Bayesian Optimization:
From Foundations to Advanced Topics

twitter

@deshwal_aryan

@syrineblk

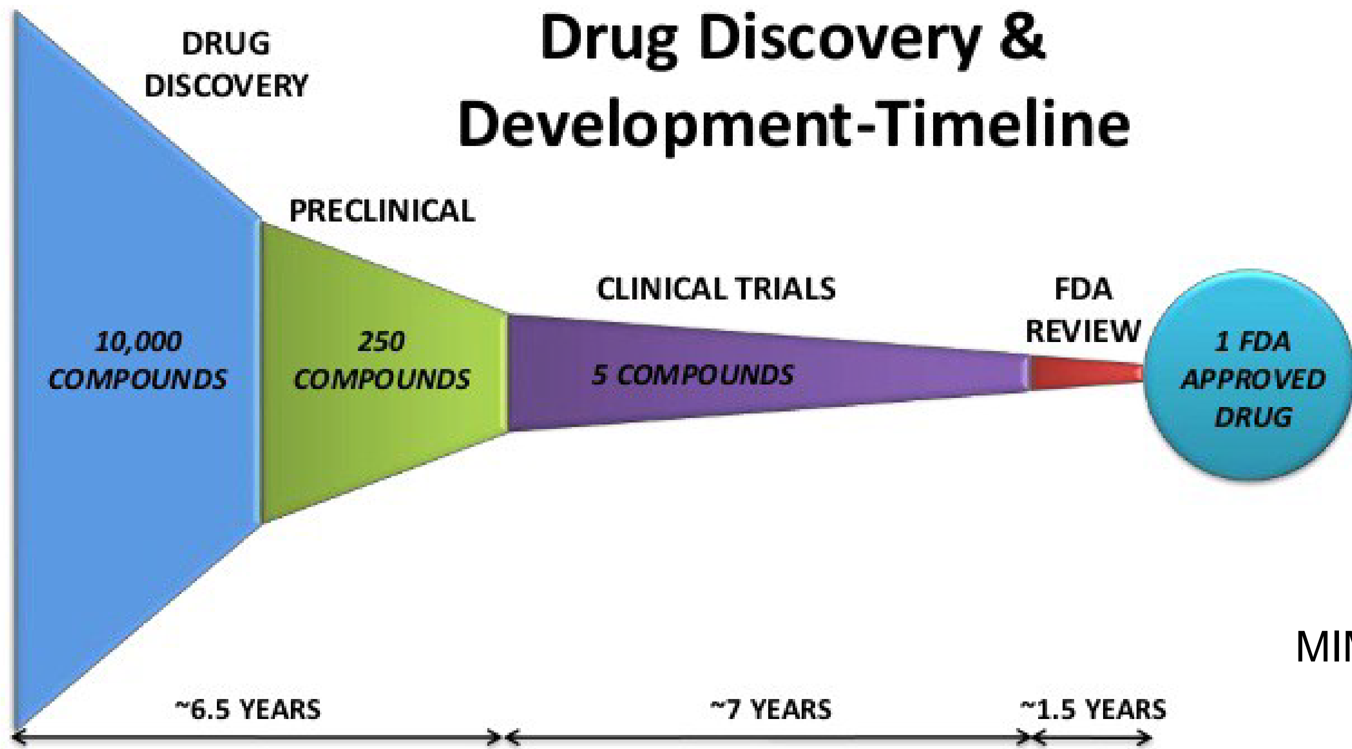
@janadoppa

NSF

Half-day Tutorial

@ AAAI-2022 Conference

Drug/Vaccine Design

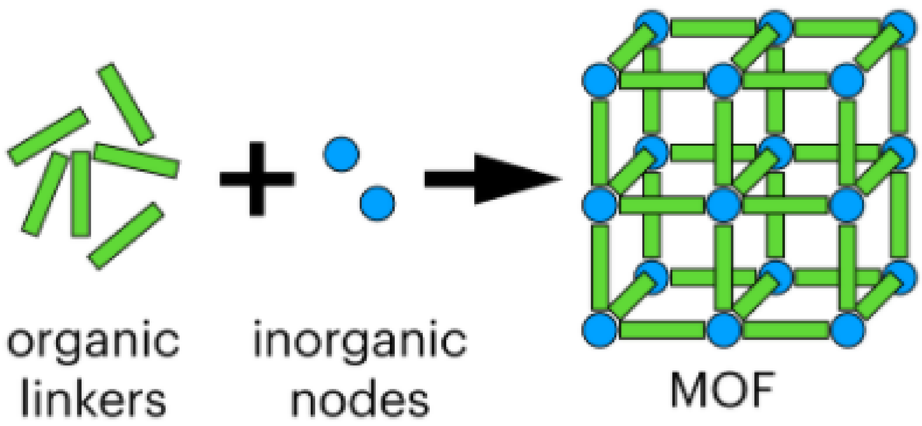


Credit:

MIMA healthcare

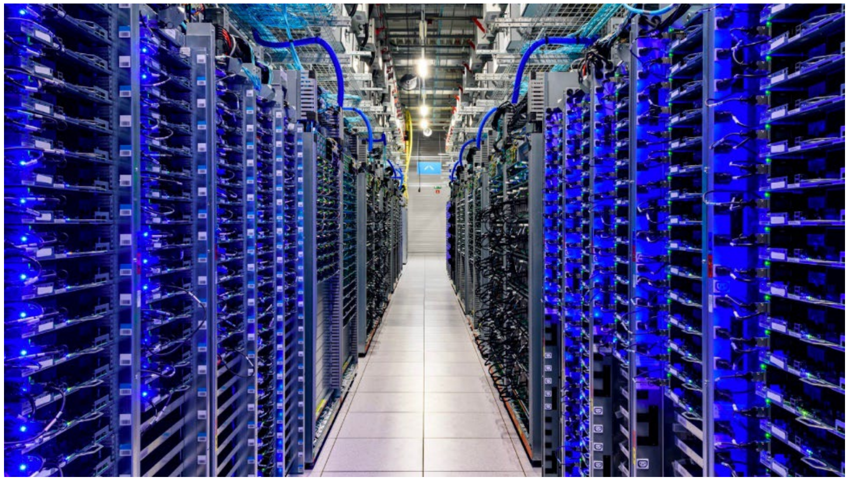
- Accelerate the discovery of promising designs

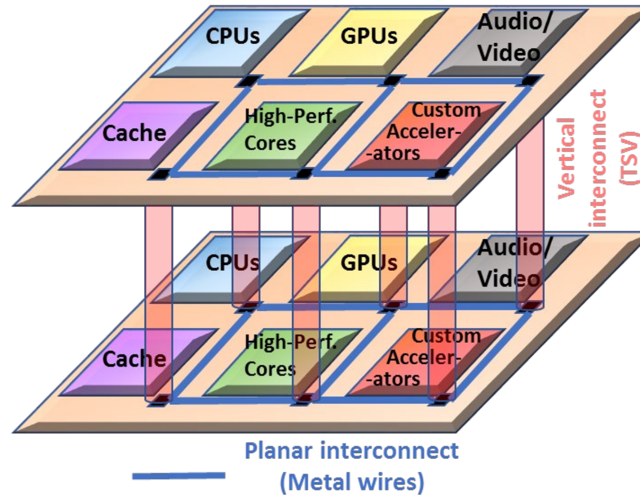
Nanoporous Materials Design



- **Sustainability applications**
  - Storing gases (e.g., hydrogen powered cars)
  - Separating gases (e.g., carbon dioxide from flue gas of coalfired power plants)
  - Detecting gases (e.g., detecting pollutants in outdoor air)

Sustainable Hardware Design for Data Centers





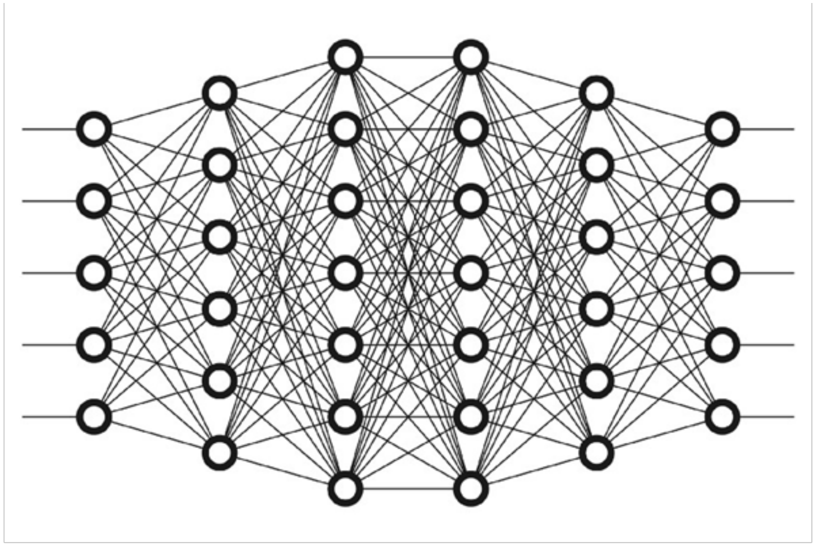
America's Data Centers Are Wasting Huge Amounts of Energy

By 2020, data centers are projected to consume roughly 140 billion kilowatt-hours annually, costing American businesses $13 billion annually in electricity bills and emitting nearly 150 million metric tons of carbon pollution

High-performance and Energy-efficient manycore chips

Report from Natural Resources Defense Council:.
https://www.nrdc.org/sites/default/files/data-center-efficiency-assessment-IB.pdf

Auto ML and Hyperparameter Tuning

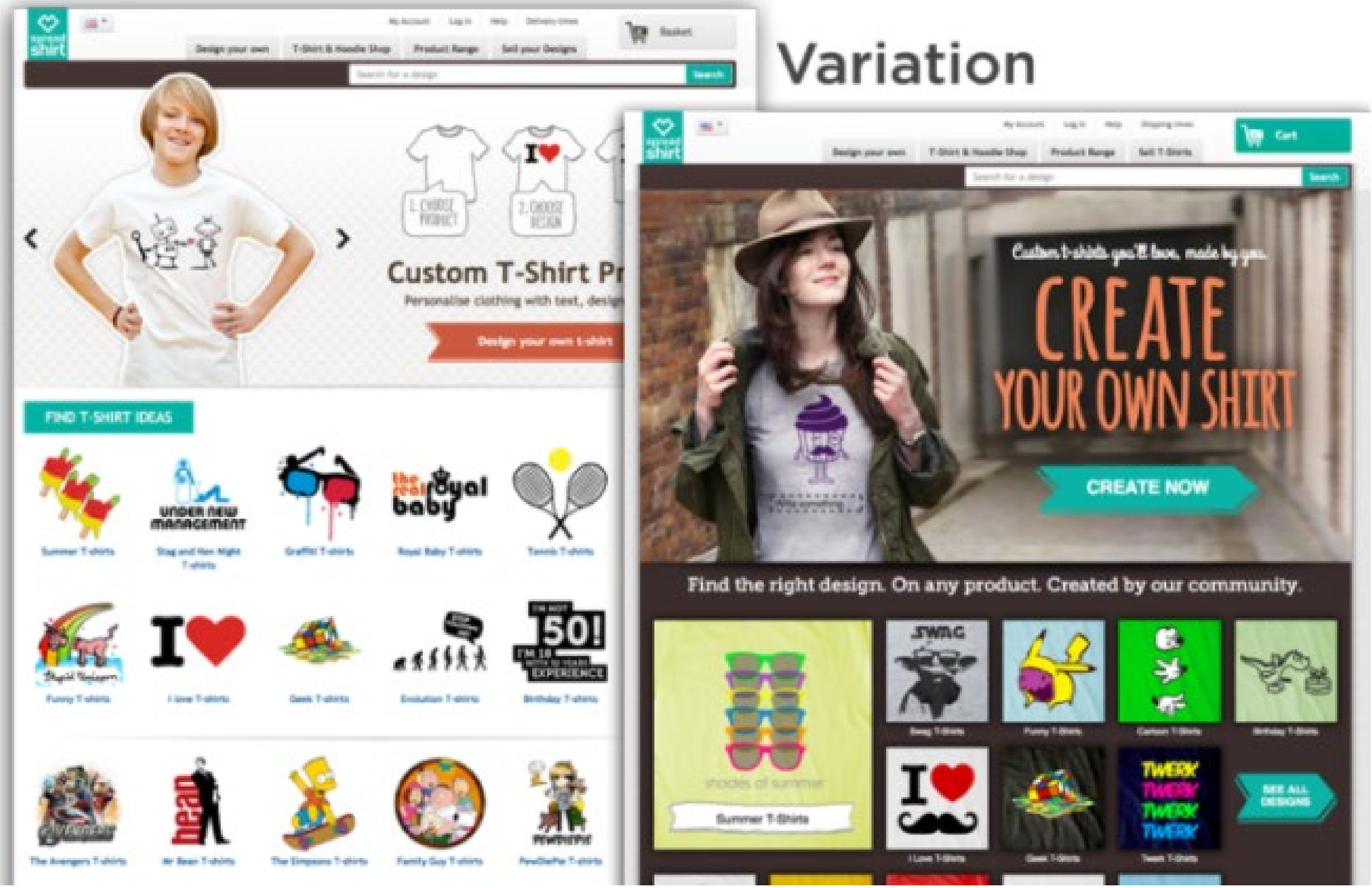


- Accuracy of models critically depends on hyper-parameters
  - Optimization algorithm, learning rates, momentum, batch normalization, batch sizes, dropout rates, weight decay, data augmentation, …

A/B Testing to Configure Websites

Original

Variation



Making Delicious Cookies

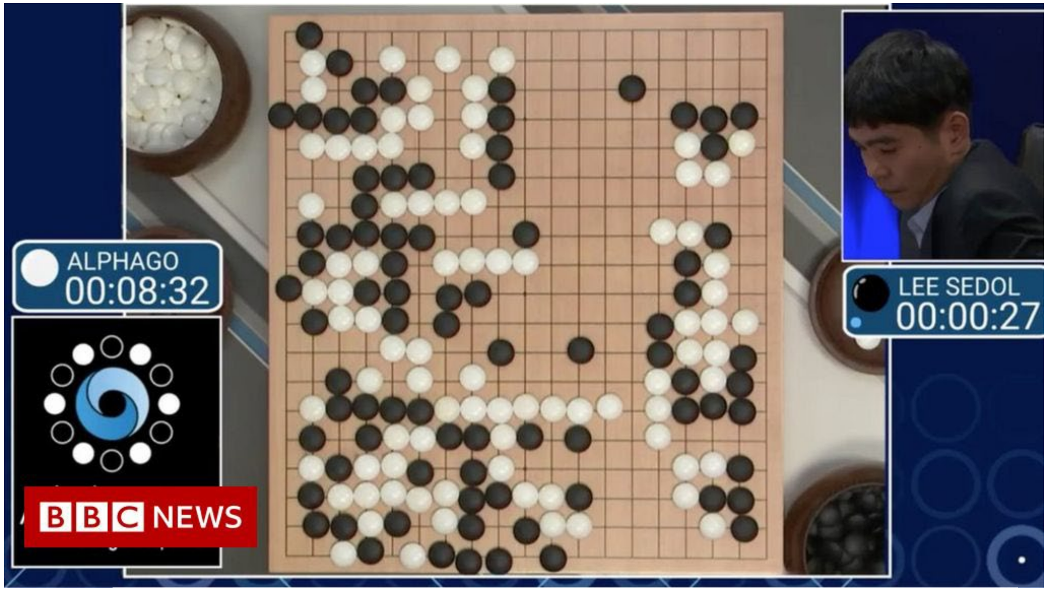
Bayesian Optimization for a Better Dessert

Greg Kochanski, Daniel Golovin, John Karro, Benjamin Solnik, Subhodeep Moitra, and D. Sculley
{gpk, dgg, karro, bsolnik, smoitra, dsculley}@google.com; Google Brain Team

Abstract

We present a case study on applying Bayesian Optimization to a complex real-world system; our challenge was to optimize chocolate chip cookies. The process was a mixed-initiative system where both human chefs, human raters, and a machine optimizer participated in 144 experiments. This process resulted in highly rated cookies that deviated from expectations in some surprising ways – much less sugar in California, and cayenne in Pittsburgh. Our experience highlights the importance of incorporating domain expertise and the value of transfer learning approaches.

Making AlphaGo Better



Bayesian Optimization in AlphaGo

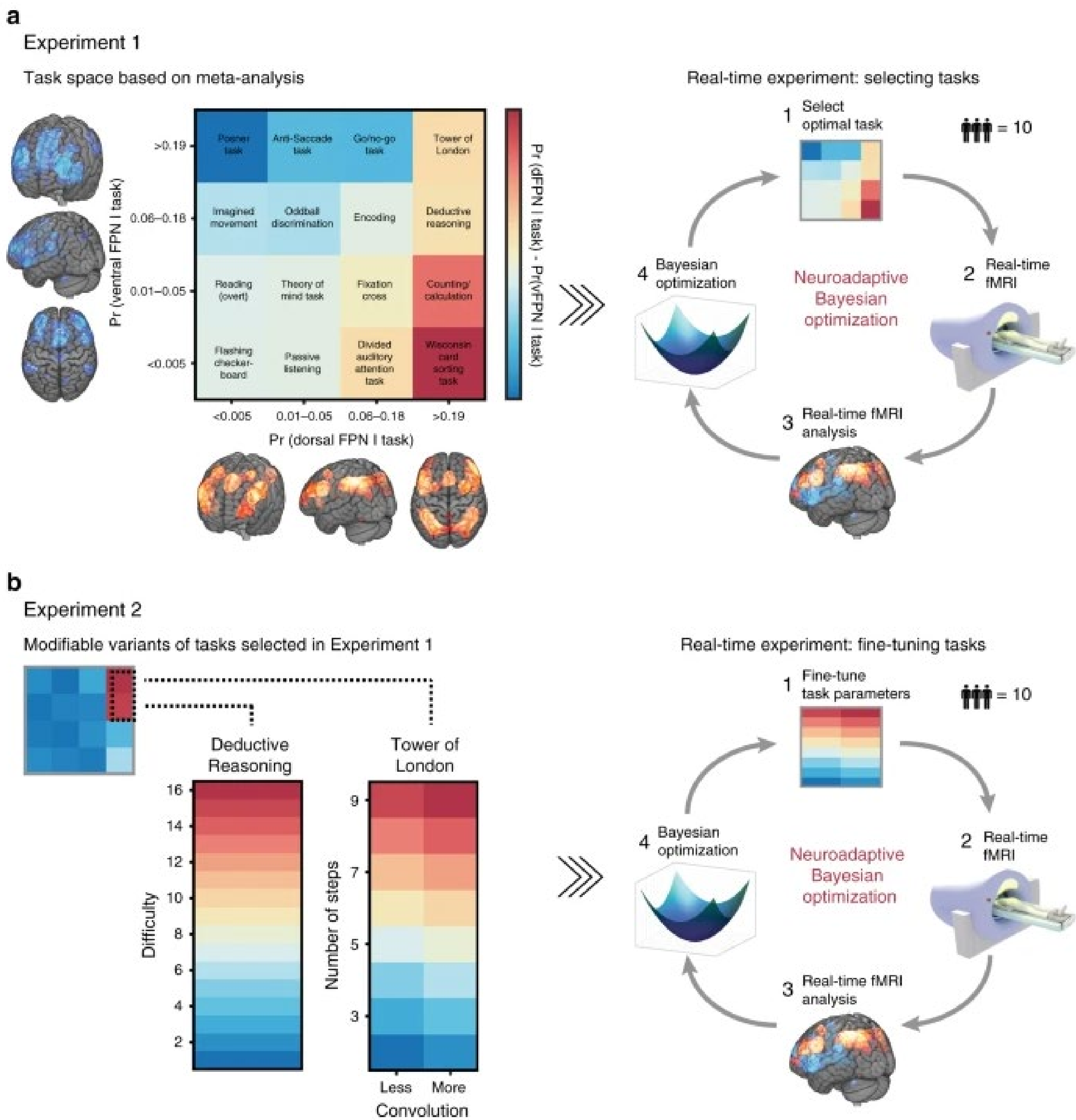
Yutian Chen, Aja Huang, Ziyu Wang, Ioannis Antonoglou, Julian Schrittwieser, David Silver & Nando de Freitas

DeepMind, London, UK
`yutianc@google.com`

Abstract

During the development of AlphaGo, its many hyper-parameters were tuned with Bayesian optimization multiple times. This automatic tuning process resulted in substantial improvements in playing strength. For example, prior to the match with Lee Sedol, we tuned the latest AlphaGo agent and this improved its win-rate from 50% to 66.5% in self-play games. This tuned version was deployed in the final match. Of course, since we tuned AlphaGo many times during its development cycle, the compounded contribution was even higher than this percentage. It is our hope that this brief case study will be of interest to Go fans, and also provide Bayesian optimization practitioners with some insights and inspiration.

Neuroscience and Brain Analytics



Credit: https://www.nature.com/articles/s41467-018-03657-3

Common Attributes of the Search Problem

- Search Space: Many candidate choices (inputs)
- Objective function: Need to perform an expensive experiment to evaluate the objective value of any input
- Optimization problem: find the candidate input with highest objective function value

Common Attributes of the Search Problem

- Search Space: Many candidate choices (inputs)
- Objective function: Need to perform an expensive experiment to evaluate the objective value of any input
- Optimization problem: find the candidate input with highest objective function value

Cannot afford exhaustive search

Common Attributes of the Search Problem

- Search Space: Many candidate choices (inputs)
- Objective function: Need to perform an expensive experiment to evaluate the objective value of any input
- Optimization problem: find the candidate input with highest objective function value

Trial and Error?

Common Attributes of the Search Problem

- Search Space: Many candidate choices (inputs)

- Objective function: Need to perform an expensive experiment to evaluate the objective value of any input

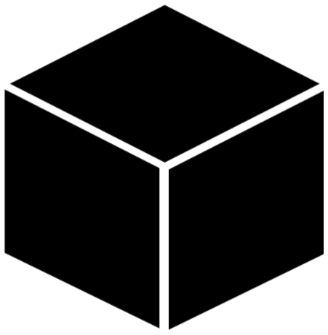
- Optimization problem: find the candidate input with highest objective function value

> Can we do better than trial-and-error?

Accelerate Search via Bayesian Optimization

- Efficiently optimize expensive black-box functions

$$x^* = arg \max_{x \in X} f(x)$$

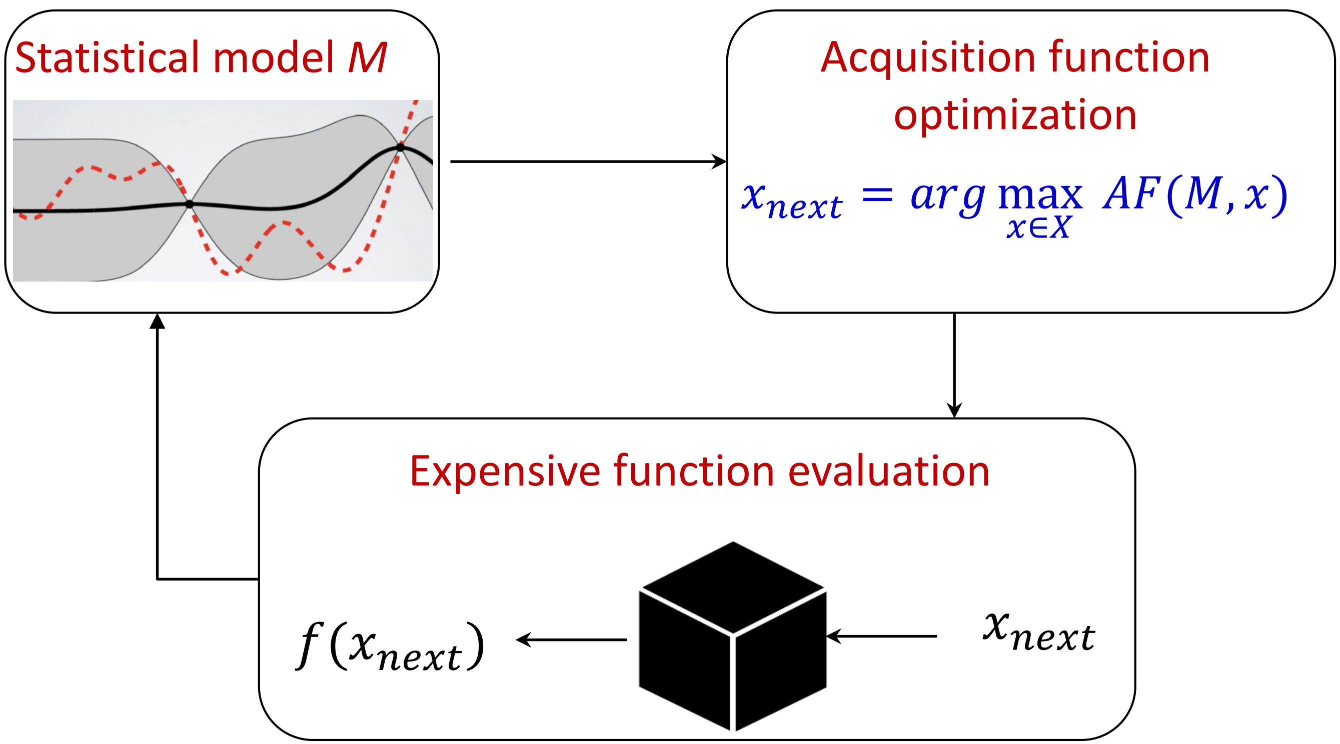
x →  → $f(x)$

input | Function evaluation

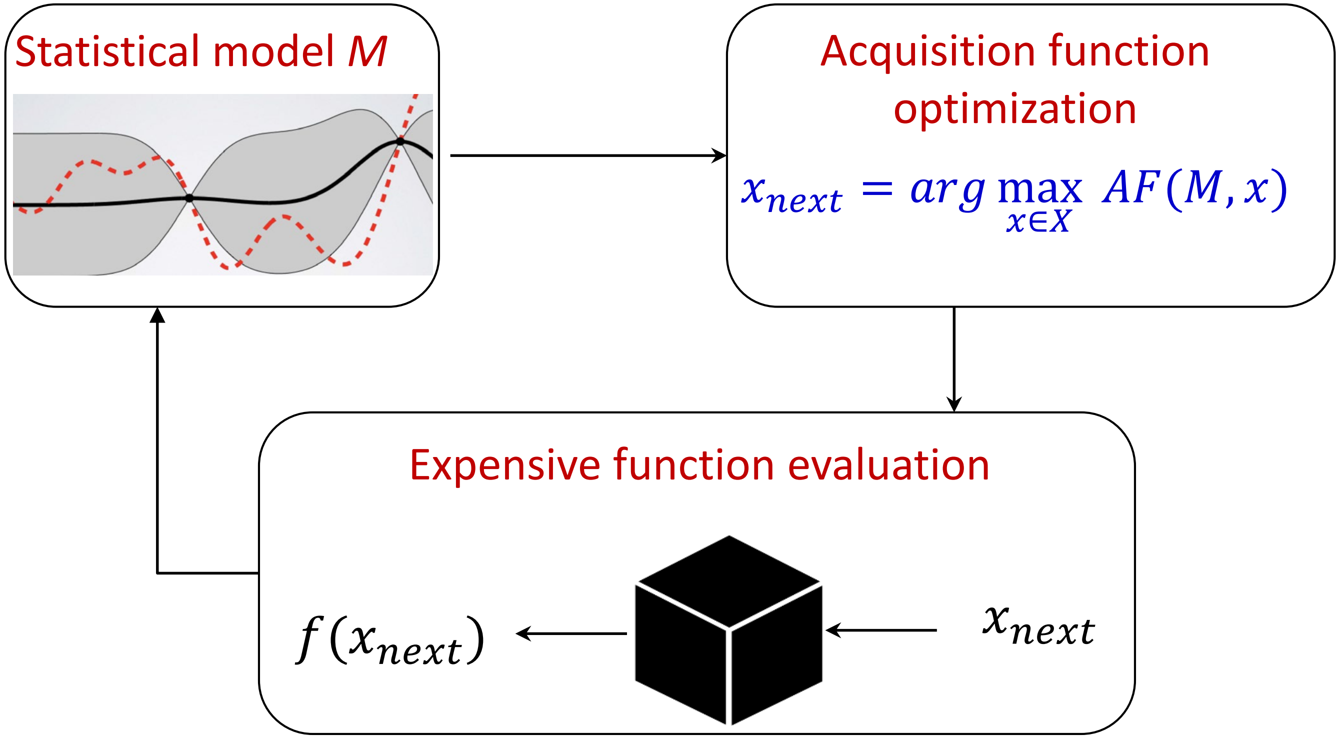
- Black-box queries (aka experiments) are expensive

Bayesian Optimization: Key Idea

- Build a surrogate statistical model and use it to intelligently search the space
  - Replace expensive queries with cheaper queries
  - Use uncertainty of the model to select expensive queries



Bayesian Optimization: Three Key Elements



- Statistical model (e.g., Gaussian process)
- Acquisition function (e.g., Expected improvement)
- Acquisition function optimizer (e.g., local search)

BO Dimensions: Input Space

- **Continuous space**
  - All variables of input x are continuous
- **Discrete / Combinatorial space**
  - Sequences, trees, graphs, sets, permutations etc.
- **Hybrid space**
  - x = mixture of x_d (discrete) and x_c (continuous) variables

BO Dimensions: Input Space

- **Continuous space**
  - All variables of inpu

Most of the focus of existing BO work

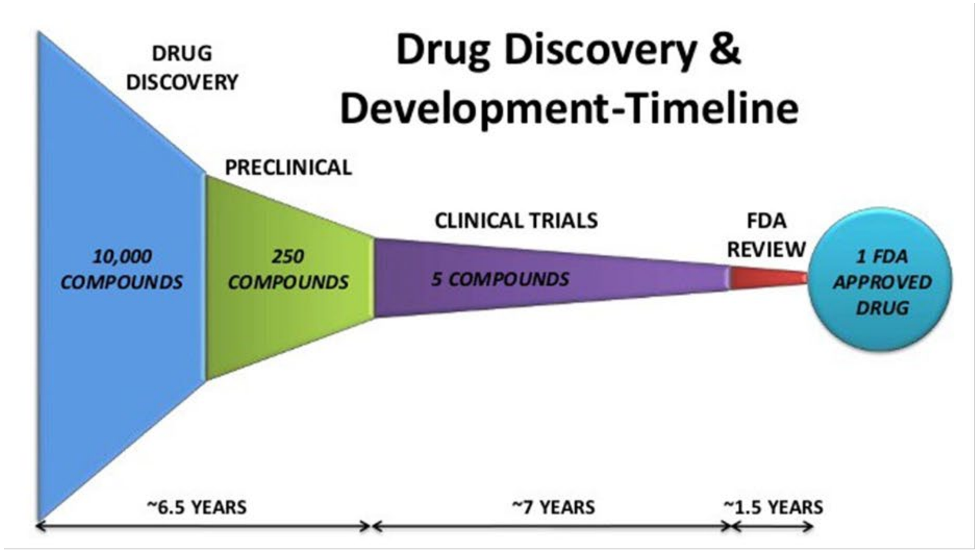
- **Discrete / Combinatorial space**
  - Sequences, trees, graphs, sets, permutations etc.

- **Hybrid space**
  - $x =$ mixture of x_d (discrete) and x_c (continuous) variables

BO Dimensions: No. of Objectives

- **Single objective**
  - For example, finding hyperparameters to optimize accuracy

- **Multiple objectives**



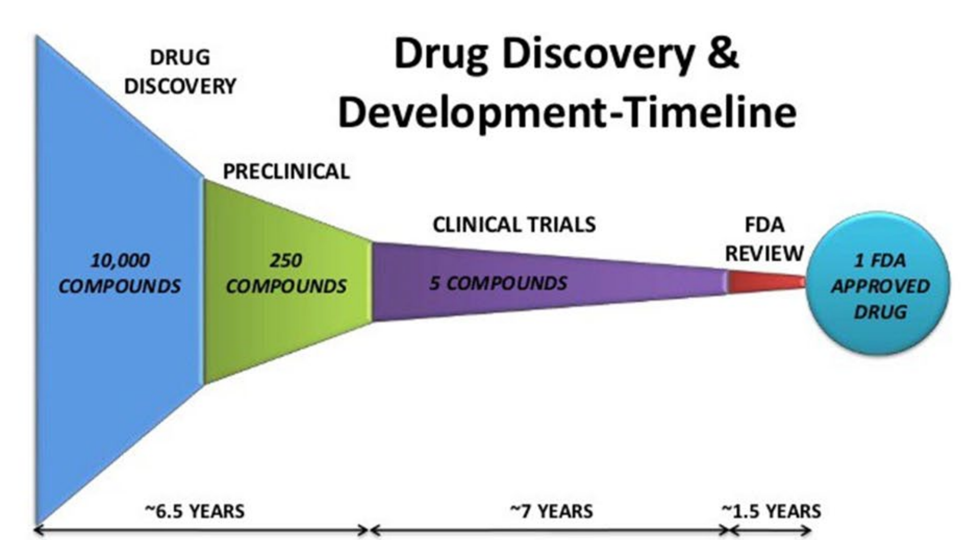
Credit: MIMA healthcare

BO Dimensions: No. of Objectives

- **Single objective**
  - For example, fin ptimize accuracy

Most of the focus of existing BO work

- **Multiple objectives**

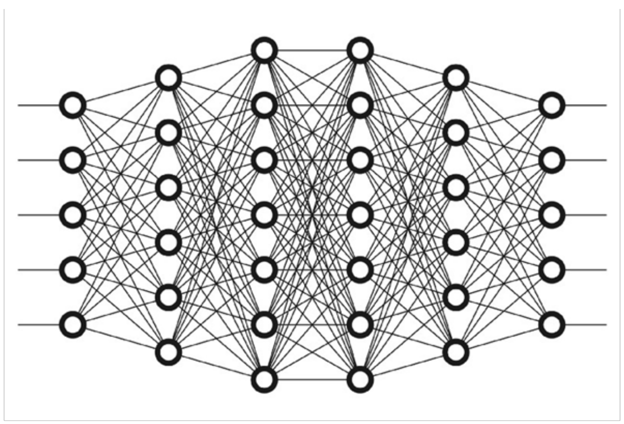


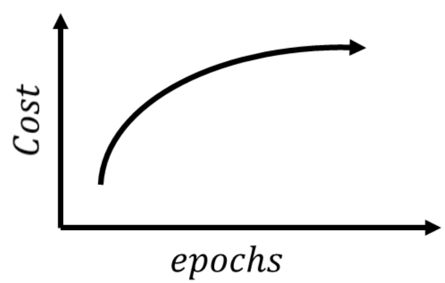
Credit: MIMA healthcare

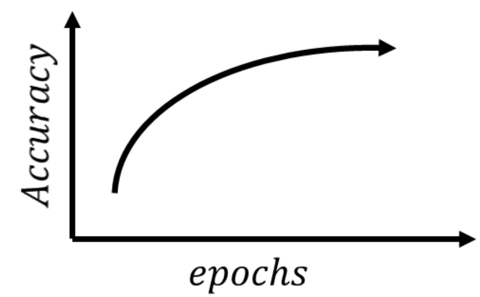
BO Dimensions: No. of Fidelities

- **Single-fidelity setting**
  - Most expensive and accurate function evaluation

- **Multi-fidelity setting**
  - Function evaluations with varying trade-offs in cost and accuracy





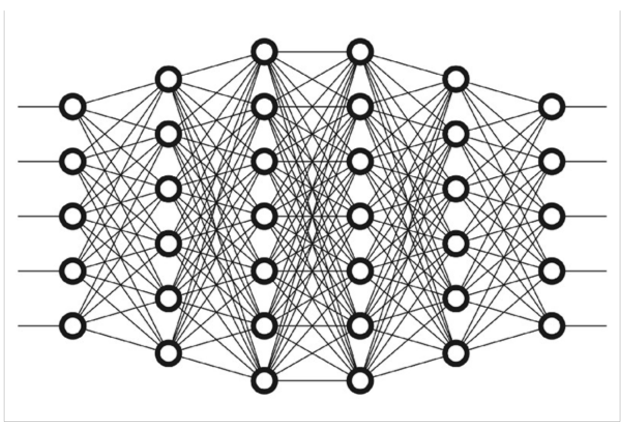


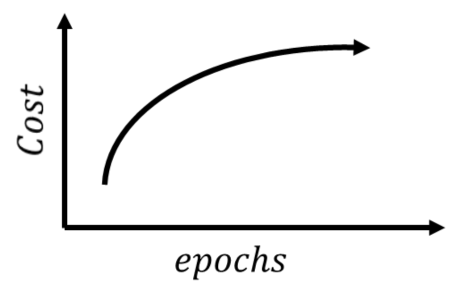
BO Dimensions: No. of Fidelities

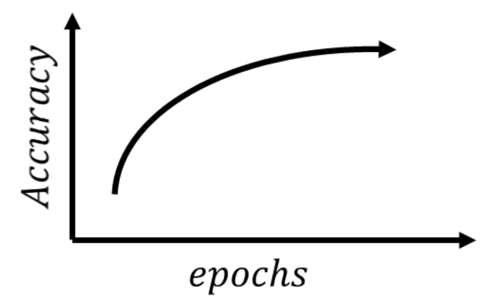
- **Single-fidelity setting**
  - Most expensive and ac

Most of the focus of existing BO work

- **Multi-fidelity setting**
  - Function evaluations with varying trade-offs in cost and accuracy

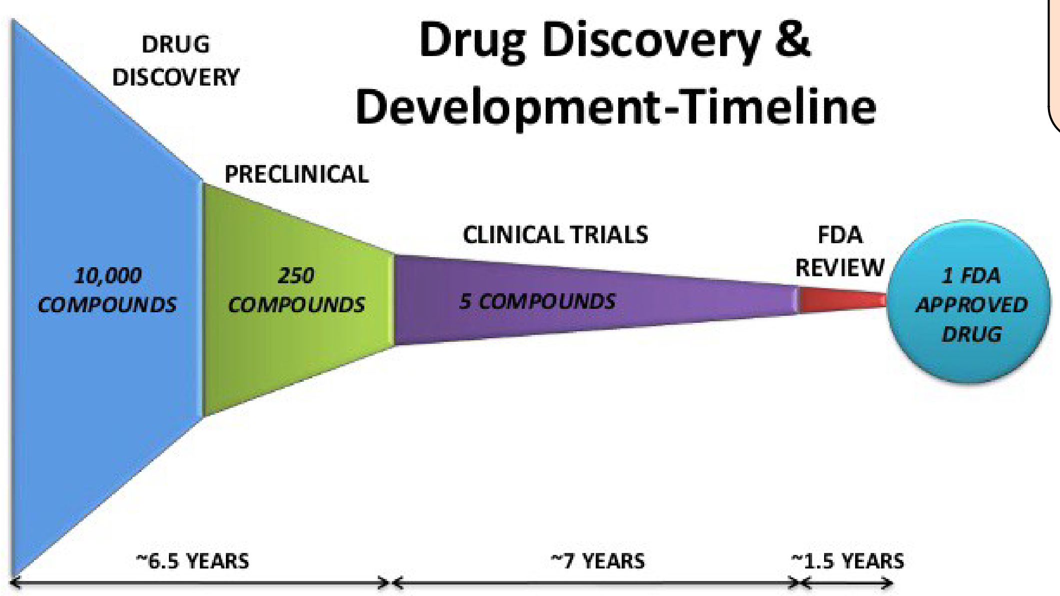






BO Dimensions: Constraints

- **Unconstrained setting**
  - all inputs are valid
- **Constrained setting**



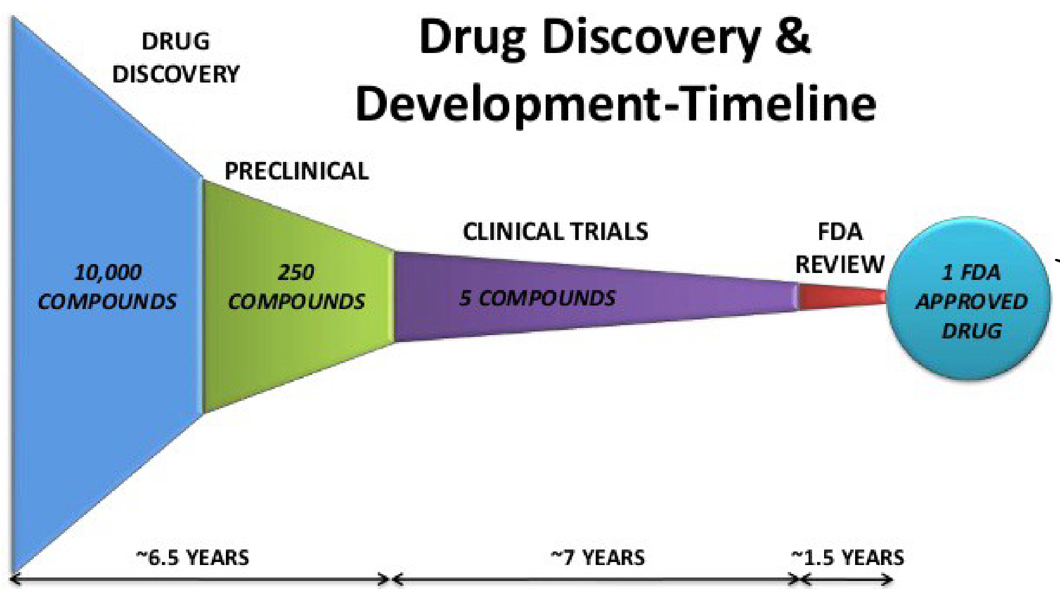
Drugs/Vaccines that are safe

BO Dimensions: Constraints

- **Unconstrained setting**
  - all inputs are valid

Most of the focus of existing BO work

- **Constrained setting**



Drugs/Vaccines that are safe

Outline of the Tutorial

- Background on GPs and Single-Objective BO
- Bayesian Optimization over Combinatorial Spaces
- Bayesian Optimization over Hybrid Spaces

Break

- Multi-Fidelity Bayesian Optimization
- Constrained Bayesian Optimization
- Multi-Objective Bayesian Optimization
- Summary and Outstanding Challenges in BO