AI for Materials Science

Lars Kotthoff

Artificially Intelligent Manufacturing Center
larsko@uwyo.edu
www.uwyo.edu/aim

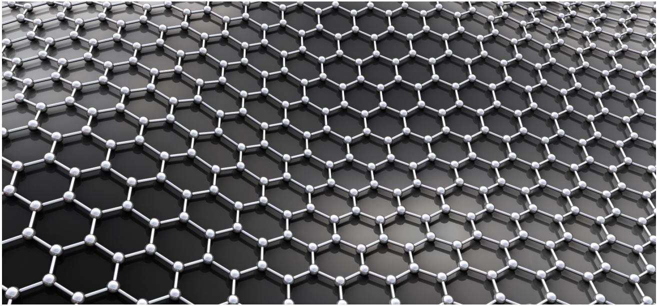
University of Wyoming

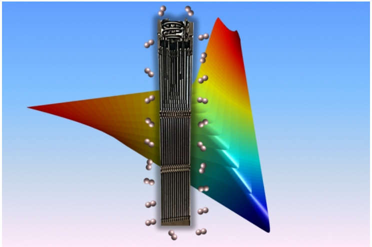
IJCAI, 10 August 2019
https://www.cs.uwyo.edu/~larsko/aimat-tut/

Outline

- Advanced Materials – Examples and Challenges
- Surrogate Models
- Advanced Materials – AI Approaches
- Bayesian Optimization Background
- Bayesian Optimization in Materials Science
- Common Themes in AI and Materials Science
- Challenges and Opportunities

Advanced Materials



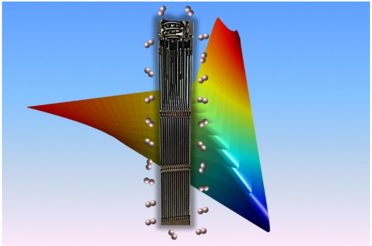


Advanced Materials – OLEDs

- efficiency, lifetime, light output need to be improved
- manufacturing expensive
- LEDs much more mature

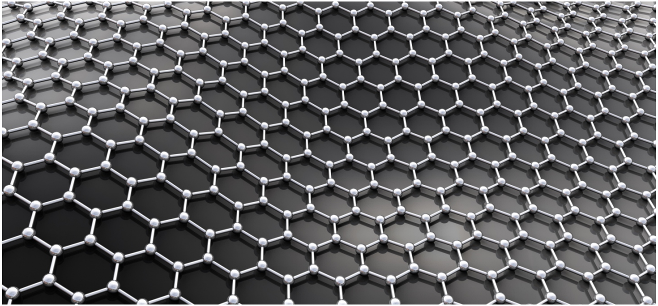
https://www.energy.gov/eere/ssl/oled-rd-challenges

Advanced Materials – Metal Alloys



- existing alloys cannot operate at high temperatures
- properties of new types of alloys not well understood
- advanced manufacturing methods required

Advanced Materials – Graphene



- ▷ scaling up production difficult
- ▷ impurities during synthesis
- ▷ manufacturing expensive

https://www.nature.com/articles/s41563-019-0341-4

Issues to Tackle

- designing and optimizing manufacturing processes
- designing and optimizing materials and their properties
- limited first-principles knowledge

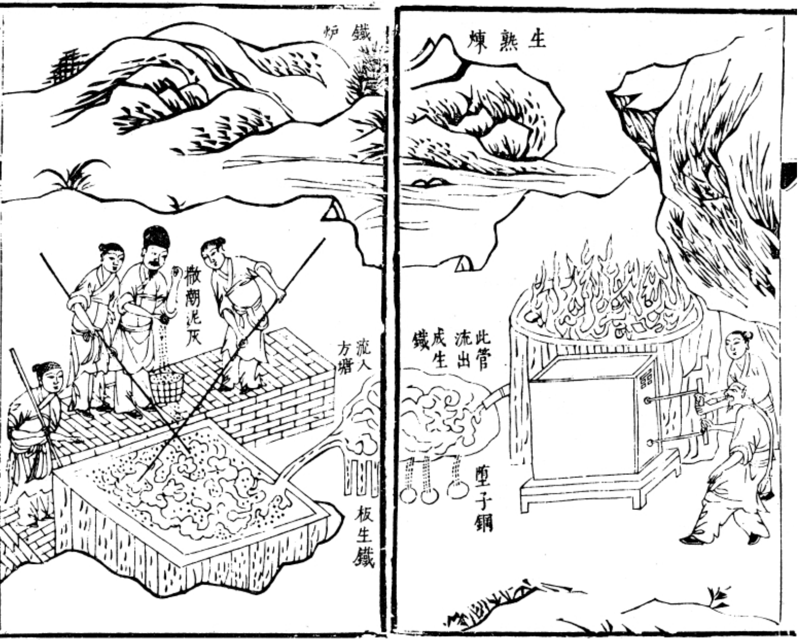
Challenges

- large design space for new materials and processes – composition, structure, manufacturing steps...
- often multiple, competing objectives
- expensive to synthesize and test

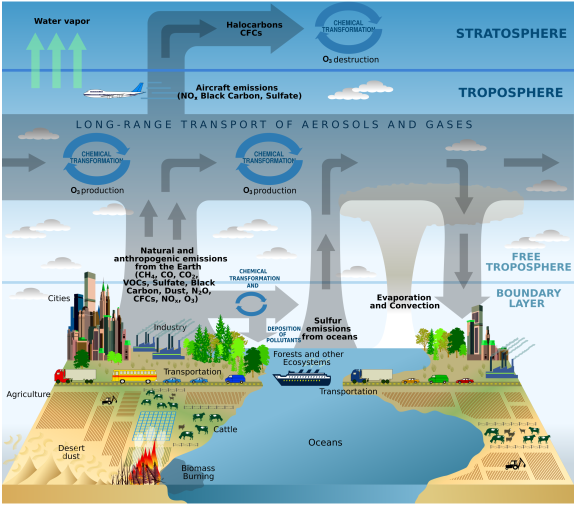
Challenges

Challenges – Sound Familiar?

- large design space for new AI approaches, ML pipelines...
- often multiple, competing objectives
- expensive to test


生熟煉
鐵炉
撒潮泥灰
流入方塘
板生鐵
此管流出成生鐵
堕子鋼

Modeling



- ▷ build models based on observations and theories
- ▷ use models to make predictions

---

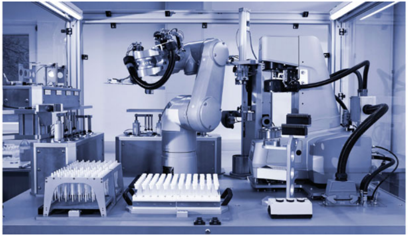
https://en.wikipedia.org/wiki/Scientific_modelling

Surrogate Models – Experiments

- around for thousands of years
- takes minutes to weeks
- ground-truth results

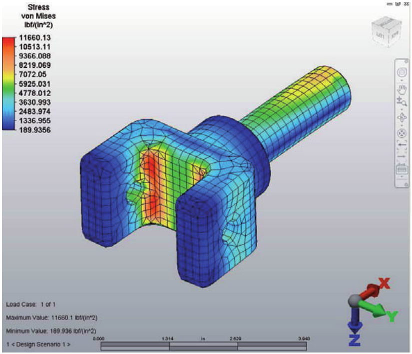
Surrogate Models – High-Throughput Experimentation

- around for tens of years
- takes seconds to days
- ground-truth results
- expensive and complex setups



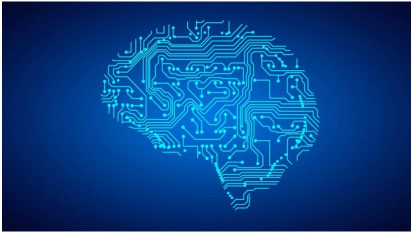
Surrogate Models – Computational Simulations

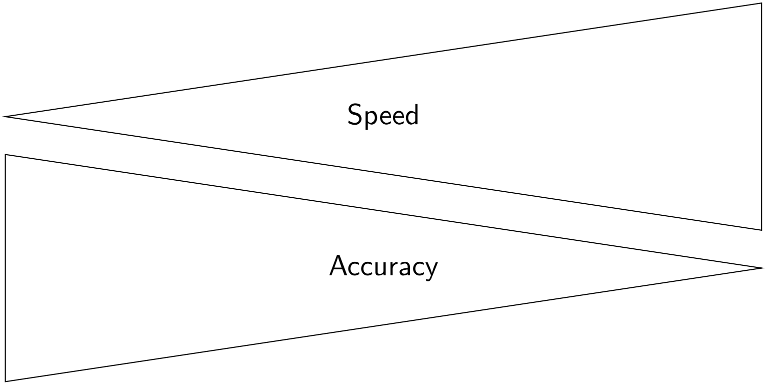
- developed since 1940s
- takes seconds to days
- results based on mathematical models that encapsulate our understanding of fundamental processes
- no expensive/dangerous/bulky experimental setup

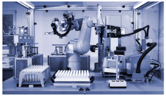


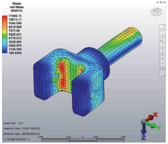
Surrogate Models – Machine Learning

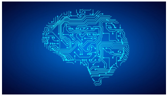
- started $\approx$20 years ago
- takes seconds
- approximate results based on statistical correlations

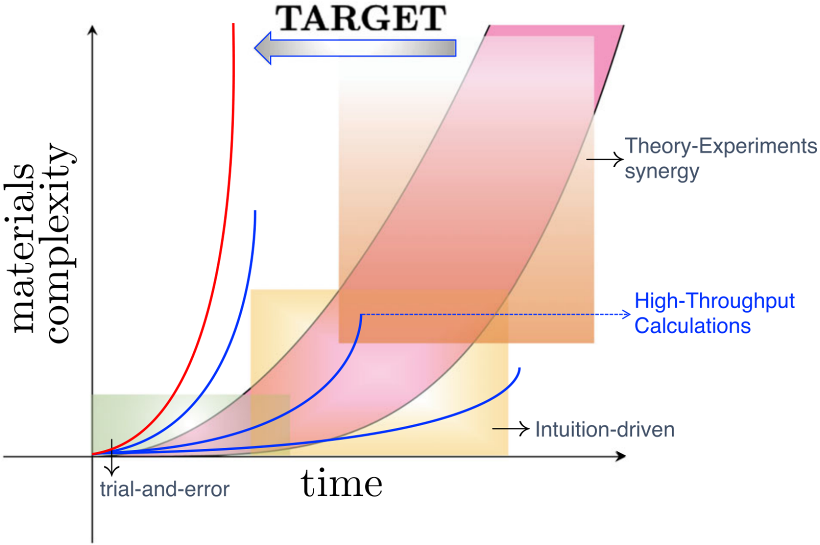



Speed
Accuracy





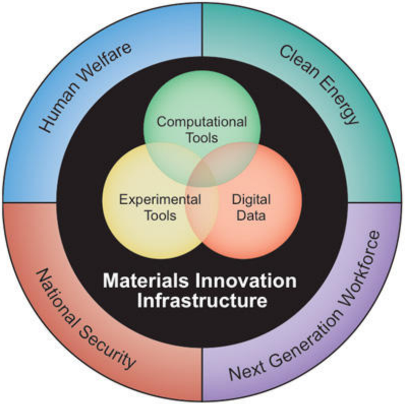




Lookman, Turab, Prasanna V. Balachandran, Dezhen Xue, and Ruihao Yuan. "Active Learning in Materials Science with Emphasis on Adaptive Sampling Using Uncertainties for Targeted Design." Npj Computational Materials 5, no. 1 (February 18, 2019): 21. https://doi.org/10.1038/s41524-019-0153-8.

Materials Genome Initiative

- launched in 2011 to "discover, manufacture, and deploy advanced materials twice as fast, at a fraction of the cost"
- US agencies and international partners
- AI, machine learning, and computation play central role



https://www.mgi.gov/

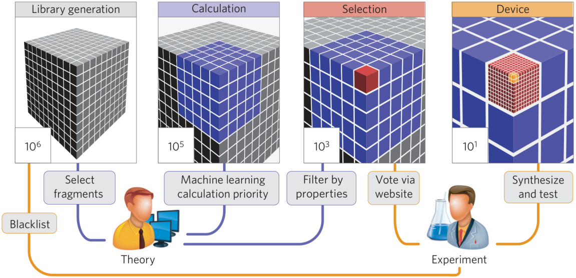
Materials Genome Initiative

The Problem	Finding a New Material is Complex, Expensive and Time-Consuming
The Answer	Computers are Good at Complexity

https://www.nist.gov/mgi/about-material-genome-initiative

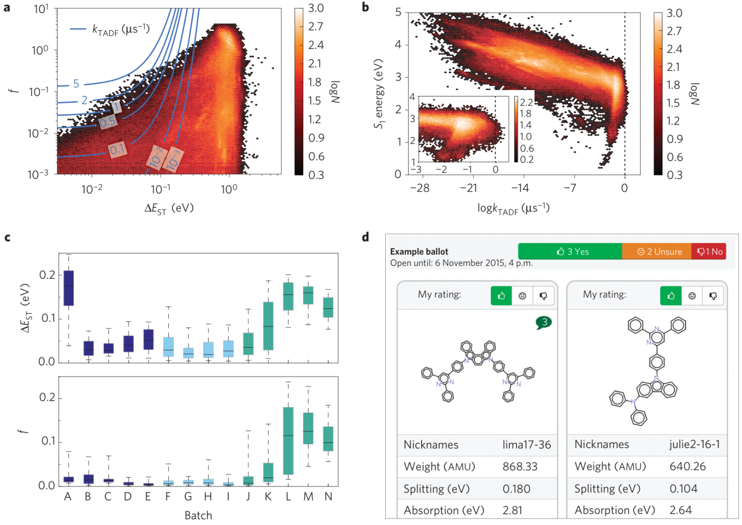
Advanced Materials – OLEDs

- compute properties of candidates with quantum chemical calculations
- machine learning model based on these calculations to pre-screen
- human decision-making on what to synthesize and test
- improve OLED efficiency by 22%



Gómez-Bombarelli, Rafael, Jorge Aguilera-Iparraguirre, Timothy D. Hirzel, David Duvenaud, Dougal Maclaurin, Martin A. Blood-Forsythe, Hyun Sik Chae, et al. "Design of Efficient Molecular Organic Light-Emitting Diodes by a High-Throughput Virtual Screening and Experimental Approach." Nature Materials 15 (August 8, 2016): 1120.

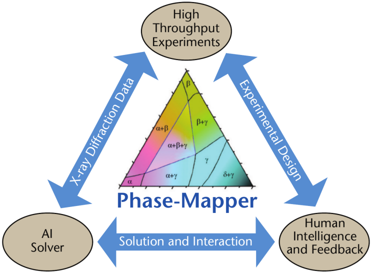
Advanced Materials – OLEDs



Gómez-Bombarelli, Rafael, Jorge Aguilera-Iparraguirre, Timothy D. Hirzel, David Duvenaud, Dougal Maclaurin, Martin A. Blood-Forsythe, Hyun Sik Chae, et al. “Design of Efficient Molecular Organic Light-Emitting Diodes by a High-Throughput Virtual Screening and Experimental Approach.” Nature Materials 15 (August 8, 2016): 1120.

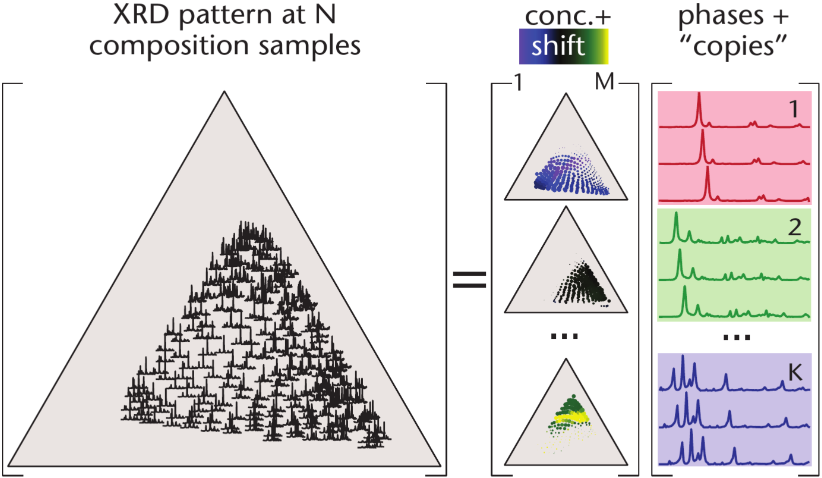
Advanced Materials – Metal Alloys

- Phase-Mapper system – identify crystal structure of materials (specifically metal alloys) from x-ray diffraction (XRD) images
- find combination of basis patterns from observed pattern
- allows to rapidly interpret XRD patterns and identify new materials with desirable properties
- automated previously manual work; speed up from days to minutes



Bai, Junwen, Yexiang Xue, Johan Bjorck, Ronan Le Bras, Brendan Rappazzo, Richard Bernstein, Santosh K. Suram, Robert Bruce van Dover, John M. Gregoire, and Carla P. Gomes. "Phase-Mapper: Accelerating Materials Discovery with AI." AI Magazine 39, no. 1 (2018): 15–26.

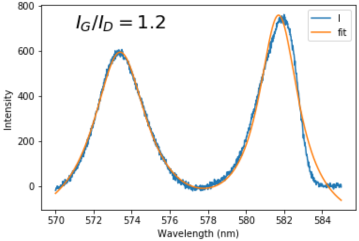
Advanced Materials – Metal Alloys

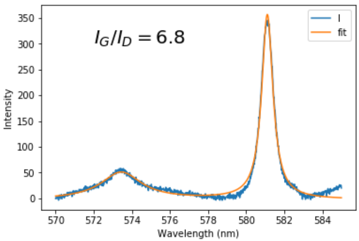


Bai, Junwen, Yexiang Xue, Johan Bjorck, Ronan Le Bras, Brendan Rappazzo, Richard Bernstein, Santosh K. Suram, Robert Bruce van Dover, John M. Gregoire, and Carla P. Gomes. "Phase-Mapper: Accelerating Materials Discovery with AI." AI Magazine 39, no. 1 (2018): 15–26.

Advanced Materials – Graphene

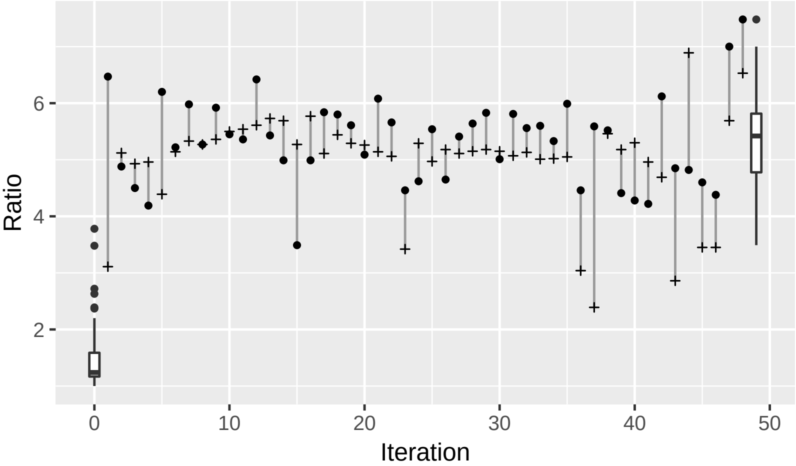
- irradiate graphene oxide film with laser to synthesize graphene
- Bayesian Optimization to tune parameters of laser
- improvement of 2x over best result in literature
- talk tomorrow 14.50h Data Science Meets Optimization Workshop, Florence 2301





Kotthoff, Lars, Vivek Jain, Alexander Tyrrell, Hud Wahab, and Patrick Johnson. "AI for Materials Science: Tuning Laser-Induced Graphene Production." In Data Science Meets Optimisation Workshop at IJCAI 2019, 2019.

Advanced Materials – Graphene



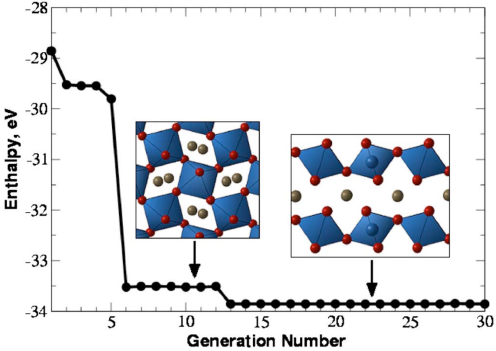
Kotthoff, Lars, Vivek Jain, Alexander Tyrrell, Hud Wahab, and Patrick Johnson. "AI for Materials Science: Tuning Laser-Induced Graphene Production." In Data Science Meets Optimisation Workshop at IJCAI 2019, 2019.

Applications of AI in Materials Science

- predicting (machine learning surrogate models of properties)
- optimizing (matching explanations to observations)
- combinations of the two

Optimize

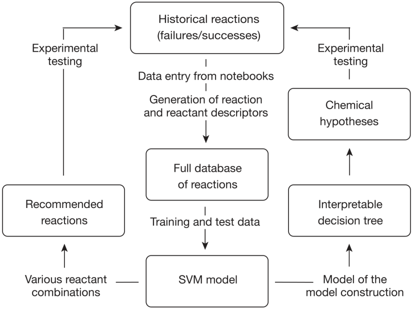
- "predict" crystal structure using genetic algorithms – crystal structure optimized to match desired properties
- evolutionary algorithm to create structures, first-principles computations to compute fitness
- identification of new high-pressure crystal structures



Oganov, Artem R., and Colin W. Glass. "Crystal Structure Prediction Using Ab Initio Evolutionary Techniques: Principles and Applications." The Journal of Chemical Physics 124, no. 24 (2006): 244704. https://doi.org/10.1063/1.2210932.

Predict

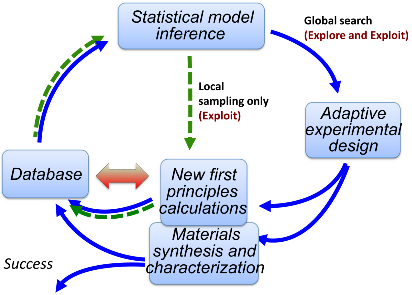
- ▷ SVM model to predict success of chemical reaction
- ▷ fit decision tree to SVM to understand model
- ▷ uses unpublished data from failed experiments
- ▷ better accuracy than humans
- ▷ model "could also be applied to exploration reactions"



Raccuglia, Paul, Katherine C. Elbert, Philip D. F. Adler, Casey Falk, Malia B. Wenny, Aurelio Mollo, Matthias Zeller, Sorelle A. Friedler, Joshua Schrier, and Alexander J. Norquist. "Machine-Learning-Assisted Materials Discovery Using Failed Experiments." Nature 533 (May 4, 2016): 73.

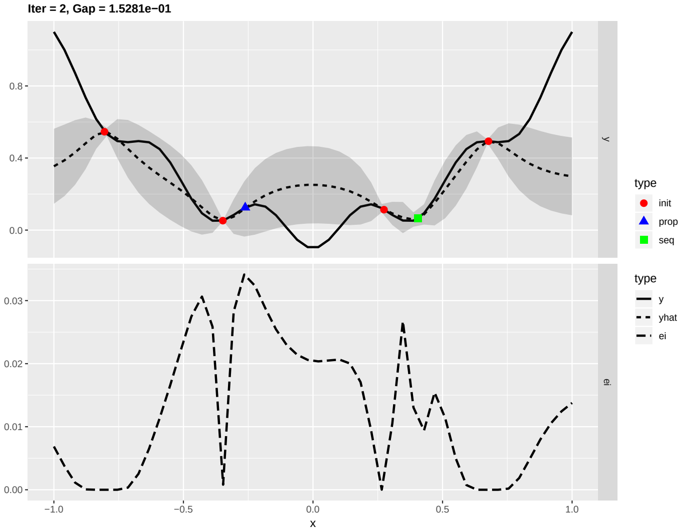
Combination

- surrogate model based on first-principles calculations
- Bayesian Optimization with infill criterion to choose next point to evaluate
- iterative process



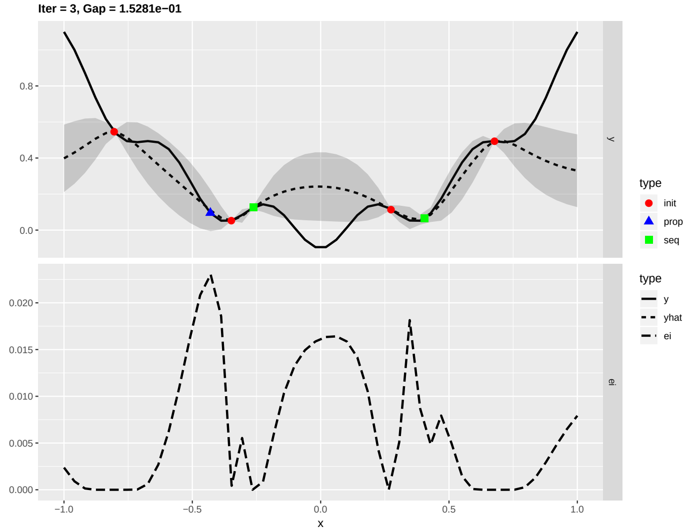
Lookman, Turab, Prasanna V. Balachandran, Dezhen Xue, and Ruihao Yuan. "Active Learning in Materials Science with Emphasis on Adaptive Sampling Using Uncertainties for Targeted Design." Npj Computational Materials 5, no. 1 (February 18, 2019): 21. https://doi.org/10.1038/s41524-019-0153-8.

Bayesian Optimization



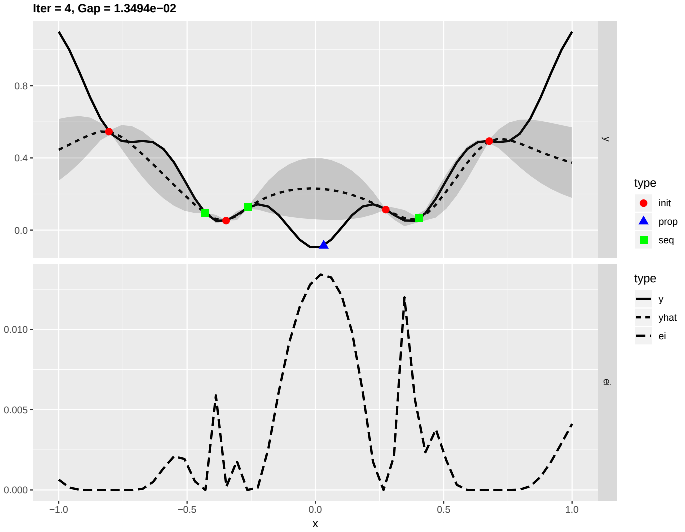
Bischl, Bernd, Jakob Richter, Jakob Bossek, Daniel Horn, Janek Thomas, and Michel Lang. "MlrMBO: A Modular Framework for Model-Based Optimization of Expensive Black-Box Functions," March 9, 2017. http://arxiv.org/abs/1703.03373.

Bayesian Optimization



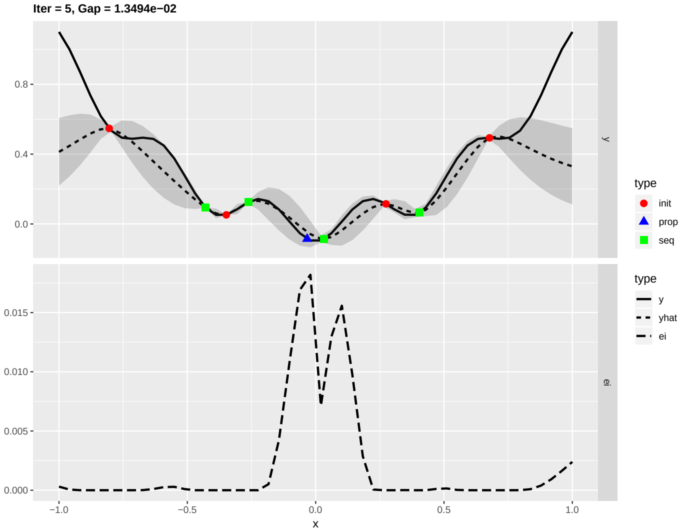
Bischl, Bernd, Jakob Richter, Jakob Bossek, Daniel Horn, Janek Thomas, and Michel Lang. "MlrMBO: A Modular Framework for Model-Based Optimization of Expensive Black-Box Functions," March 9, 2017. http://arxiv.org/abs/1703.03373.

Bayesian Optimization



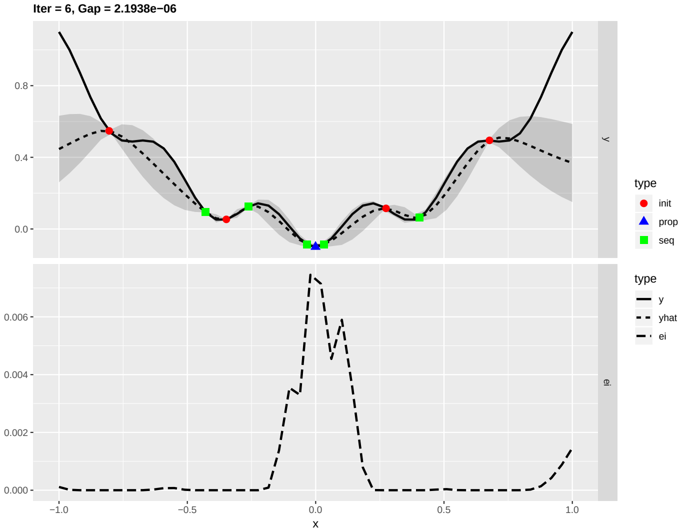
Bischl, Bernd, Jakob Richter, Jakob Bossek, Daniel Horn, Janek Thomas, and Michel Lang. "MlrMBO: A Modular Framework for Model-Based Optimization of Expensive Black-Box Functions," March 9, 2017. http://arxiv.org/abs/1703.03373.

Bayesian Optimization



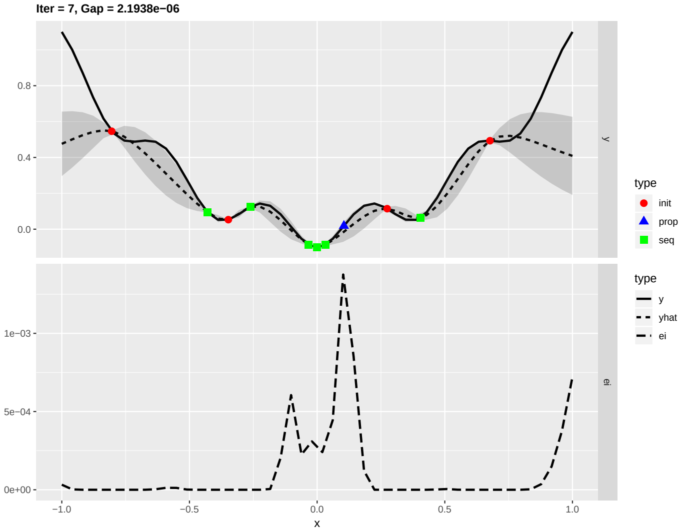
Bischl, Bernd, Jakob Richter, Jakob Bossek, Daniel Horn, Janek Thomas, and Michel Lang. "MlrMBO: A Modular Framework for Model-Based Optimization of Expensive Black-Box Functions," March 9, 2017. http://arxiv.org/abs/1703.03373.

Bayesian Optimization

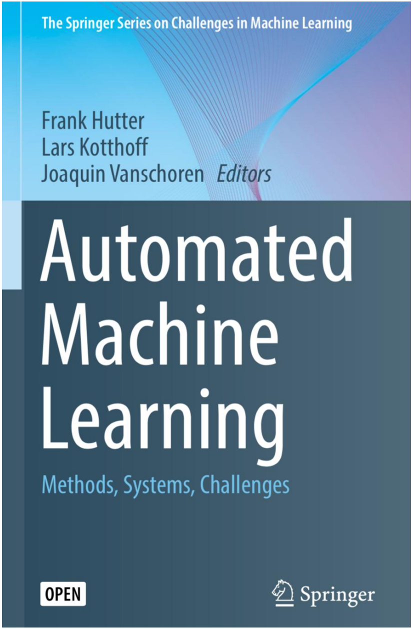


Bischl, Bernd, Jakob Richter, Jakob Bossek, Daniel Horn, Janek Thomas, and Michel Lang. "MlrMBO: A Modular Framework for Model-Based Optimization of Expensive Black-Box Functions," March 9, 2017. http://arxiv.org/abs/1703.03373.

Bayesian Optimization



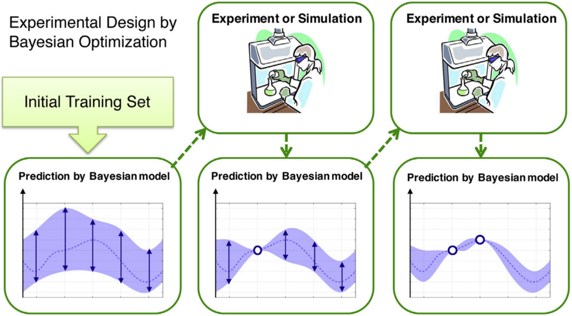
Bischl, Bernd, Jakob Richter, Jakob Bossek, Daniel Horn, Janek Thomas, and Michel Lang. "MlrMBO: A Modular Framework for Model-Based Optimization of Expensive Black-Box Functions," March 9, 2017. http://arxiv.org/abs/1703.03373.



https://www.automl.org/book/

COMBO

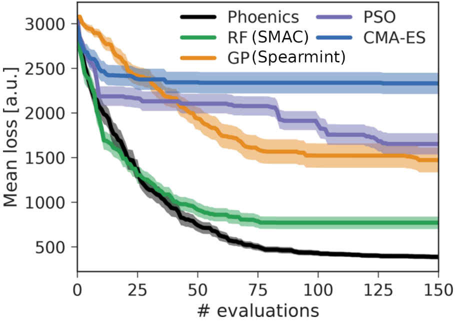
- ▷ Bayesian Optimization framework for iterative design space exploration
- ▷ optimize grain boundary in copper, evaluations are first-principles simulations
- ▷ specific focus on scalability – thousands of data points



Ueno, Tsuyoshi, Trevor David Rhone, Zhufeng Hou, Teruyasu Mizoguchi, and Koji Tsuda. "COMBO: An Efficient Bayesian Optimization Library for Materials Science." Materials Discovery 4 (2016): 18–21. https://doi.org/10.1016/j.md.2016.04.001.

Phoenics

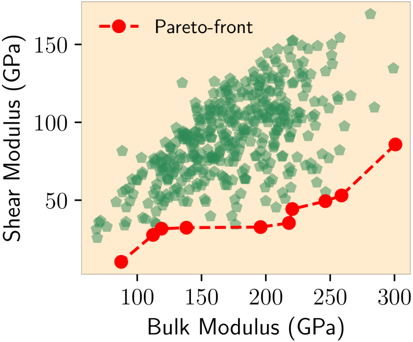
- ▷ Bayesian Optimization for optimizing benchmarks of chemical reactions, evaluations are first-principles simulations
- ▷ Bayesian Neural Networks to estimate parameters of Bayesian kernel density distribution
- ▷ parallel evaluations through batching of different acquisition functions



Häse, Florian, Loïc M. Roch, Christoph Kreisbeck, and Alán Aspuru-Guzik. "Phoenics: A Bayesian Optimizer for Chemistry." ACS Central Science 4, no. 9 (September 26, 2018): 1134–45. https://doi.org/10.1021/acscentsci.8b00307.

Matpredict

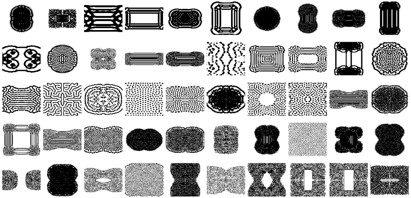
- Bayesian Optimization for shear and bulk modulus in MAX-phase materials, which consist of layers of different elements and can behave like metals or ceramics
- Bayesian Model Averaging for feature selection
- evaluations are first-principles simulations

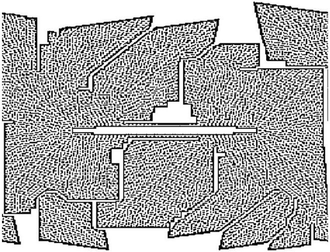


Talapatra, Anjana, Shahin Boluki, Thien Duong, Xiaoning Qian, Edward Dougherty, and Raymundo Arróyave. "Autonomous Efficient Experiment Design for Materials Discovery with Bayesian Model Averaging." Phys. Rev. Materials 2, no. 11 (November 2018): 113803. https://doi.org/10.1103/PhysRevMaterials.2.113803.

Materials Generation

- Bayesian Optimization for surface area (energy absorption) of material
- evaluations are first-principles simulations
- patterns can be 3D-printed and evaluated experimentally





Common Themes

- fundamentally, efficient design space exploration problems
- avoid expensive evaluations
- Bayesian Optimization and variants emerging as state of the art
- application-agnostic black-box methods

Common Themes

AI, Machine Learning

- approaches
  - Spearmint
  - SMAC
  - mlrMBO
  - ...
- repositories
  - OpenML
  - UCI ML repo
  - ...

Materials Science

- approaches
  - COMBO
  - Phoenics
  - Matpredict
  - ...
- repositories
  - JARVIS `https://jarvis.nist.gov/`
  - Materials Project `https://www.materialsproject.org/`
  - ...

Challenges and Opportunities

- sparsity of data – common repositories
- parallelization, scalability
- multi-scale measurements
- combination with experiments and simulations
- avoiding duplication

Beyond Design Space Exploration

- high-throughput experiments – robotics, computer vision...
- natural language processing for mining papers for data
- understanding of models/experiments/data

Do Try This at Home

Simulator optimizers available at
`https://www.cs.uwyo.edu/~larsko/aimat-tut/`

- build surrogate model based on (relatively) large amount of data
- Bayesian Optimization based on this surrogate model
- playground to try your own approaches

Summary

- lots of applications for AI in Materials Science, especially Bayesian Optimization and surrogate modeling
- tools used in Materials Science are not as comprehensive and mature as in AI
- applications to real-world problems pose interesting challenges for AI