No closed analytical form
- High dimensionality
- For *M* vibrational degrees of freedom:
  - Gauss-Hermite quadrature: Number of points ~ P^M
  - Taylor expansion truncated at order *T:* Number of points ~ M^T

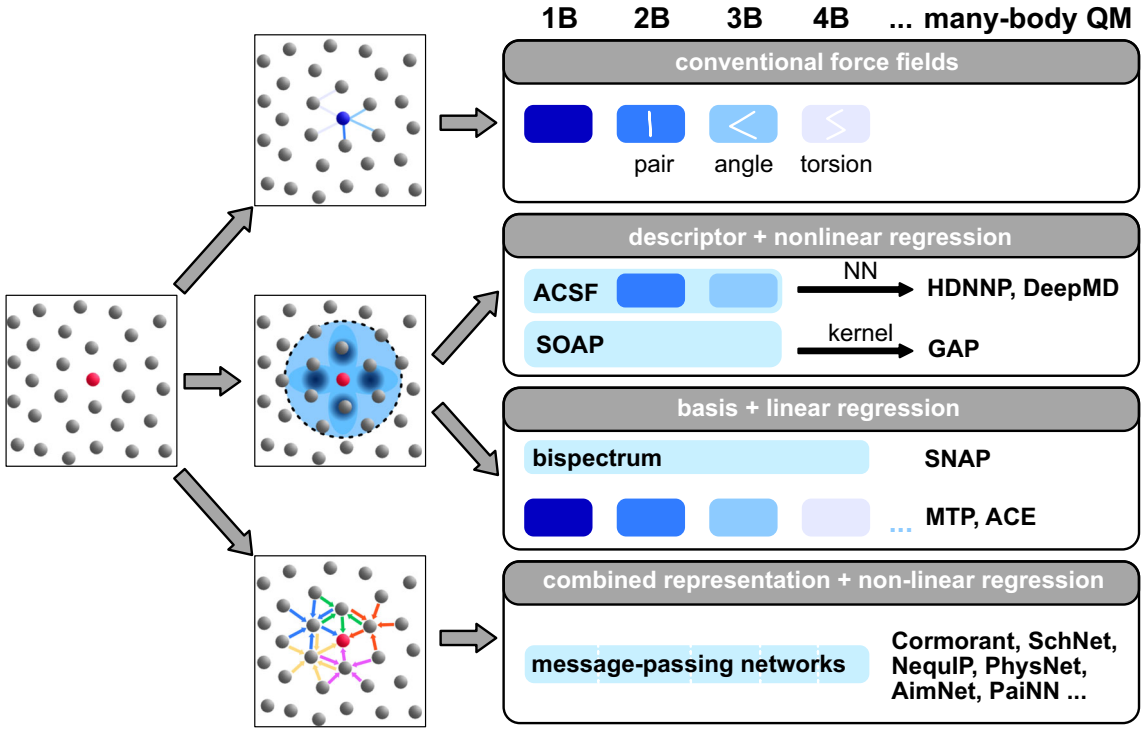


$$V = V_0 + \sum_i F_i Q_i + \frac{1}{2}\sum_{i,j} F_{ij} Q_i Q_j + \frac{1}{6}\sum_{i,j,k} F_{ijk} Q_i Q_j Q_k + \frac{1}{24}\sum_{i,j,k,l} F_{ijkl} Q_i Q_j Q_k Q_l + \ldots$$

Workflow for Machine Learning Potentials



Descriptors for Machine Learning Potentials



- **Descriptors + nonlinear fitting.**
  A compact symmetry preserving representation of the (usually local) atomic geometry is defined, and used as input to nonlinear regression schemes such as neural networks or kernel regression
- **Basis + linear fitting**
  A basis of many symmetric functions of local atomic geometries is constructed, and coefficients are determined using regularised least squares fitting
- **Message passing networks**
  The representation and regression problems are solved simultaneously, starting with just atomic identities, which are iteratively passed along the neighbour graph of the system and combined into a symmetric nonlinear function

Behler, J.; Csányi, G. Machine Learning Potentials for Extended Systems: A Perspective. Eur. Phys. J. B 2021, 94 (7), 142.

Neural Network Potentials

- Artificial neural networks are universal approximators[1]; hence they can be used to fit complex PESs with arbitrary accuracy,
- They do not require any prior functional form,
- They are applicable to any type of bonding without any bias,
- Sampling, training and inference can be done in parallel.

Neural network models of potential energy surfaces

Thomas B. Blank, Steven D. Brown, August W. Calhoun,[a] and Douglas J. Doren
Department of Chemistry and Biochemistry, University of Delaware, Newark, Delaware 19716

(Received 6 February 1995; accepted 7 June 1995)

Neural networks provide an efficient, general interpolation method for nonlinear functions of several variables. This paper describes the use of feed-forward neural networks to model global properties of potential energy surfaces from information available at a limited number of configurations. As an initial demonstration of the method, several fits are made to data derived from an empirical potential model of CO adsorbed on Ni(111). The data are error-free and geometries are selected from uniform grids of two and three dimensions. The neural network model predicts the potential to within a few hundredths of a kcal/mole at arbitrary geometries. The accuracy and efficiency of the neural network in practical calculations are demonstrated in quantum transition state theory rate calculations for surface diffusion of CO/Ni(111) using a Monte Carlo/path integral method. The network model is much faster to evaluate than the original potential from which it is derived. As a more complex test of the method, the interaction potential of H_2 with the Si(100)-2×1 surface is determined as a function of 12 degrees of freedom from energies calculated with the local density functional method at 750 geometries. The training examples are not uniformly spaced and they depend weakly on variables not included in the fit. The neural net model predicts the potential at geometries outside the training set with a mean absolute deviation of 2.1 kcal/mole.

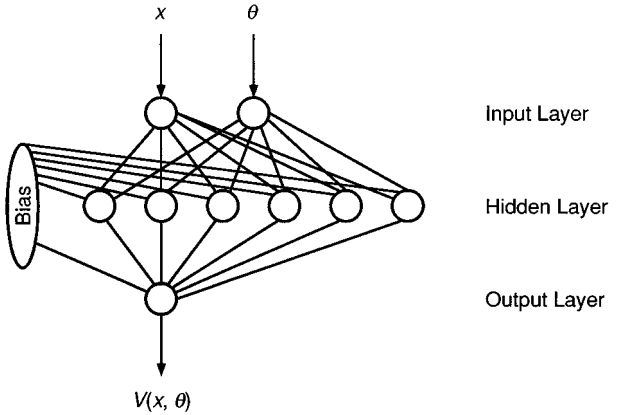


FIG. 1. Schematic of the feed-forward neural network used for the two degree-of-freedom model.

[1]Hornik, K.; Stinchcombe, M.; White, H. Multilayer Feedforward Networks Are Universal Approximators. *Neural Networks* **1989**, *2* (5), 359–366

Behler-Parrinello NNPs

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

Jörg Behler and Michele Parrinello

Department of Chemistry and Applied Biosciences, ETH Zurich, USI-Campus, Via Giuseppe Buffi 13, CH-6900 Lugano, Switzerland

(Received 27 September 2006; published 2 April 2007)

The accurate description of chemical processes often requires the use of computationally demanding methods like density-functional theory (DFT), making long simulations of large systems unfeasible. In this Letter we introduce a new kind of neural-network representation of DFT potential-energy surfaces, which provides the energy and forces as a function of all atomic positions in systems of arbitrary size and is several orders of magnitude faster than DFT. The high accuracy of the method is demonstrated for bulk silicon and compared with empirical potentials and DFT. The method is general and can be applied to all types of periodic and nonperiodic systems.

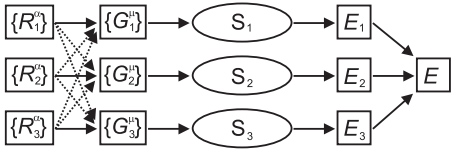


FIG. 2. Structure of the neural network as applied in this Letter to a system containing three atoms. The Cartesian coordinates of atom i are given by R_i^α. These are transformed to a set of μ symmetry function values G_i^μ describing the local geometric environment of atom i, which depends on the positions of all atoms in the system as indicated by the dotted arrows. The symmetry function values of atom i then enter the subnet S_i yielding the energy contribution E_i of atom i to the total energy of the system E. The structure of the subnets corresponds to the neural network shown in Fig. 1.

- Earlier NNPs are limited to low-dimensional PESs since the order of input coordinates were not arbitrary (any change in the order of atoms change the energy) and a trained model can only be used for the molecule it was trained.
- BP-NNPs represent total energy as a sum of atomic contributions inspired by the empirical potentials.

$$E = \sum_i E_i.$$

- Introduced atom centered symmetry functions to describe local environment.

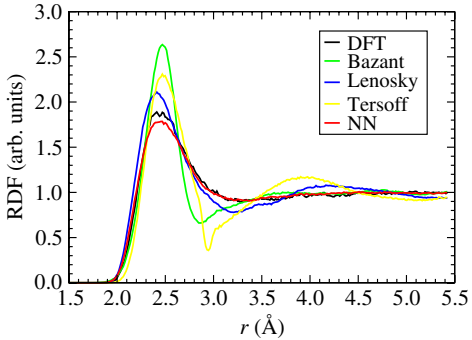
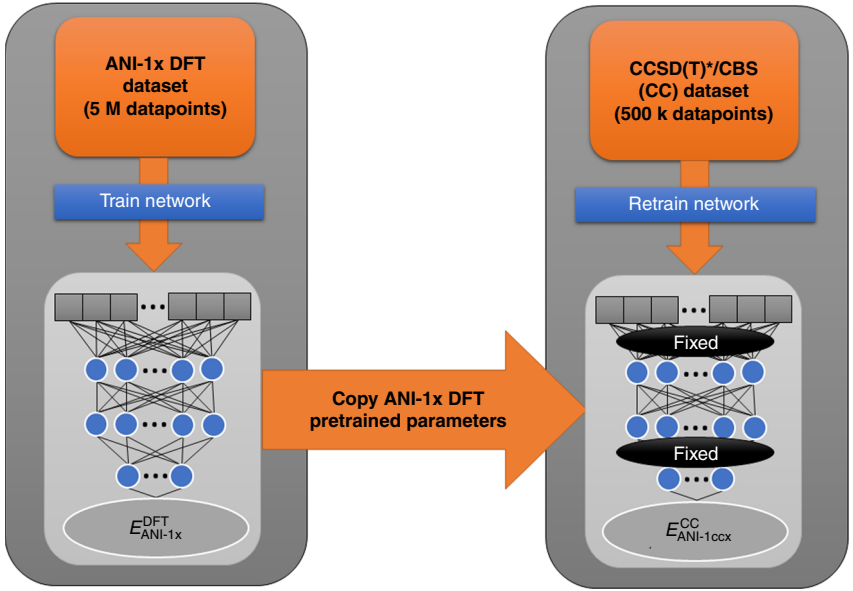


FIG. 3 (color online). Radial distribution function (RDF) of a silicon melt at 3000 K as obtained using a cubic 64 atom cell ($a = 20.526$ bohr). The curves shown were obtained from the Bazant [17,19], the Lenosky [15,19], the Tersoff [16,20], a neural network (NN) potential, and from density-functional theory (DFT) [18].

Transfer learning with BP-NNPs: ANI-1ccx

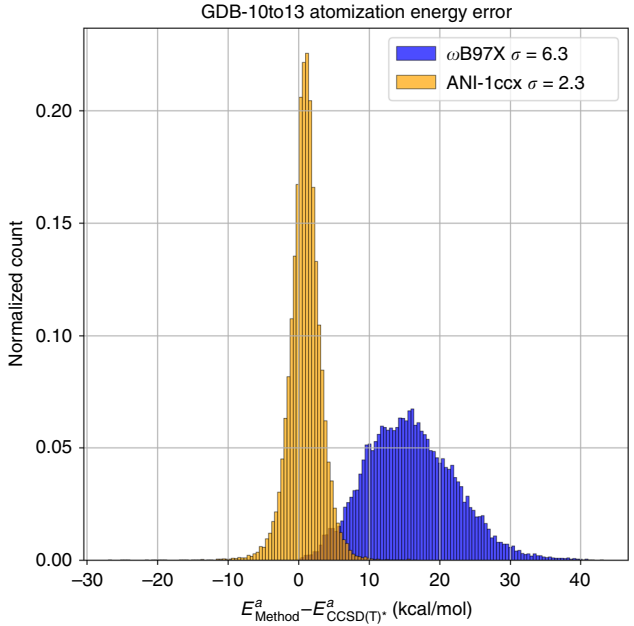
- Train NNP initially with 5M DFT calculations. After freezing some parameters, they retrain with a smaller dataset (500k) DLPNO-CCSD(T)/CBS calculations.

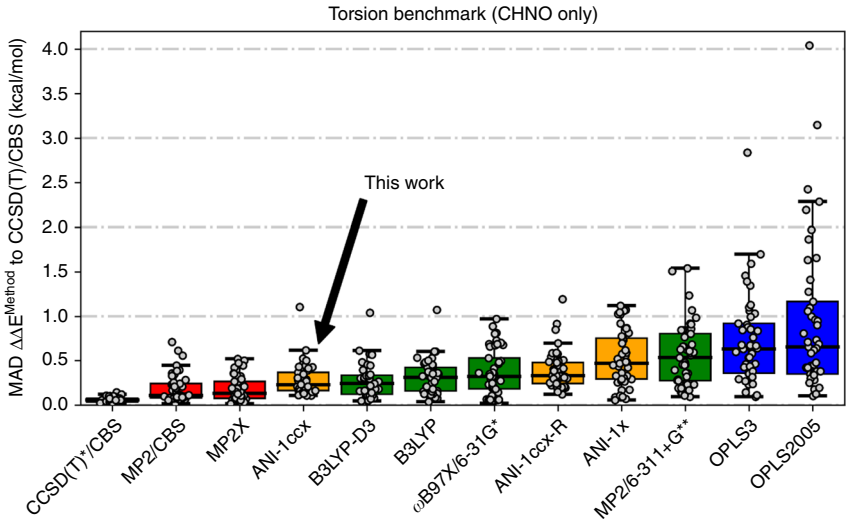


Smith, J. S.; Nebgen, B. T.; Zubatyuk, R.; Lubbers, N.; Devereux, C.; Barros, K.; Tretiak, S.; Isayev, O.; Roitberg, A. E. Approaching Coupled Cluster Accuracy with a General-Purpose Neural Network Potential through Transfer Learning. Nat. Commun. 2019, 10 (1), 2903.

ANI-1ccx performance

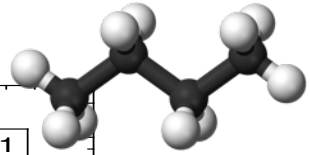
- Training a single ANI model to the original 5.2 million molecule data set takes ~4 h on a NVIDIA Titan V GPU, while retraining to the 500k molecule DLPNO-CCSD(T)/CBS data set takes around 30 min.





Smith, J. S.; Nebgen, B. T.; Zubatyuk, R.; Lubbers, N.; Devereux, C.; Barros, K.; Tretiak, S.; Isayev, O.; Roitberg, A. E. Approaching Coupled Cluster Accuracy with a General-Purpose Neural Network Potential through Transfer Learning. Nat. Commun. 2019, 10 (1), 2903.

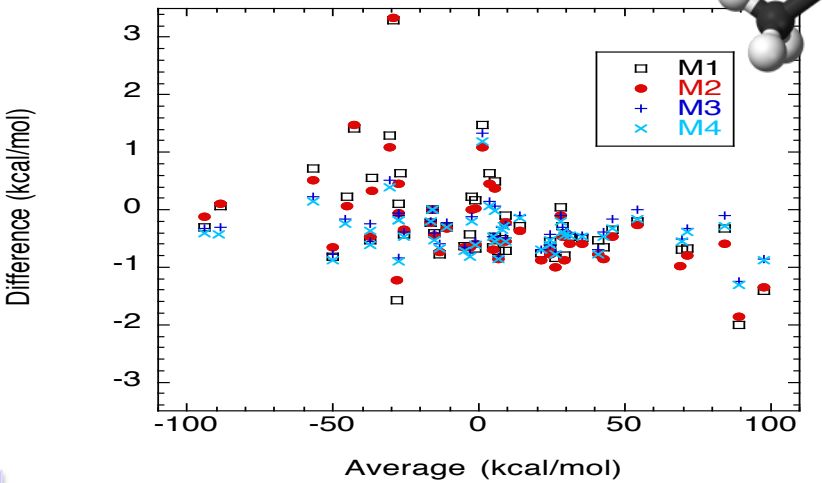
Heats of Formations for Butane Combustion



- Butane combustion model generated by RMG.
- Up to 87 electrons, 22 atoms (11 heavy), 6 rotors.

Comparison of different methods

	ATcT	ANL0
M1	0.72	0.84
M2	0.70	0.83
M3	0.47	0.51
M4	0.50	0.58



Comparison of DFT vs ANI results

Multiplicity	Count	MAD (kcal/mol)	Rel. Diff. (%)
1	143	1.25	3.37
2	153	2.27	8.70
3	16	2.22	14.14

M0 = torsscan/m062x/tz, opt/b2plypd3/tz, freq/b2plypd3/tz
M1 = M0, CCSD(T)-F12/cc-pVDZ-F12
M2 = M0, CCSD(T)-F12/cc-pVDZ-F12 + MP2/cc-pVTZ-F12 – MP2/cc-pVDZ-F12
M3 = M0, CCSD(T)-F12/cc-pVTZ-F12
M4 = M0, CCSD(T)-F12/cc-pVTZ-F12 + MP2/cc-pVQZ-F12 – MP2/cc-pVTZ-F12

1. B. Ruscic, R.E. Pinzon, G. von Laszewski, D. Kodeboyina, A. Burcat, D. Leahy, D. Montoya, A.F. Wagner, J. Phys. Conf. Ser. 16 (2005) 561-570
2. S.J. Klippenstein, L.B. Harding, B. Ruscic, J. Phys. Chem. A 121 (2017) 6580-6602.
3. M. Keçeli, S. N. Elliott, Y.-P. Li, M. S. Johnson, C. Cavallotti, Y. Georgievskii, W. H. Green, M. Pelucchi, J. M. Wozniak, A. W. Jasper, and S. J. Klippenstein, Proc. Combust. Inst. 37, 363 (2019).

SchNet

- End-to-end continuous filter convolutional neural network that can learn a representation directly from atom types and positions.
- Reflects fundamental physical laws including invariance to atom indexing and translation, a smooth energy prediction w.r.t. atom positions as well as energy-conservation of the predicted force fields.
- Can be trained on both energies and forces.
- Predicted energies are rotationally invariant and force predictions are rotationally equivariant.

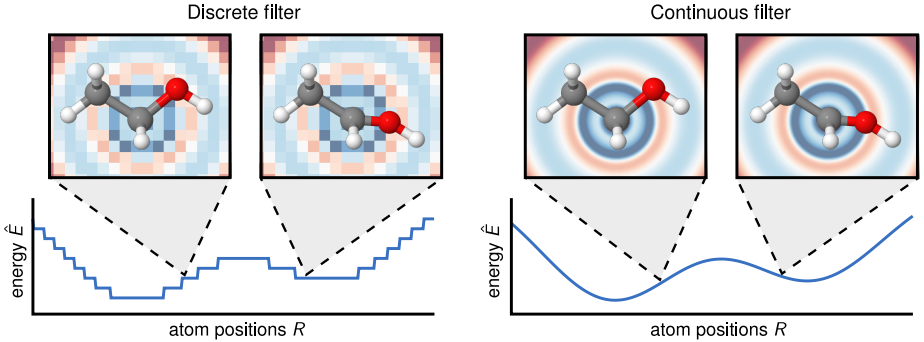


Figure 1: The discrete filter (left) is not able to capture the subtle positional changes of the atoms resulting in discontinuous energy predictions $\hat{E}$ (bottom left). The continuous filter captures these changes and yields smooth energy predictions (bottom right).

Schütt, K. T.; Kindermans, P.-J.; Sauceda, H. E.; Chmiela, S.; Tkatchenko, A.; Müller, K.-R. SchNet: A Continuous-Filter Convolutional Neural Network for Modeling Quantum Interactions. NeurIPS 2017, http://arxiv.org/abs/1706.08566

Schütt, K. T.; Sauceda, H. E.; Kindermans, P. J.; Tkatchenko, A.; Müller, K. R. SchNet - A Deep Learning Architecture for Molecules and Materials. J. Chem. Phys. 2018, 148 (24).

SchNet Architecture

- Atom embedding:
  - Atoms are represented using a tuple of features:

    $X^l = (\mathbf{x}_1^l, \dots \mathbf{x}_n^l)$,

    with $\mathbf{x}_i^l \in \mathbb{R}^F$
    - *F:* number of features
    - *l*: layer index
    - *n*: number of atoms
- Atom-wise layer:
  - Dense layers that are applied separately to the representations:

    $$\mathbf{x}_i^{l+1} = W^l \mathbf{x}_i^l + \mathbf{b}^l$$
- Interaction layers:
  - Add refinements to the atom representation based on pairwise interactions with the surrounding atoms.

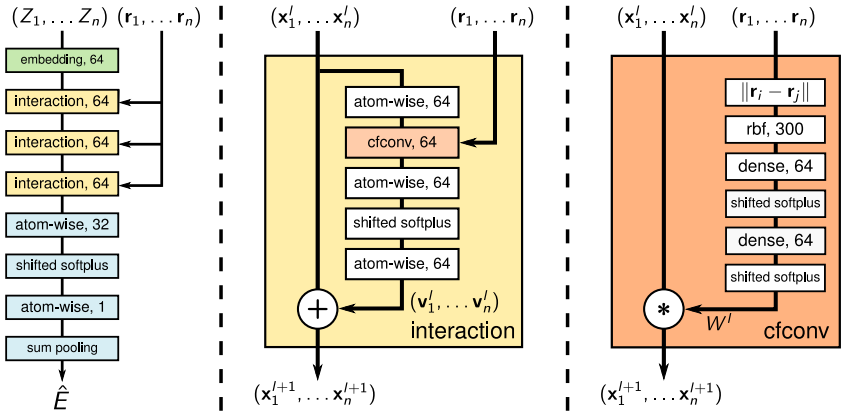


Figure 2: Illustration of SchNet with an architectural overview (left), the interaction block (middle) and the continuous-filter convolution with filter-generating network (right). The shifted softplus is defined as $\text{ssp}(x) = \ln(0.5e^x + 0.5)$.

Schütt, K. T.; Kindermans, P.-J.; Sauceda, H. E.; Chmiela, S.; Tkatchenko, A.; Müller, K.-R. SchNet: A Continuous-Filter Convolutional Neural Network for Modeling Quantum Interactions. NeurIPS 2017, http://arxiv.org/abs/1706.08566

Schütt, K. T.; Sauceda, H. E.; Kindermans, P. J.; Tkatchenko, A.; Müller, K. R. SchNet - A Deep Learning Architecture for Molecules and Materials. J. Chem. Phys. 2018, 148 (24).

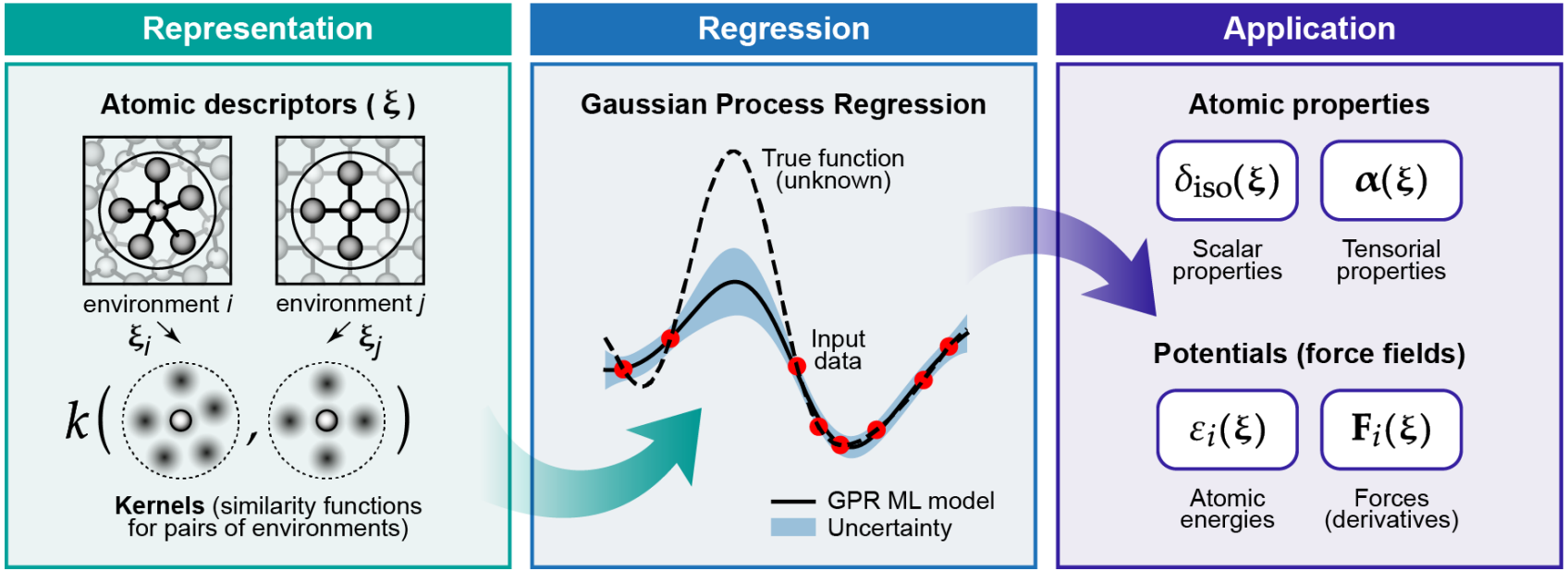
SchNet Performance

property	unit	model	MAE	RMSE	time
		Dataset: Acetylsalicylic acid (N = 50k)			
energy	kcal mol^{-1}	SchNet	0.11	0.14	2 d, 11.5 h
energy	kcal mol^{-1}	ACSF	0.40	0.53	1 d, 6 h
energy	kcal mol^{-1}	wACSF	1.20	2.69	1 d, 6 h
atomic forces	kcal mol^{-1} Å^{-1}	SchNet	0.14	0.19	2 d, 11.5 h
atomic forces	kcal mol^{-1} Å^{-1}	ACSF	0.88	1.26	1 d, 6 h
atomic forces	kcal mol^{-1} Å^{-1}	wACSF	2.31	3.14	1 d, 6 h
		Dataset: QM9 (N = 110k)			
U_0	kcal mol^{-1}	SchNet	0.26	0.54	12 h
U_0	kcal mol^{-1}	ACSF	0.49	0.92	8 h
U_0	kcal mol^{-1}	wACSF	0.43	0.81	6 h
dipole moment	Debye	SchNet	0.020	0.038	13 h
dipole moment	Debye	ACSF	0.064	0.100	8 h
dipole moment	Debye	wACSF	0.064	0.095	8 h
		Dataset: ANI-1 (N = 10.1M)			
energy	kcal mol^{-1}	SchNet	0.55	0.89	9 d, 7 h[b]
		Dataset: ANI-1 (N = 19.8M)			
energy	kcal mol^{-1}	SchNet	0.47	0.77	12 d, 15 h[c]
		Dataset: Materials Project (N = 62k)			
formation energy	eV/atom	SchNet	0.041	0.088	1 d, 14 h

[a] N is used to denote the size of the combined training set and validation set. [b] Four Tesla P100 GPUs were used for data-parallel training. [c] Two Tesla P100 GPUs were used for data-parallel training.

Schütt, K. T et.al SchNetPack: A Deep Learning Toolbox For Atomistic Systems. *J. Chem. Theory Comput.* **2019**, *15* (1), 448–455.

Gaussian Approximation Potentials



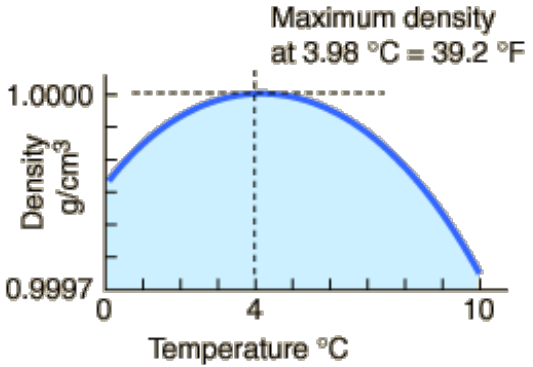
Bartók, A. P.; Payne, M. C.; Kondor, R.; Csányi, G. Gaussian Approximation Potentials: The Accuracy of Quantum Mechanics, without the Electrons. Phys. Rev. Lett. 2010, 104 (13), 136403.

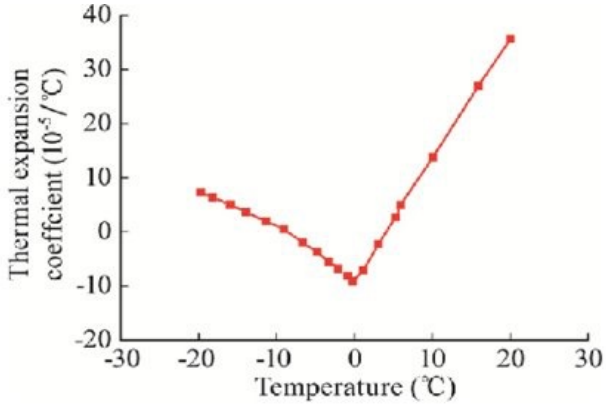
Deringer, V. L.; Bartók, A. P.; Bernstein, N.; Wilkins, D. M.; Ceriotti, M.; Csányi, G. Gaussian Process Regression for Materials and Molecules. Chem. Rev. 2021, 121 (16), 10073–10141.

Thermal Expansion

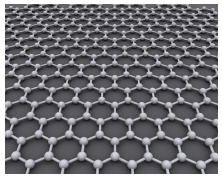
- Change in temperature results in change in volume under constant pressure.
- Consequence of anharmonicity, no thermal expansion in harmonic approximation.
- Thermal expansion coefficient is a measure of the volume change of a material in response to a temperature change.

$$\alpha = \frac{1}{V}\left(\frac{\partial V}{\partial T}\right)_P$$

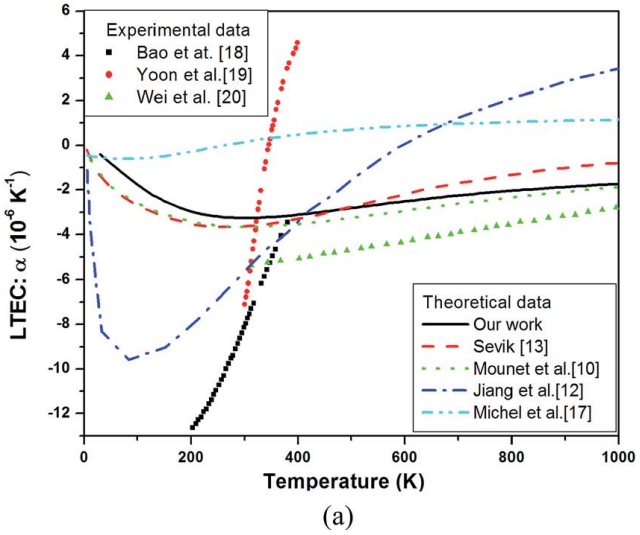




Thermal Expansion Coefficient of Graphene



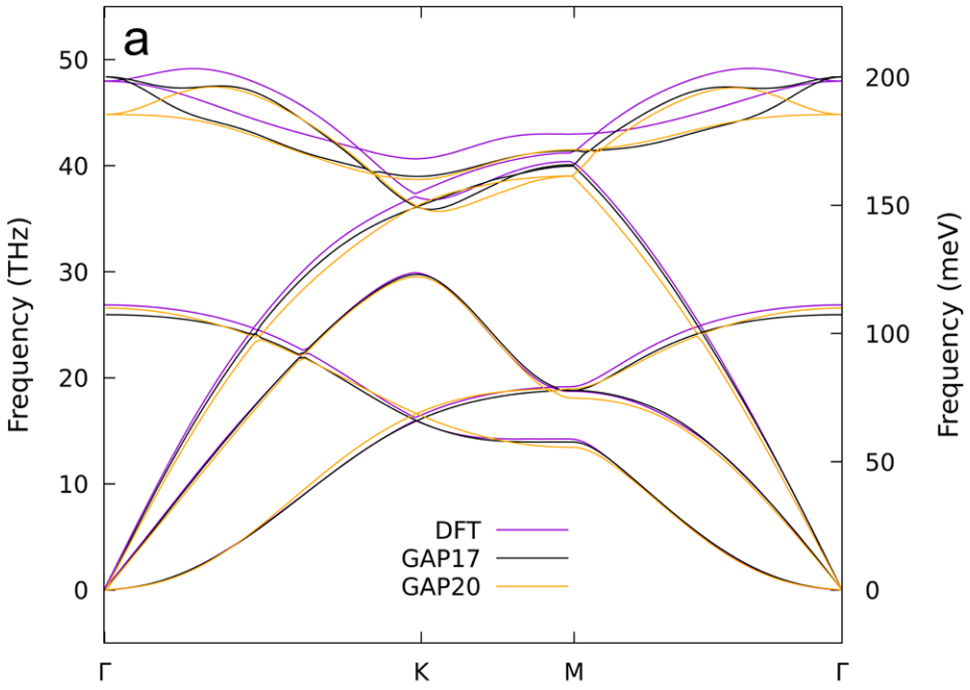
- Graphene has a negative TEC, but there is an uncertainty on the details of the temperature dependence.
- Computing TEC is difficult since it requires incorporating anharmonic effects with large supercell simulations.
- One can compute TEC with three different methods:
  - Quasi Harmonic Approximation (QHA)
  - Grüneisen parameters
  - Molecular Dynamics
- We tested two different GAP potentials and compared them with DFT calculations. GAP17 trained for Graphene only, GAP20 trained for all Carbon allotropes.



(a)

Plot from Mann, Kumar and Jindal RSC Adv., 2017, 7, 22378.
Collaboration with İlker Demiroğlu, Yenal Karaaslan, Tuğbey Kocabaş, and Cem Sevik from Eskisehir Technical University in Turkey and Álvaro Vázquez-Mayagoitia.

Phonon Dispersion Curves with DFT and GAP



	MAD (meV)	MAX (meV)
GAP17	2.73	8.05
GAP20	4.14	13.03

Table 1. Average Simulation Times for Simulations Performed with GAP Models and DFT on the Same Computer Architecture

method	system size	no. of steps	no. of CPU	total time (h)	time per step (s)
DFT	112 atom	10 000	28 (1 node)	71	26
DFT	240 atom	10 000	56 (2 nodes)	170	61
DFT	336 atom	10 000	84 (3 nodes)	293	106
DFT	448 atom	10 000	112 (4 nodes)	574	207
GAP	448 atom	500 000	28 (1 node)	18	0.2
GAP	1008 atom	500 000	28 (1 node)	38	0.4
GAP	2232 atom	500 000	28 (1 node)	84	0.7
GAP	4032 atom	500 000	28 (1 node)	150	1.1
GAP	6240 atom	500 000	28 (1 node)	240	1.7

Demiroğlu, İ.; Karaaslan, Y.; Kocabaş, T.; Keçeli, M.; Vázquez-Mayagoitia, Á.; Sevik, C. *J. Phys. Chem. C* **2021**, *125* (26), 14409–14415.

Conclusions and Outlook

- Gaussian approximation and neural network potentials are very promising candidates to accelerate material science simulations and extend the domain of predictive simulations.
- The field is growing rapidly in terms of method and code development, but applications to challenging problems are still rare.
- Predicting properties directly or using RL to design materials are other promising application

- Challenges and opportunities:
  - More accurate and diverse data sets are required for training.
  - Automating the workflow for sampling, neural architecture search, hyperparameter optimization, and training.
  - Number of different elements that can be trained is still limited
  - Modeling long-range interactions

Recommendations

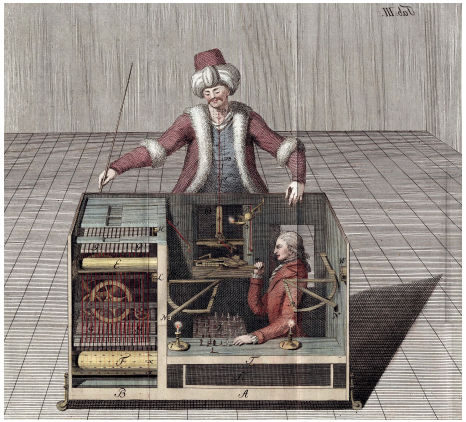
- Master python
  - numpy, scipy for numerical methods
  - pandas, scikit-learn for data analysis and ML
  - TensorFlow or PyTorch for deep learning
  - ASE for atomistic simulations
  - Open Babel, RDKit for cheminformatics
  - DScribe for descriptors
- Many free resources online to learn ML. Kaggle competitions can be fun.
- Use Jupyter Lab/Notebook for testing, data visualization, and code development
- Use a computer with GPU for training & inference with DL models
- Open a GitHub account and actively use it for code development
- When you are stuck don’t be shy to post questions on GitHub issues or StackOverflow. Provide a minimal example to reproduce your problem.
- Visual Studio Code is a smart IDE that can help with code development
- Learn using Docker to easily deploy different applications.

THANK YOU!
QUESTIONS OR COMMENTS?

https://keceli.github.io/
https://github.com/keceli/
twitter @ hpc4science

What is Machine Learning?

- AI is defined broadly as the study of systems that can perceive their environment and take actions to achieve their goals.
  - Formal design of Turing-complete artificial neurons by McCullouch and Pitts in 1943 is considered as the first AI work.
- ML is the study of models that can learn patterns from data and make predictions without being explicitly programmed.
  - Original definition by Arthur Samuel in 1959.



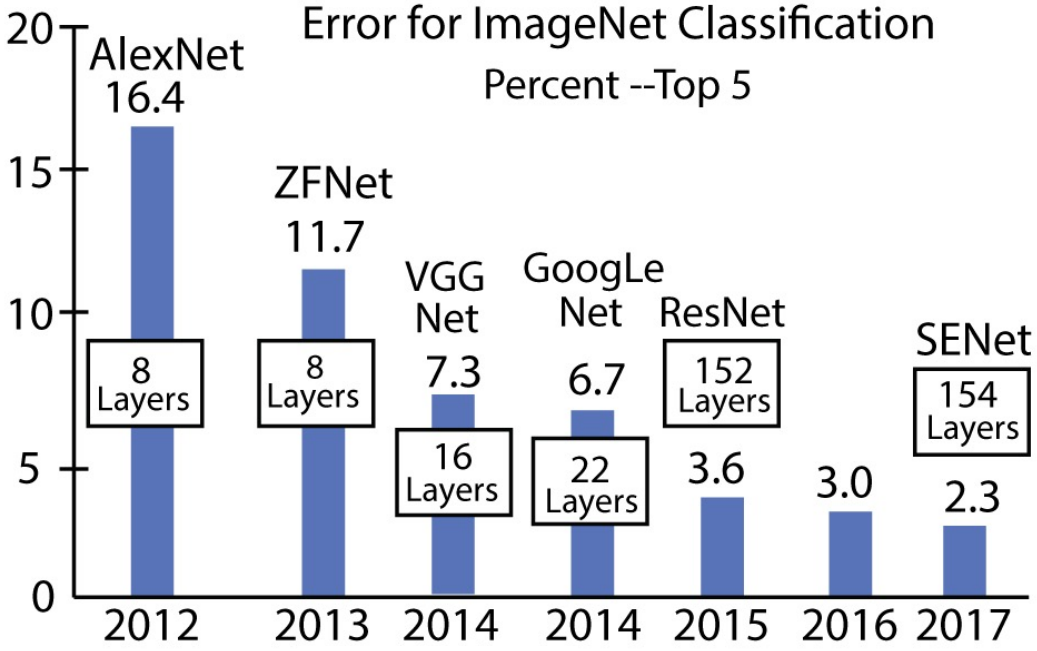
The Turk, 1770

Deep Blue, 1996

DeepMind

AlphaZero, 2017

Deeper


Error for ImageNet Classification
Percent --Top 5
20
15
10
5
0
AlexNet
16.4
8
Layers
ZFNet
11.7
8
Layers
VGG
Net
7.3
16
Layers
GoogLe
Net
6.7
22
Layers
ResNet
152
Layers
3.6
SENet
154
Layers
3.0
2.3
2012
2013
2014
2014
2015
2016
2017


I WAS WINNING IMAGENET
UNTIL A DEEPER MODEL CAME ALONG
imgflip.com

Weighted atom-centered symmetry functions

- The number of ACSFs grows quadratically with the number of different chemical species. This can lead to an impractical number of ACSFs for systems containing more than four elements.
- wACSFs introduce weight functions that depend on element type to circumvent this problem.

Radial ACSFs:

$$G_i^{\mathrm{rad}} = \sum_{j \neq i}^{N} e^{-\eta(r_{ij}-\mu)^2} f_{\mathrm{c}}(r_{ij}),$$

Radial wACSFs:

$$W_i^{\mathrm{rad}} = \sum_{j \neq i}^{N} g(Z_j) e^{-\eta(r_{ij}-\mu)^2} f_{ij}$$

wACSF—Weighted atom-centered symmetry functions as descriptors in machine learning potentials

M. Gastegger, L. Schwiedrzik, M. Bittermann, F. Berzsenyi, and P. Marquetand[a)]
Institute of Theoretical Chemistry, Faculty of Chemistry, University of Vienna, Währinger Str. 17, 1090 Vienna, Austria

(Received 15 December 2017; accepted 29 January 2018; published online 15 March 2018)

We introduce weighted atom-centered symmetry functions (wACSFs) as descriptors of a chemical system's geometry for use in the prediction of chemical properties such as enthalpies or potential energies via machine learning. The wACSFs are based on conventional atom-centered symmetry functions (ACSFs) but overcome the undesirable scaling of the latter with an increasing number of different elements in a chemical system. The performance of these two descriptors is compared using

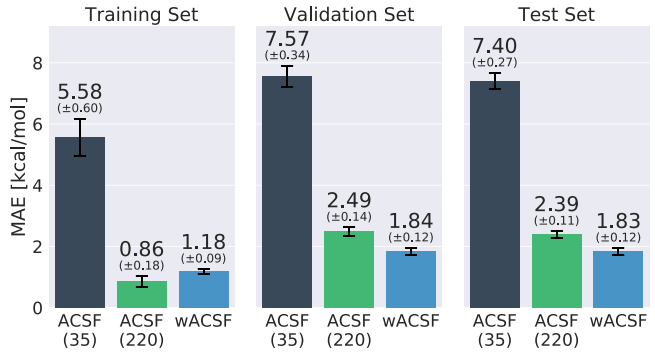
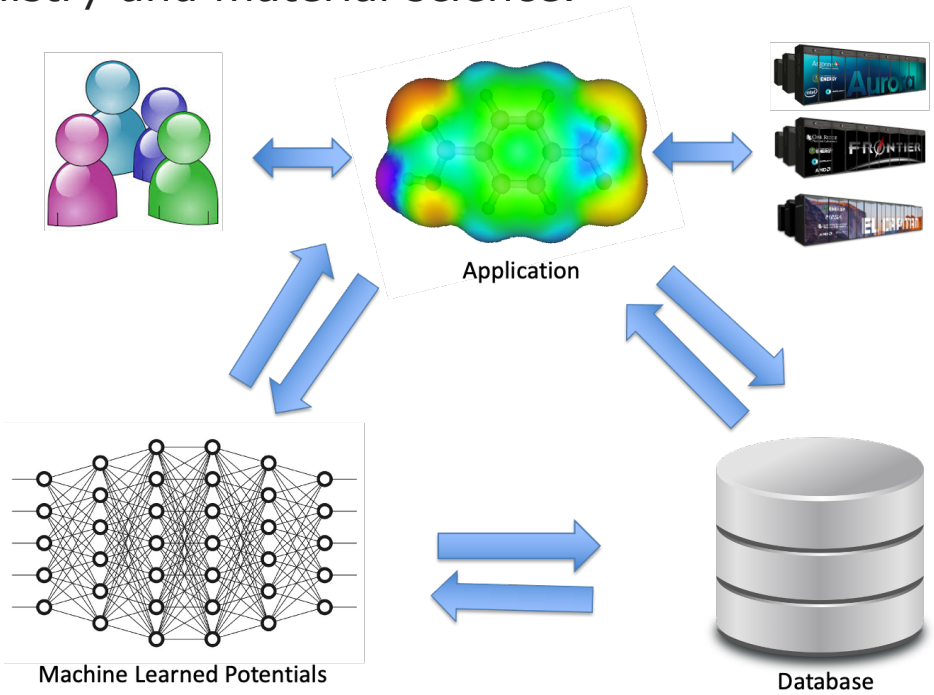


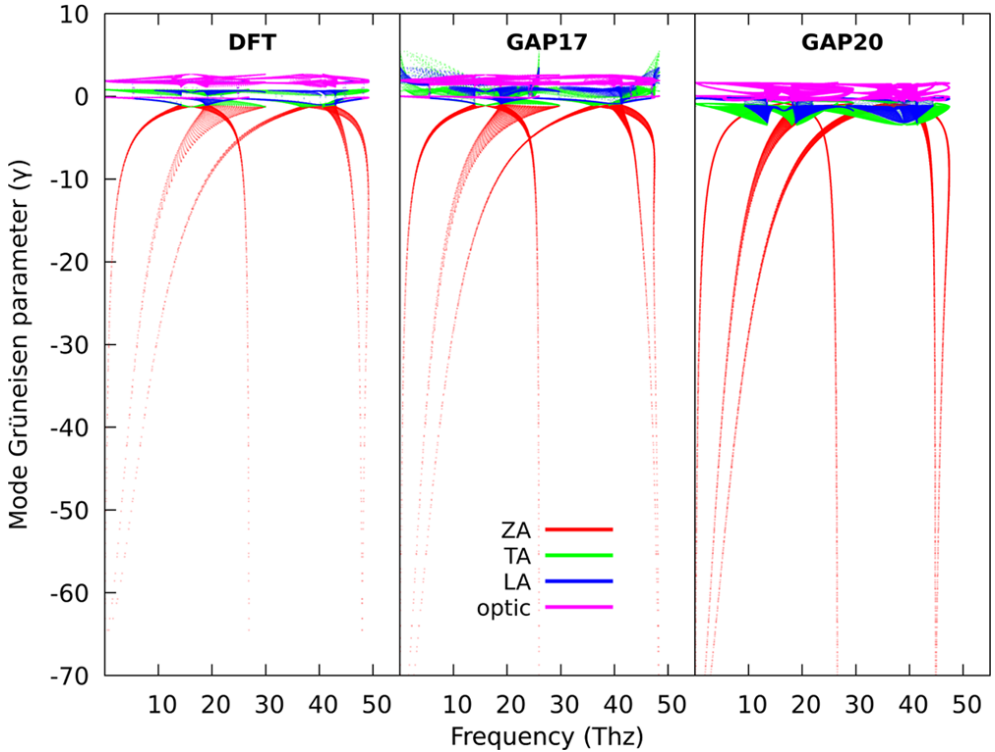
FIG. 3. MAEs obtained for ACSF and wACSF type descriptors. In the case of the wACSF-based model, 32 symmetry functions are used to describe the chemical environment of each element (26 radial and 6 angular functions). The error bars and values in brackets correspond to the standard deviations obtained via the 5-fold cross validation procedure. For the ACSFs, two different descriptor vectors are shown, one using a minimal set of 35 symmetry functions (5 radial and 30 angular), while the other one uses the same spatial resolution as the wACSF vector, leading to a total of 220 symmetry functions (130 radial and 90 angular). Note that better prediction errors may be achieved with larger descriptor vectors and larger NNs, but this was not the goal of this study.

What can we do?

- Ever-increasing computing power, simulation data, and machine learning tools.
- Coupling machine learning and simulations is the key to solve challenging problems in chemistry and material science.



Mode Grüneisen Parameters

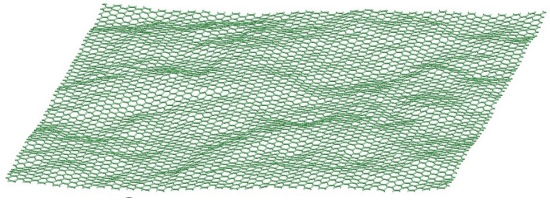


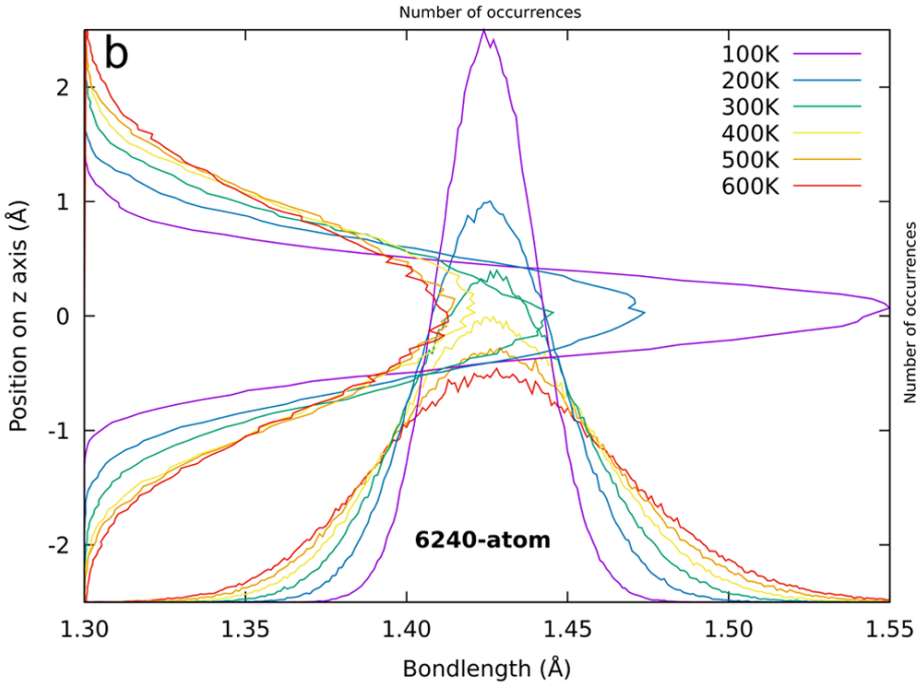
$$\gamma_{q,j} = \frac{-a_0}{\omega_{q,j}} \frac{\partial \omega_{q,j}}{\partial a}$$

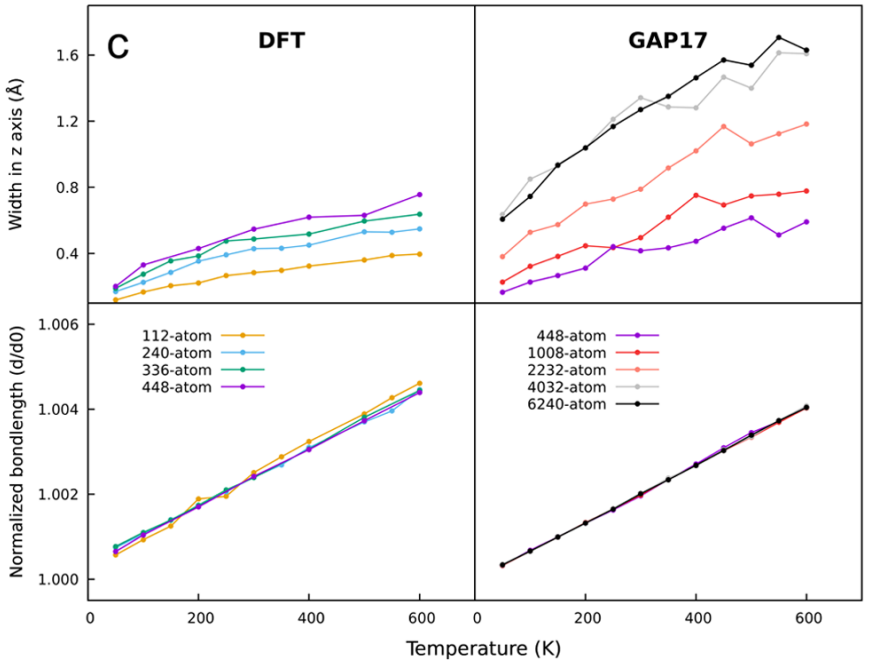
$$\alpha = \frac{\gamma_{\text{mean}} C^V}{B_T V_0} \Bigg|_{a,T}$$

Demiroğlu, İ.; Karaaslan, Y.; Kocabaş, T.; Keçeli, M.; Vázquez-Mayagoitia, Á.; Sevik, C. *J. Phys. Chem. C* **2021**, *125* (26), 14409–14415.

Structural Analysis

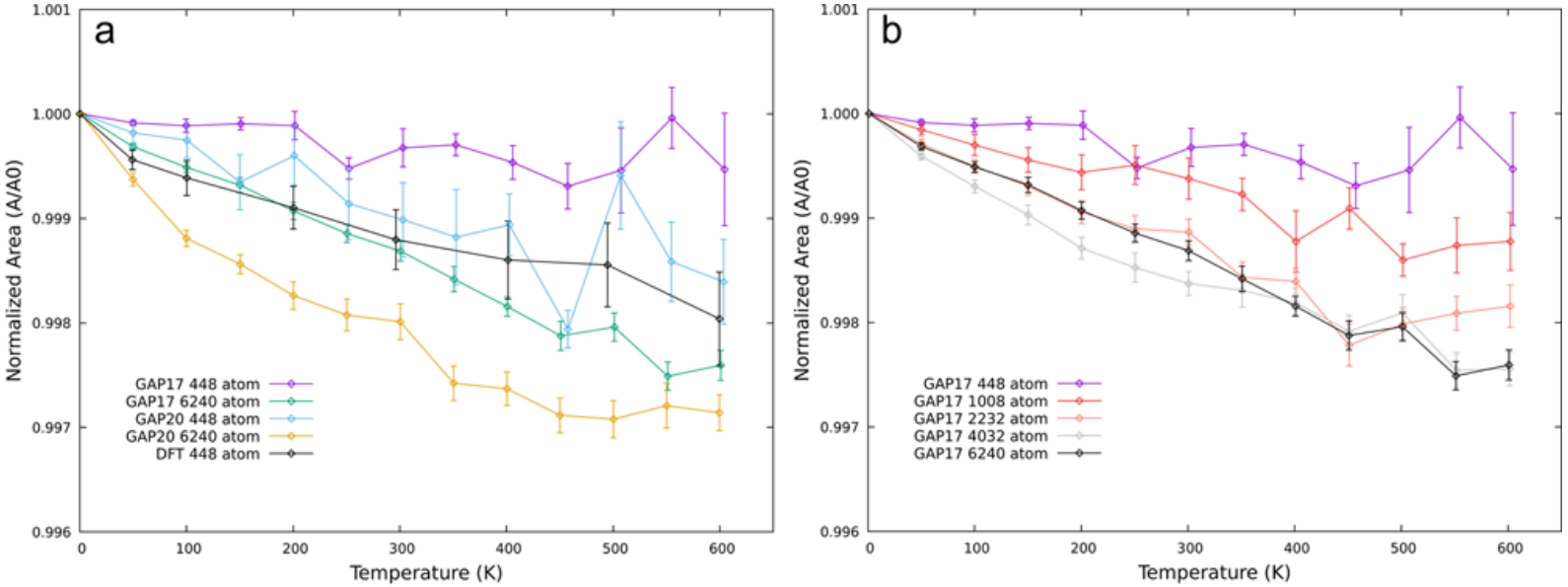




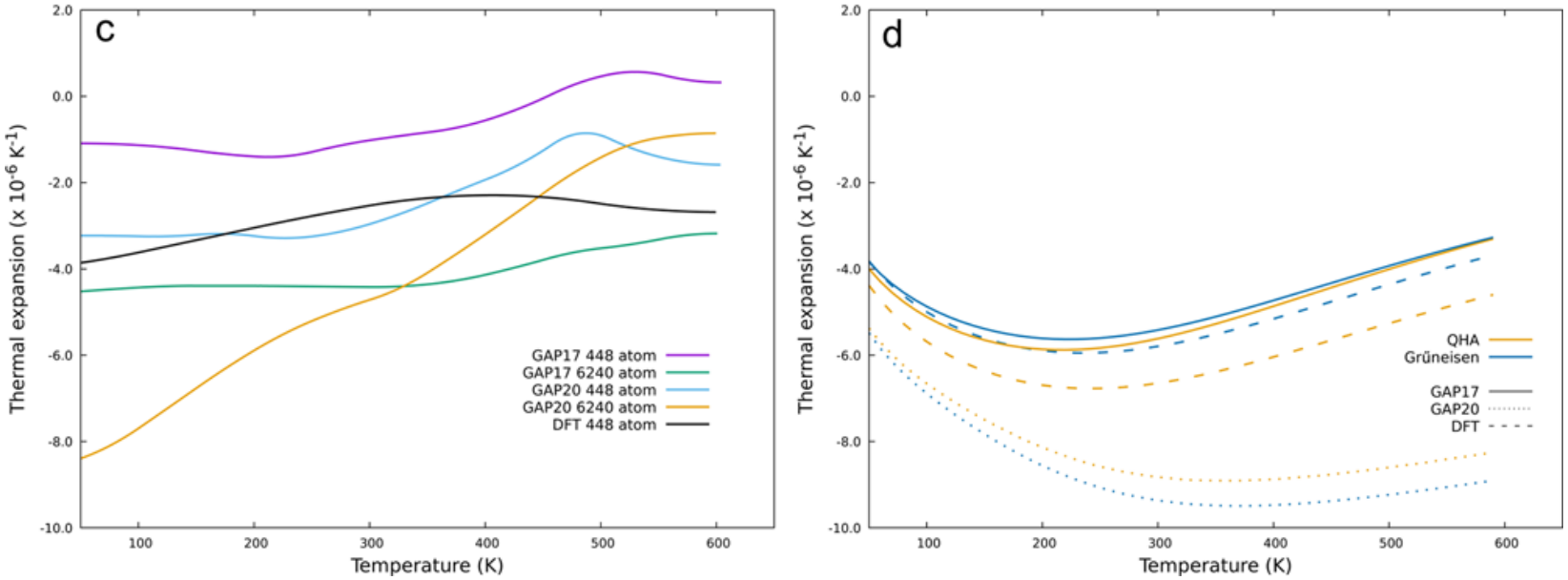


C-C bond length and rippling width increases with temperature.

Normalized area vs Temperature



Thermal Expansion Coefficient



Demiroğlu, İ.; Karaaslan, Y.; Kocabaş, T.; Keçeli, M.; Vázquez-Mayagoitia, Á.; Sevik, C. *J. Phys. Chem. C* **2021**, *125* (26), 14409–14415.