How to Optimize Black-box Functions using Machine Learning: An Introduction to Bayesian Optimization

IER Biweekly Science Lecture,
24 March 2023

Hometown

Photos references: Mehr News Agency (https://en.mehrnews.com/)

Hometown

Horse racing

Turkmen carpet

B.Sc in Petroleum Engineering ,

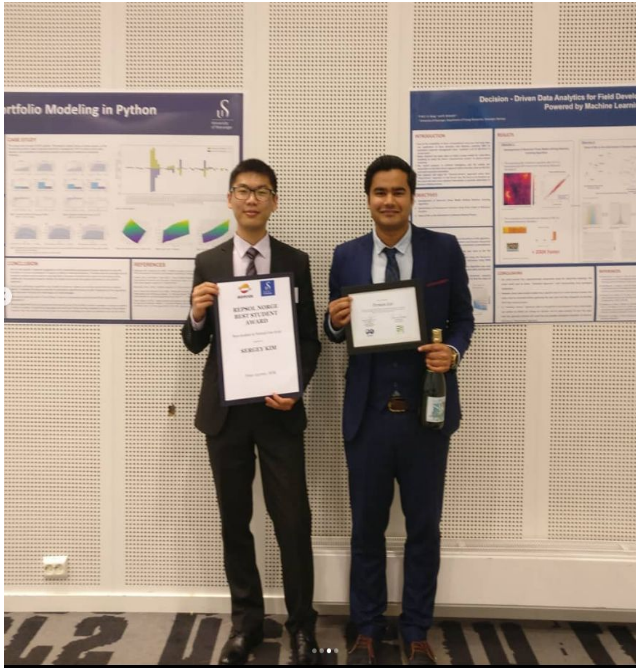
Petroleum University of Technology (2011-2016) - Iran 🇮🇷

B.Sc in Petroleum Engineering ,

Petroleum University of Technology (2011-2016) - Iran 🇮🇷

M.Sc in Reservoir Engineering,

University of Stavanger (2017-2019) - Norway 🇳🇴



M.Sc in Reservoir Engineering,

University of Stavanger (2017-2019) - Norway 🇳🇴

Decision-Driven Data Analytics for Well Placement Optimization in Field Development Scenario - Powered by Machine Learning

M.Sc Thesis

Peyman Kor

June 2019

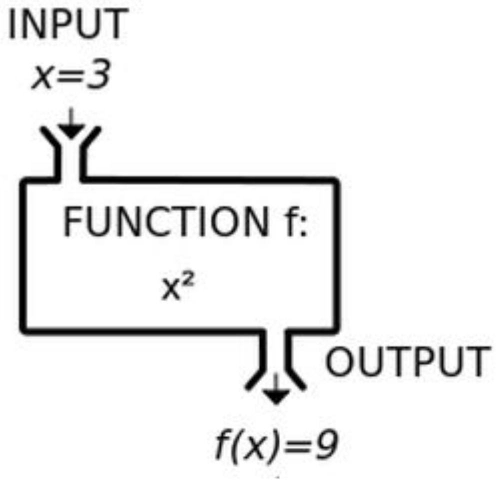
M.Sc in Applied Mathematics and Computation,

Technical University of Denmark (2019 - 2020) - Denmark 🇩🇰

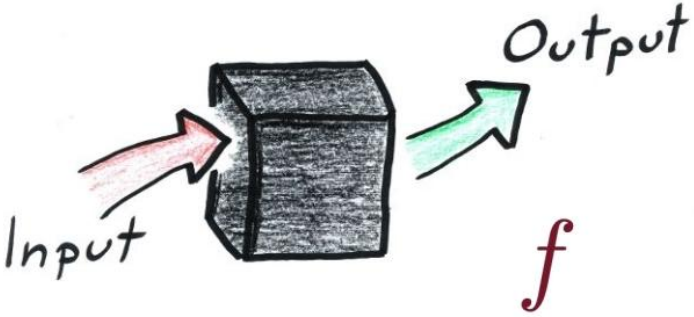
Black-Box Function Optimization?

What is Black-box Function?

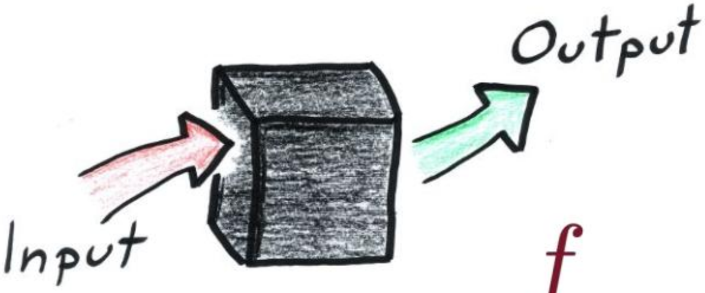
White-box function



Black-box function



Black-box Function Optimization

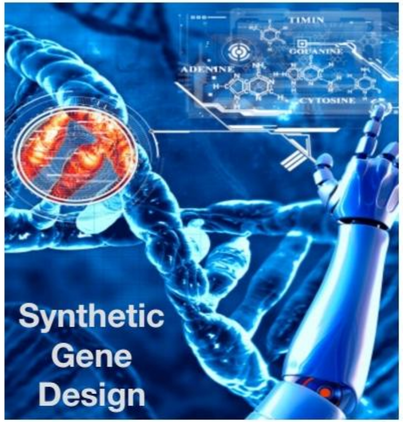


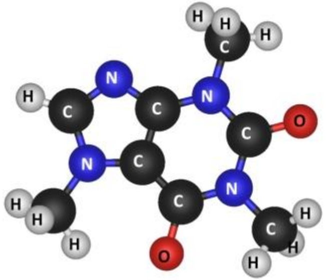
f

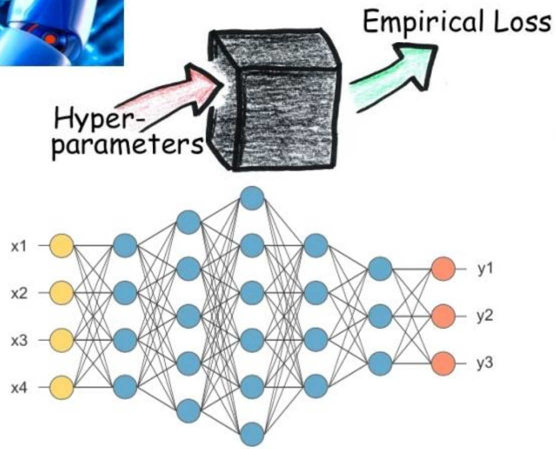
Goal:

$$x_* = \underset{\mathcal{X} \subset \mathbb{R}^d}{\operatorname{argmax}} \; f(x)$$

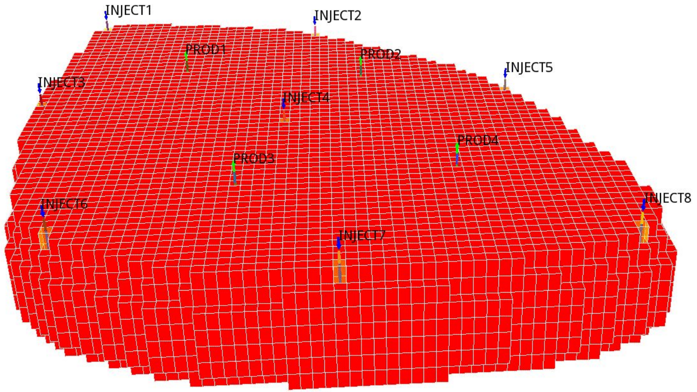
Black-box Function Optimization







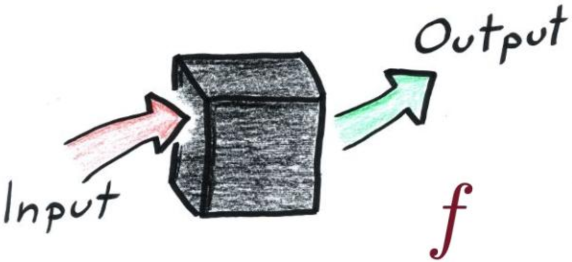
Output of Reservoir Simulator



Why Black-box Function Optimization is Difficult?

Three main reasons: A, B, C

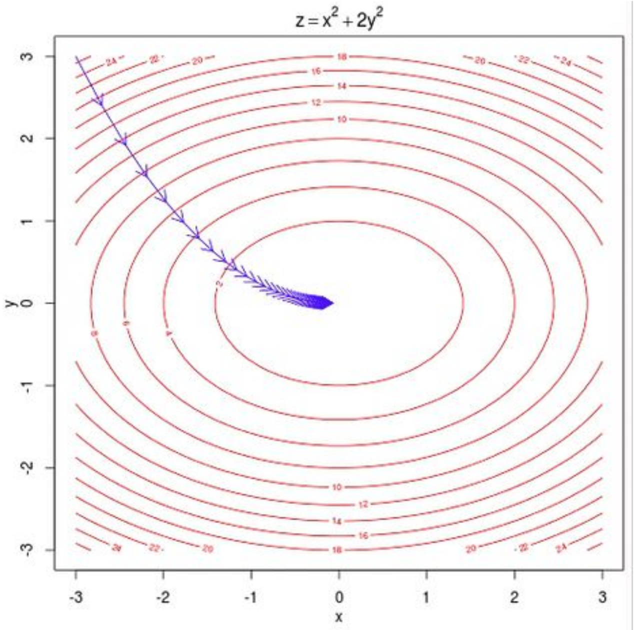
Why Black-box Function Optimization is Difficult? (A)



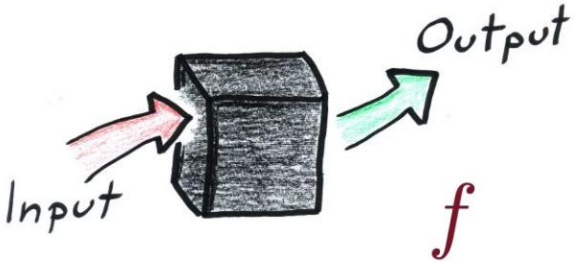
Goal:

$$x_* = \underset{\mathcal{X} \subset \mathbb{R}^d}{\operatorname{argmax}} f(x)$$

A) No gradient information!



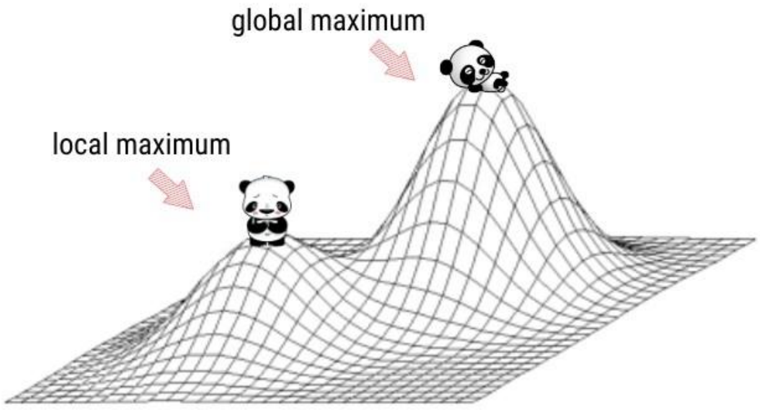
Why Black-box Function Optimization is Difficult? (B)



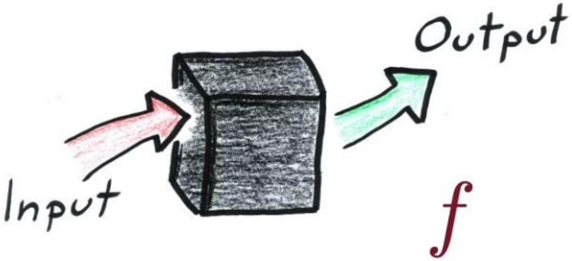
Goal:

$$x_* = \underset{\mathcal{X} \subset \mathbb{R}^d}{\operatorname{argmax}} \ f(x)$$

B) f(x) is multi-peak!



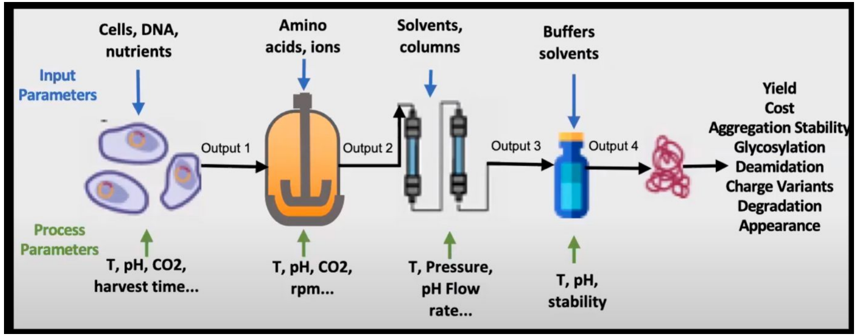
Why Black-box Function Optimization is Difficult? (C)



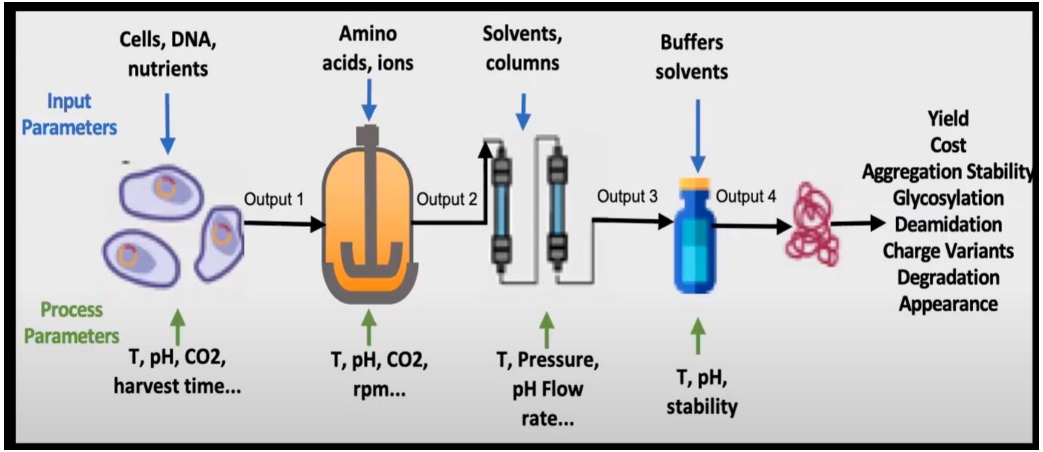
Goal:

$$x_* = \underset{\mathcal{X} \subset \mathbb{R}^d}{\operatorname{argmax}} f(x)$$

C) forward evaluation of is time-consuming!



Why Black-box Function Optimization is Difficult? (C)



- 1)Temperature, 2)pH, 3)Co2 Concentration, 4)Harvest Time
- According to them, each of the process parameters can take 5 different value.
- Then total possible combinations are 5^4 = 625

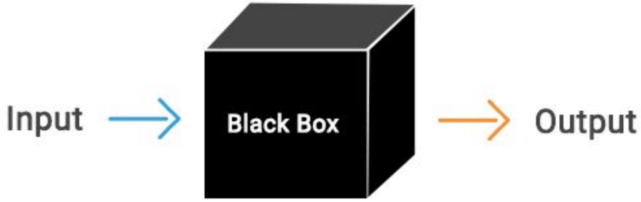
- Each forward evaluation of the experiments takes around 1 month.

$$5^4 \times \text{month} = 625 \times \text{month} \sim 52 \text{ years!}$$

- We need an optimization method that can handle all three difficulties!
- Bayesian Optimization
- Is a well-suited method for the global optimization of black -box functions.

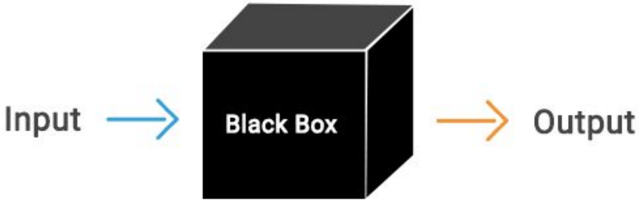
A 1-D Example of How Bayesian Optimization Works.

Maximize the Black-box Function:



Example of Bayesian Optimization:

Step I) Initialization:

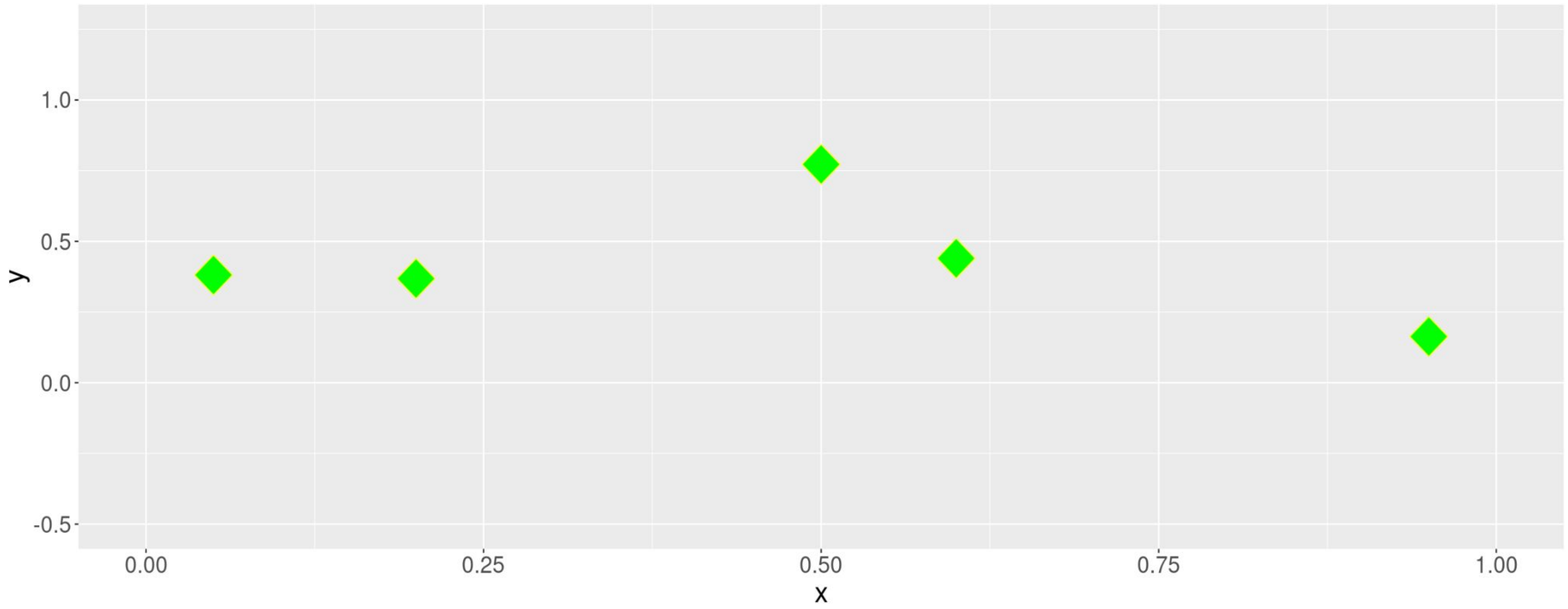


$$X = \begin{bmatrix} 0.05 \\ 0.2 \\ 0.5 \\ 0.6 \\ 0.95 \end{bmatrix} , Y = \begin{bmatrix} 0.38 \\ 0.37 \\ 0.77 \\ 0.44 \\ 0.16 \end{bmatrix}$$

Example of Bayesian Optimization:

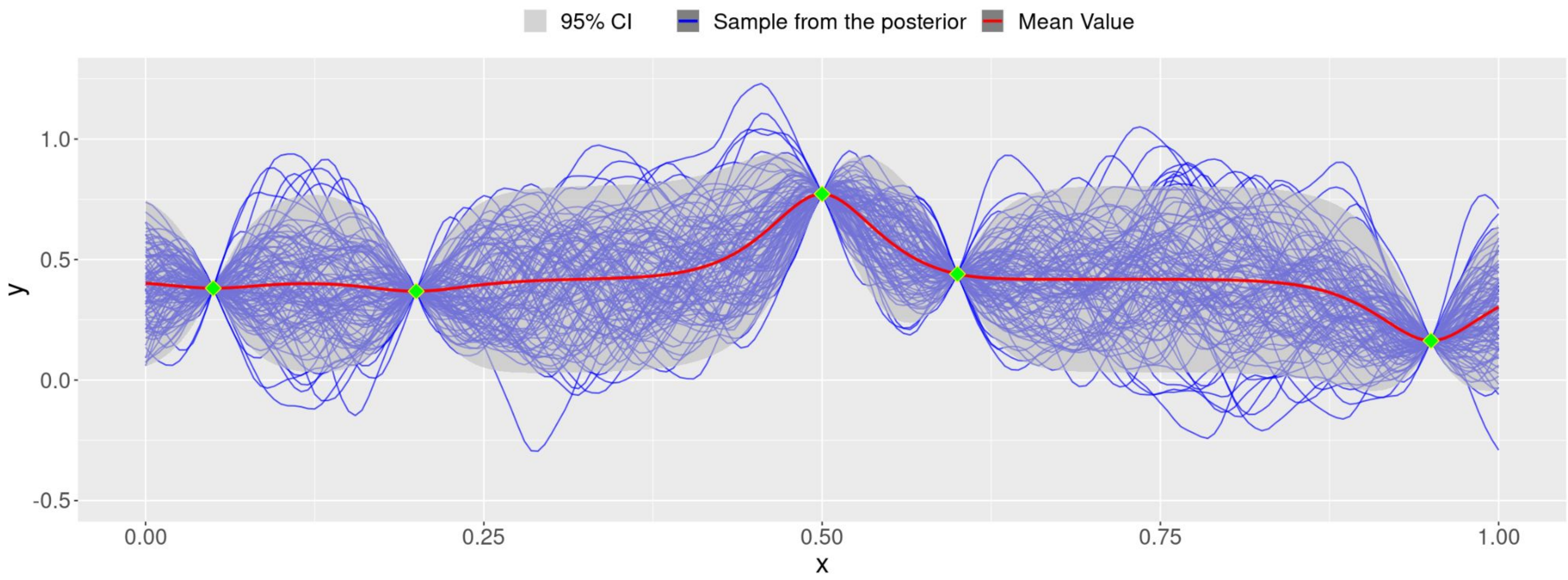
Step I) Plot of Initialization:

$$D: \quad X = \begin{bmatrix} 0.05 \\ 0.2 \\ 0.5 \\ 0.6 \\ 0.95 \end{bmatrix}, Y = \begin{bmatrix} 0.38 \\ 0.37 \\ 0.77 \\ 0.44 \\ 0.16 \end{bmatrix}$$



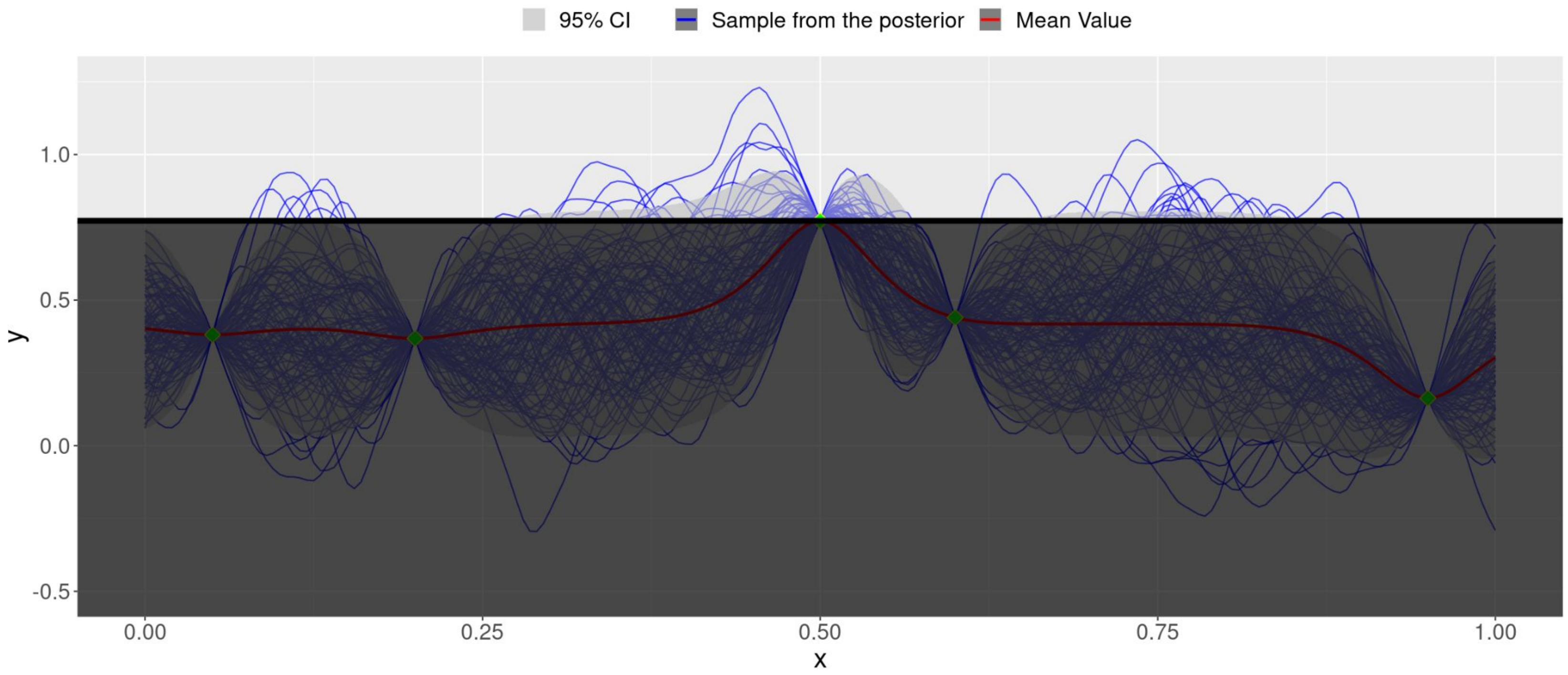
Example of Bayesian Optimization:

Step II) Build a probabilistic machine learning model (Gaussian Process Regression)



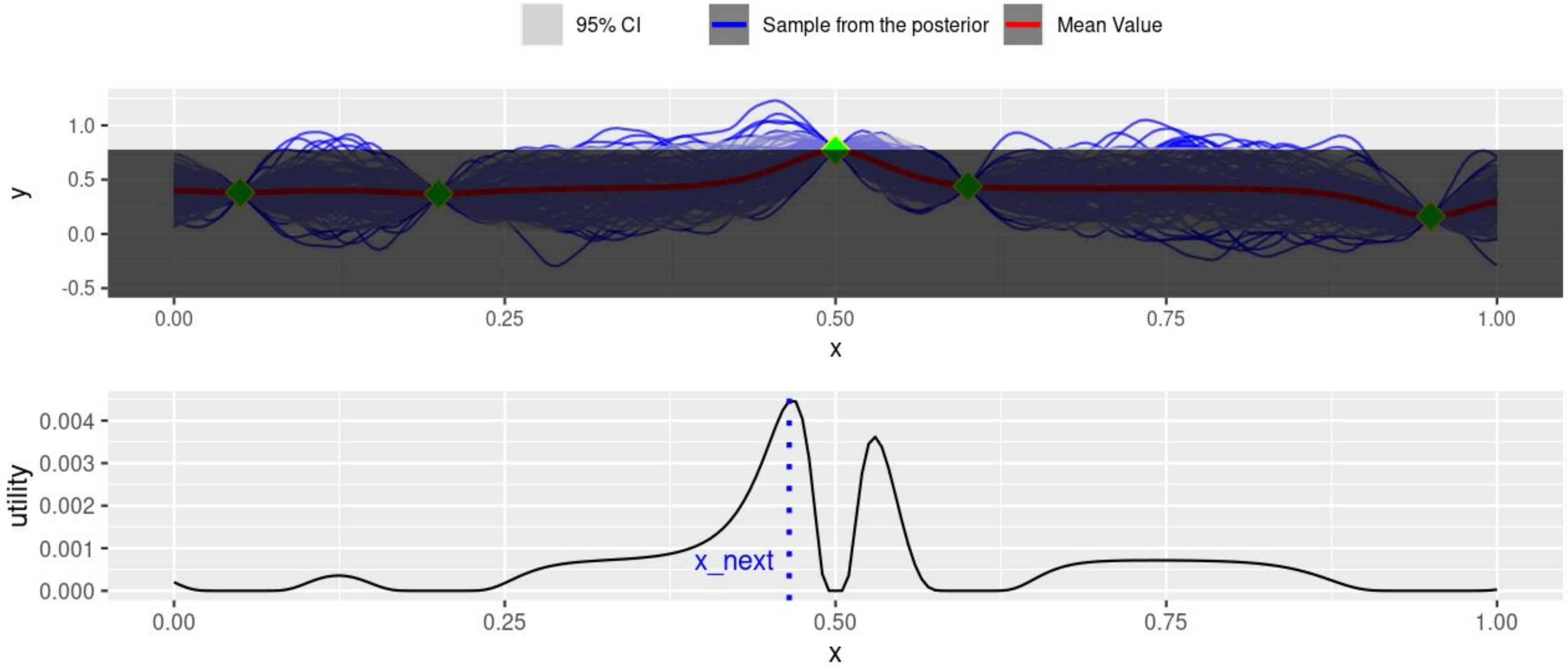
Example of Bayesian Optimization:

Step III) How to find the next (*x_next*) , to feed into the expensive function?



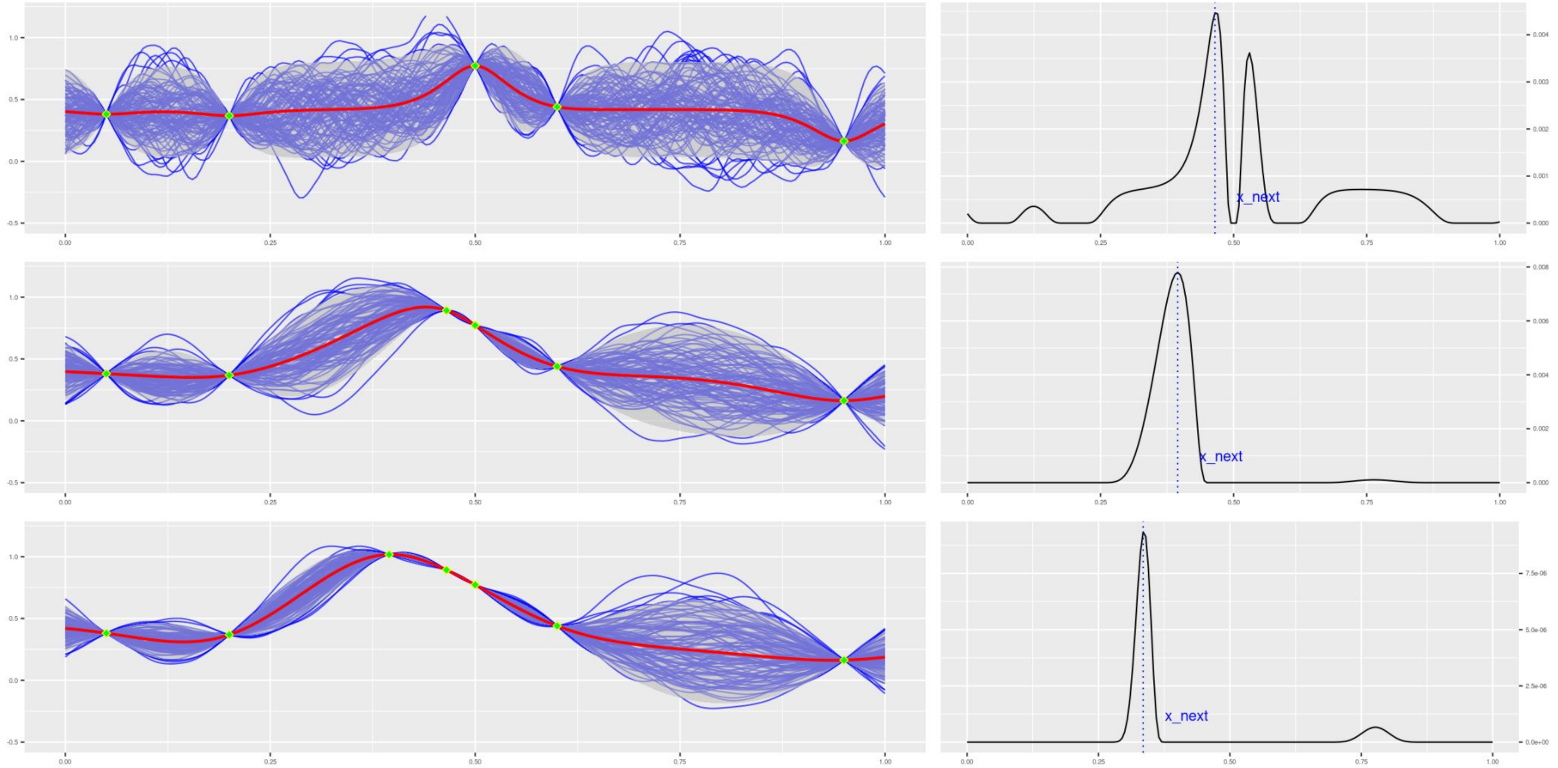
Example of Bayesian Optimization:

Step III) How to find the next (*x_next*) , to feed into the black-box (expensive) function?



Example of Bayesian Optimization:

Two Iterations of BO



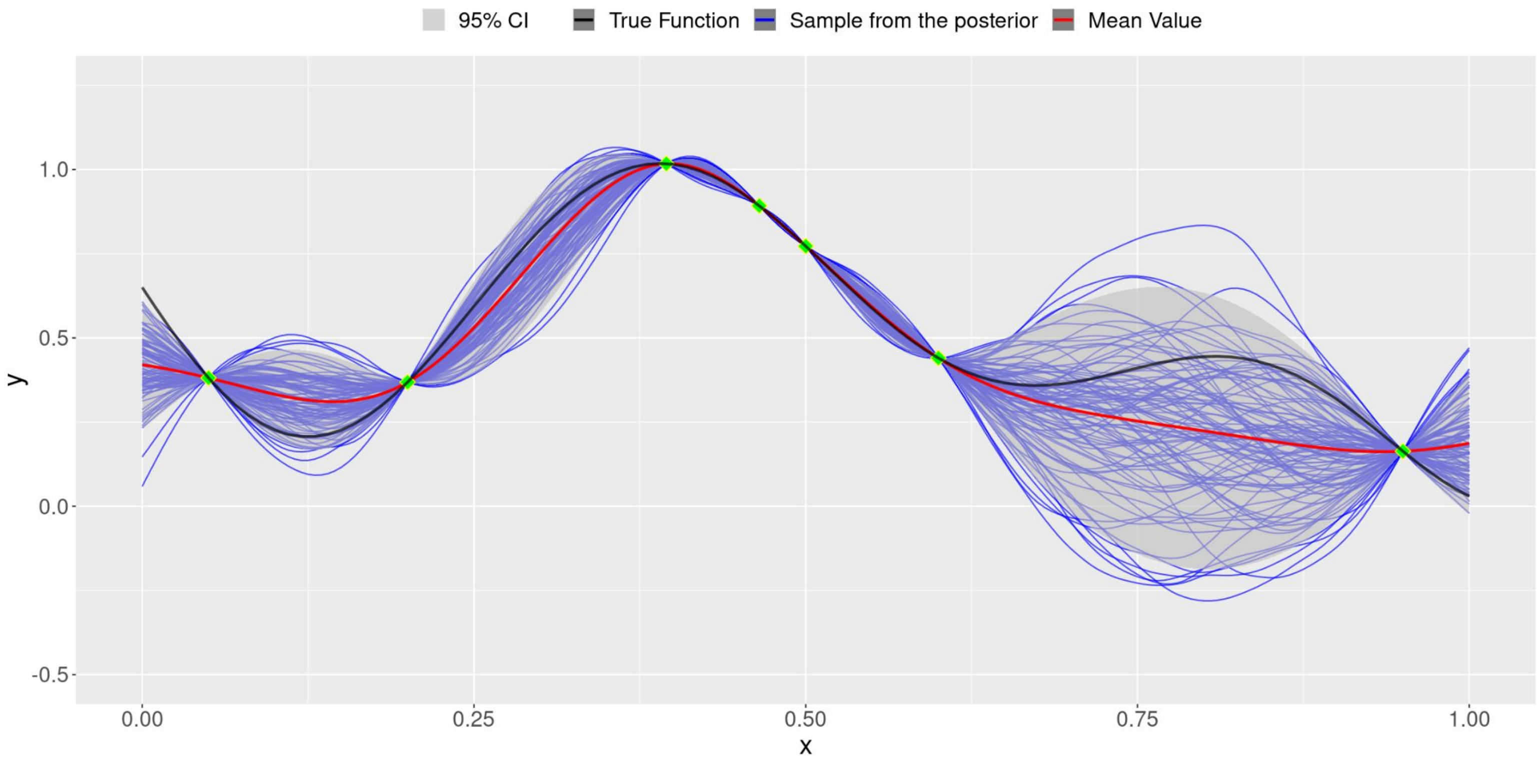
Final Solution of Bayesian Optimization (I):

- The True Function

- $y = 1 - \frac{1}{2}\left(\frac{\sin(12\mathbf{x})}{1+\mathbf{x}} + 2\cos(7\mathbf{x})\mathbf{x}^5 + 0.7\right)$

- The global optima of the function is x_M = 0.39
- After just two iterations, BO solution is = 0.385
- We are around 0.385/0.390 = 98%

Final Solution of Bayesian Optimization (II)



Utilizing Bayesian Optimization to Optimize Water Flooding Management in Reservoirs

BO Manuscript:

Reservoir Production Management with Bayesian Optimization: Achieving Robust Results in a Fraction of the Time

Peyman Kor*, **Aojie Hong** and **Reidar Brumer Bratvold**, Energy Resources Department, University of Stavanger, Norway

*Corresponding author; email: peyman.kor@uis.no

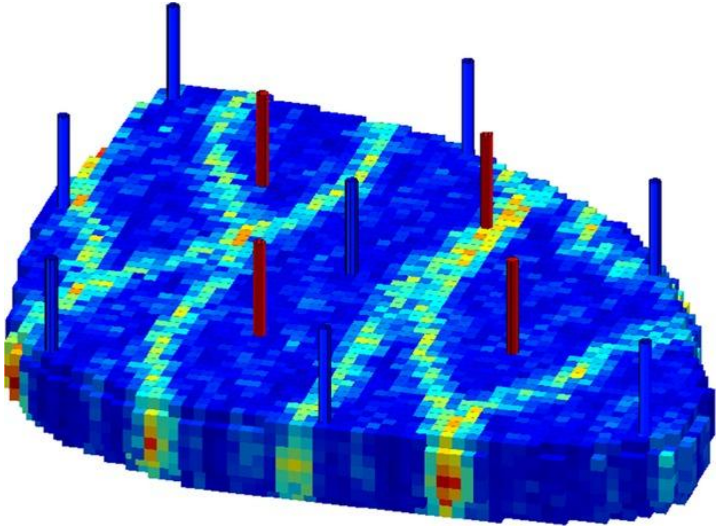
Keywords: Optimization; Gaussian Process; Probabilistic Modeling; Bayesian

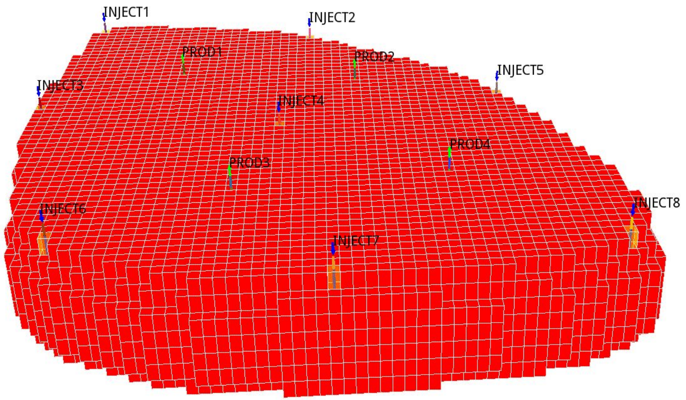
Summary

In well-control (production) optimization, the computational cost of conducting a full-physics flow simulation on a 3D, rich grid-based model poses a significant challenge. This challenge is exacerbated in a robust optimization (RO) setting, where flow simulation must be repeated for numerous geological realizations, rendering RO impractical for many field-scale cases. In this paper, we introduce and discuss a new optimization workflow that addresses this issue by providing computational efficiency, i.e., achieving a near-global optimum of the predefined objective function with minimal forward model (flow-simulation) evaluations. In this workflow, referred to as "Bayesian Optimization" the objective function for samples of decision (control) variables is first computed using a proper design experiment. Then, given the samples, a Gaussian Process Regression (GPR) is trained to mimic the surface of the objective function as a surrogate model. While balancing the dilemma to select the next control variable between high mean, low uncertainty (exploitation) or low mean, high uncertainty (exploration), a new control variable is selected, and flow simulation is run for this new point. Later, the GPR is updated, given the output of the flow simulation. This process continues sequentially until the termination criteria are satisfied. To validate the workflow and obtain a better insight into the detailed steps, we first optimized a 1D problem. The workflow is then implemented for a 3D synthetic reservoir model to perform robust optimization in a realistic field scenario. Finally, the workflow is compared with two other commonly used gradient-free algorithms in the literature: Particle Swarm Optimization (PSO) and Genetic Algorithm (GA). Our results demonstrate that the workflow achieves the same near-optimal solution as GA and PSO while requiring only a fifth of the computational time. We conclude that our method significantly accelerates the optimization process for real-world 3D optimization tasks, potentially reducing CPU times by days or months while producing robust results that lead to a near-optimal solution.

Problem Definition:

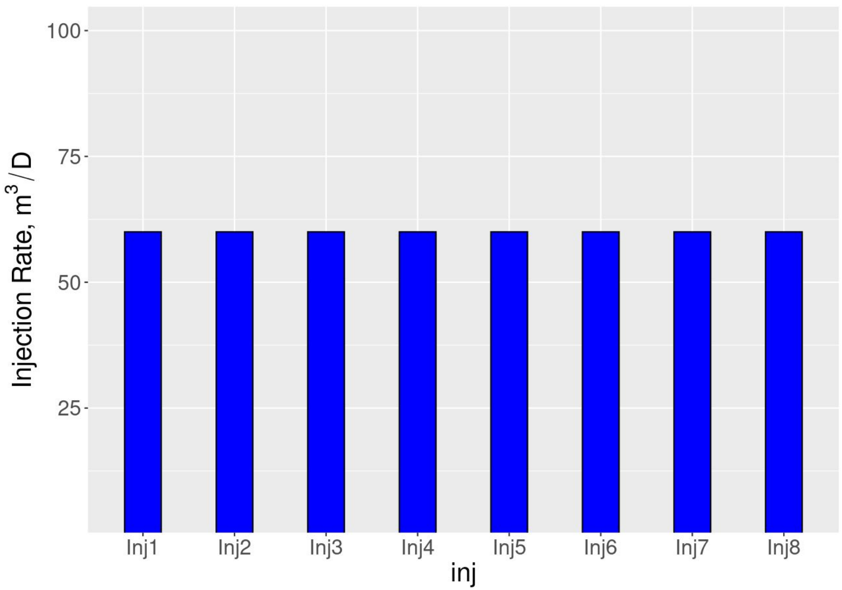
- What is the optimum water injection rate for the eight injector wells?





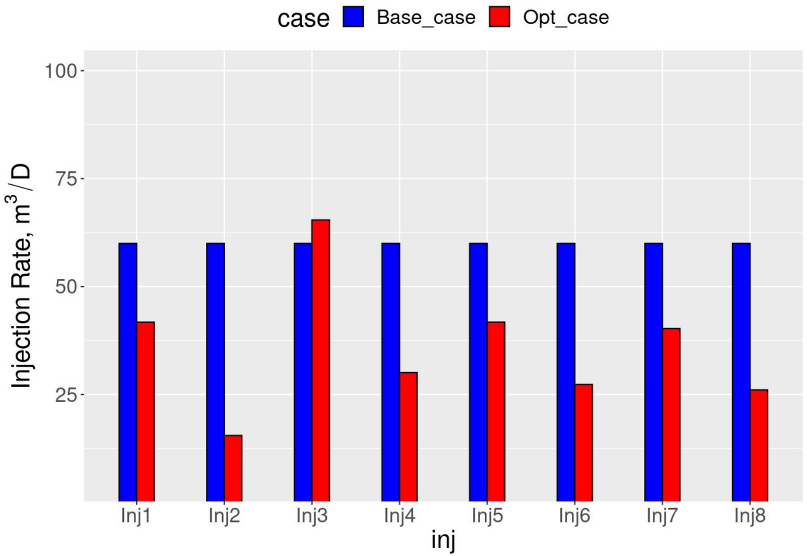
Improvement in Profit Compared to Base-case

- Base Case



Objective Function : Net Present Value (NPV) =
~ 34.5 MM USD

- Optimum Case:



Objective Function : Net Present Value (NPV) =
~ 36.8 MM USD

Tech Companies Currently Active in the Area of Bayesian Optimization?

DeepMind

Bayesian Optimization in AlphaGo

Nando de Freitas
@NandoDF

Bayesian optimisation played an important role in tuning AlphaGo, spearheaded the field of interactive machine learning, and will continue growing as one of the most important topics in AI because interaction with humans matters.

> **Jasper** @latentjasper · Apr 27
> Can we please now put to rest the idea that random search was ever even remotely competitive? "Bayesian Optimization is Superior to Random Search for Machine Learning Hyperparameter Tuning: Analysis of the Black-Box Optimization Challenge 2020" arxiv.org/abs/2104.10201

11:46 AM · Jun 8, 2021 · Twitter for iPhone

30 Retweets **207** Likes

Paper: Bo in AlphaGo

Bayesian Optimization in AlphaGo

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

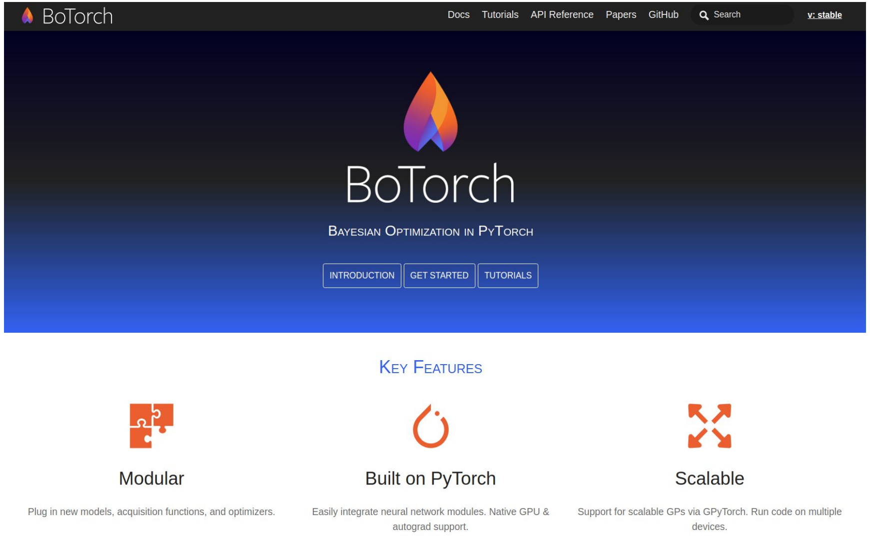
DeepMind, London, UK
yutianc@google.com

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.

Meta (Facebook)

Python Package



Research Team:

Thank You!