BOAT: Building Auto-Tuners with Structured Bayesian Optimization

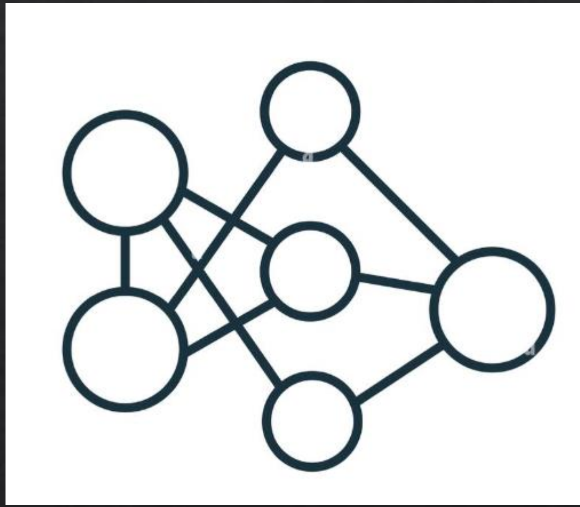
Paper Authors: Valentin Dalibard, Michael Schaarschmidt, Eiko Yoneki
Presenter: Jiahao Gai

Background: Black Box Optimizer

Learning rate

…

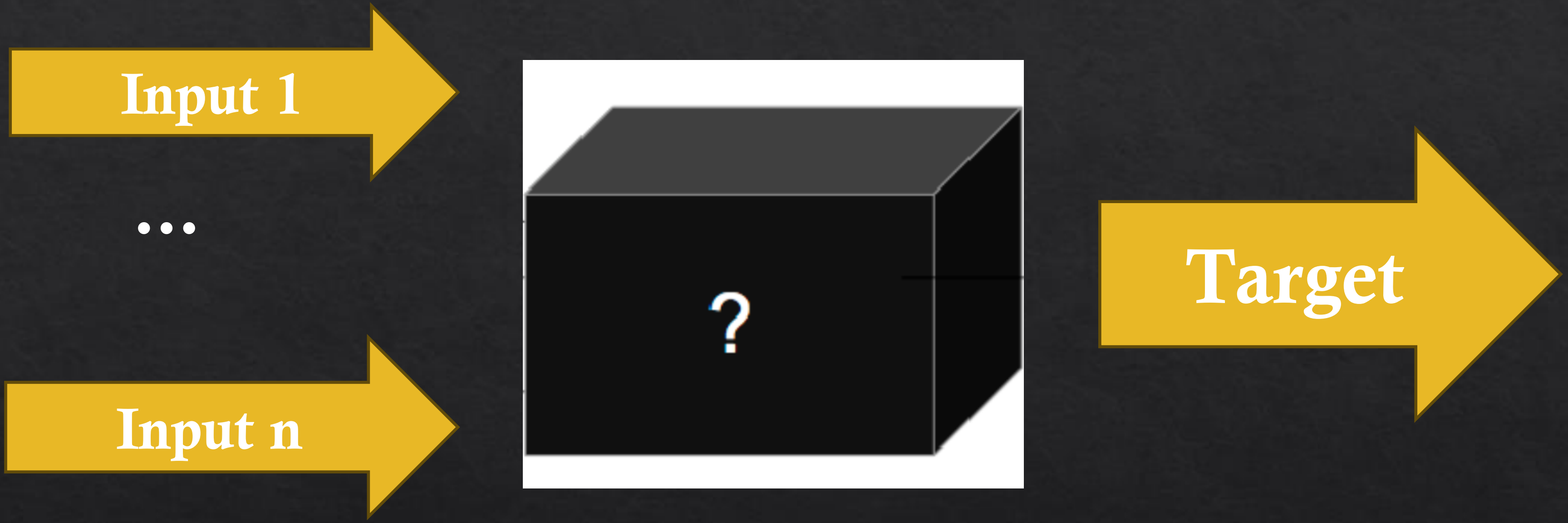
Batch size



Neural network

Precision

Background: Black Box Optimizer



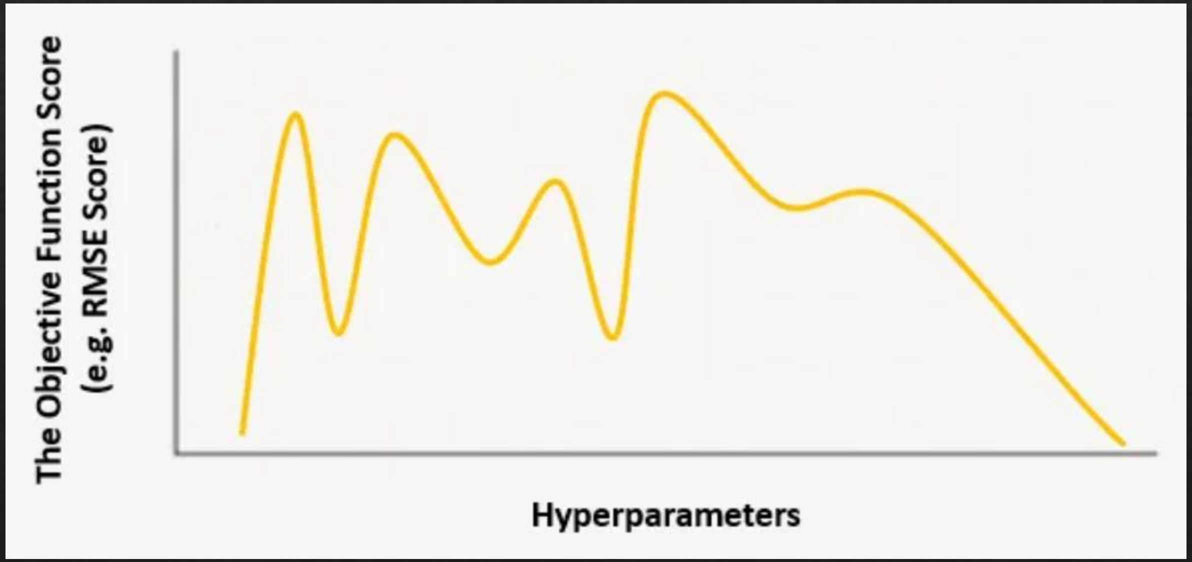
Black Box Problem

Background: Bayesian optimization

Bayesian Optimization builds a probability model of the objective function, which is used to select hyperparameters to evaluate in the true objective function.

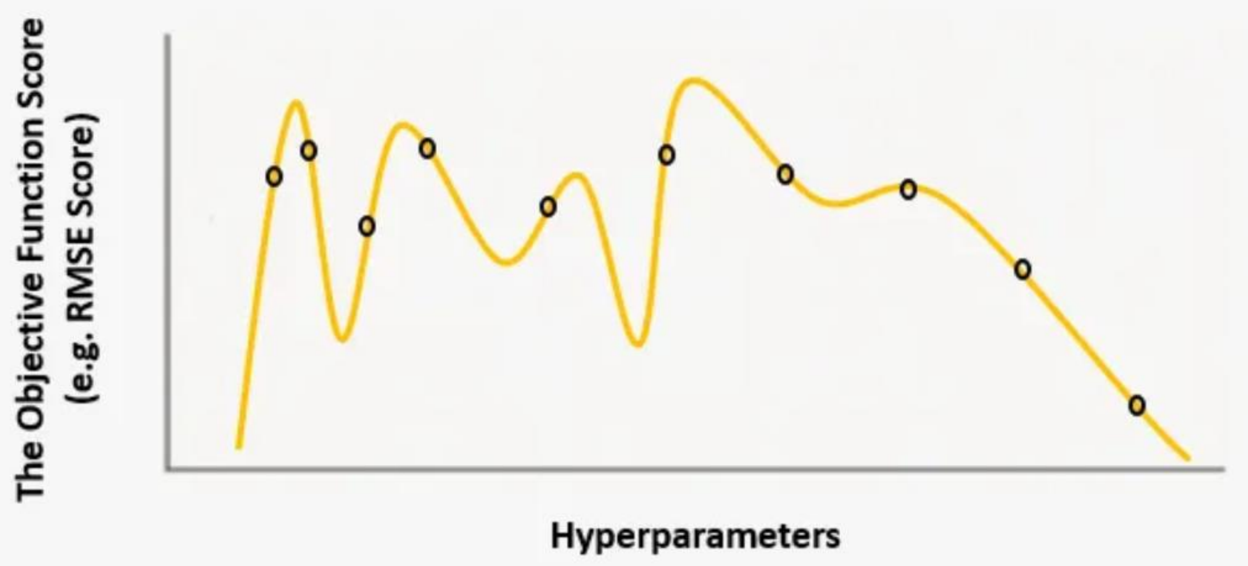
Background: Bayesian optimization

"Bayesian Optimization builds a probability model of the objective function"



Background: Bayesian optimization

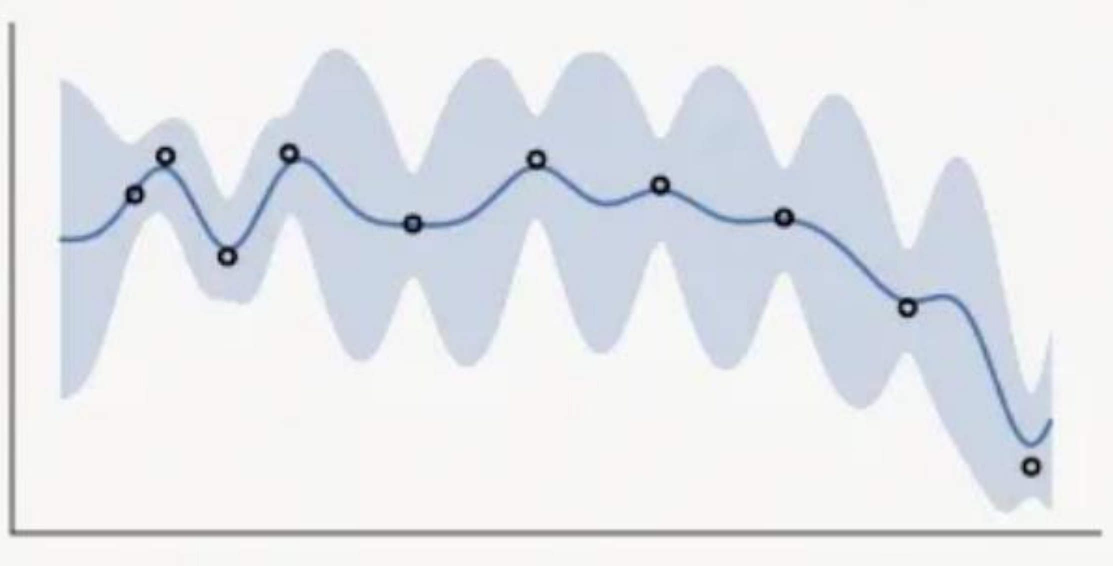
"Bayesian Optimization builds a probability model of the objective function"



10 samples from the true objective function

Background: Bayesian optimization

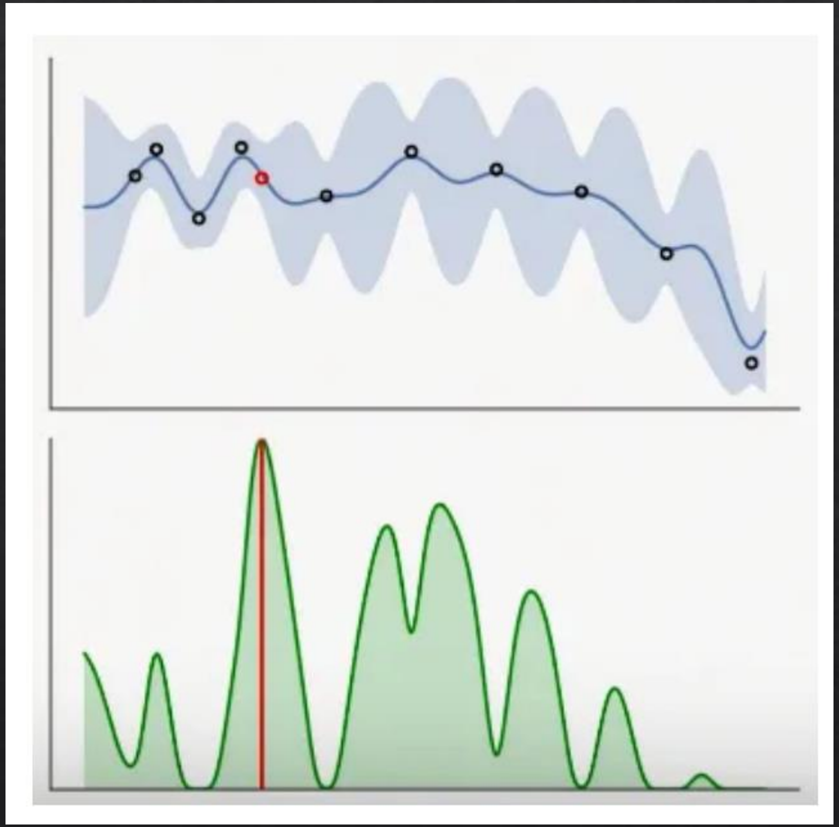
"Bayesian Optimization builds a probability model of the objective function"



Initiate the surrogate model

Background: Bayesian optimization

“and use it to select hyperparameters”



Initiate the surrogate model

Background: Bayesian optimization

Put it Altogether

Algorithm 1 The Bayesian optimization methodology

Input: Objective function $f()$
Input: Acquisition function $\alpha()$
1: Initialize the Gaussian process G
2: **for** $i = 1, 2, \ldots$ **do**
3: Sample point: $\mathbf{x}_t \leftarrow \arg\max_{\mathbf{x}} \alpha(G(\mathbf{x}))$
4: Evaluate new point: $y_t \leftarrow f(\mathbf{x}_t)$
5: Update the Gaussian process: $G \leftarrow G \mid (\mathbf{x}_t, y_t)$
6: **end for**

Background: Bayesian optimization

AKA. **Sequential Model-Based** Optimization

Problem Statement

unsuccessful at tackling optimizations in high dimensional space

1) Cannot accurately capture the objective function landscape after a reasonable number of iterations due to the **curse of dimensionality**

2) The numerical optimization algorithm, used in each iteration, fails to converge and find a promising point.

Structured Bayesian Optimization (SBO)

A novel extension of Bayesian optimization capable of leveaging bespoke probabilistic models to rapidly converge to high-performance configurations.

Structured Bayesian Optimization (SBO)

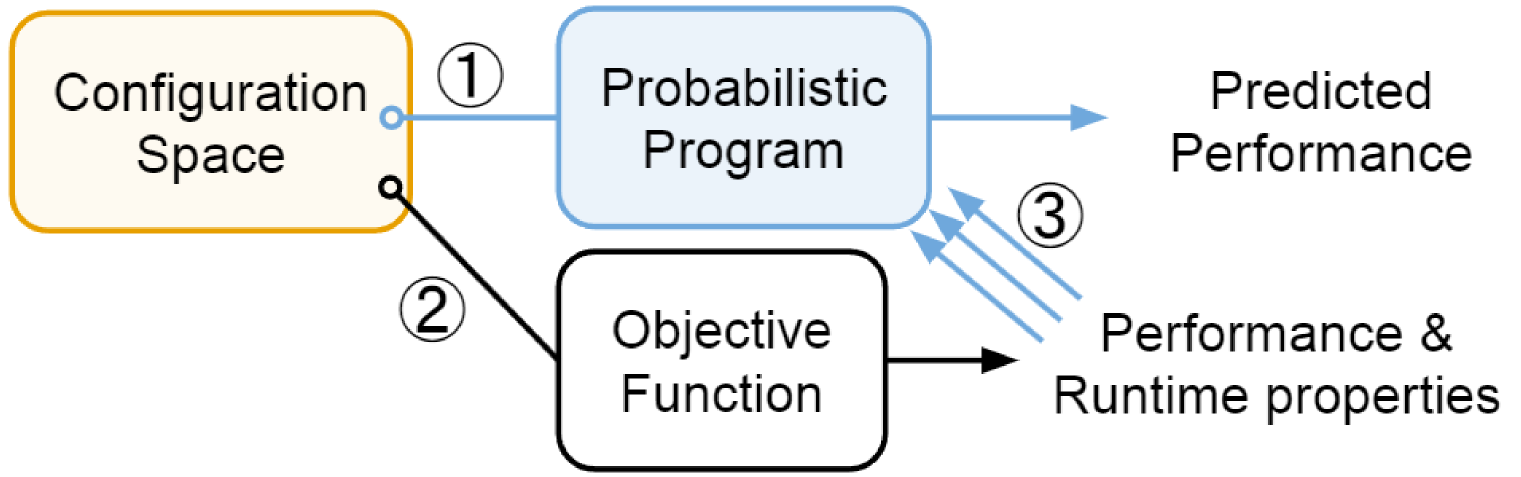


Figure 1: Procedure of Structured Bayesian Optimization

Two advantages of SBO

1. It captures the user's understanding of the behavior of the system

2. Using such a model allows us to monitor runtime properties reflected in the model and use them for inference.

BOAT (BespOke Auto-Tuner)

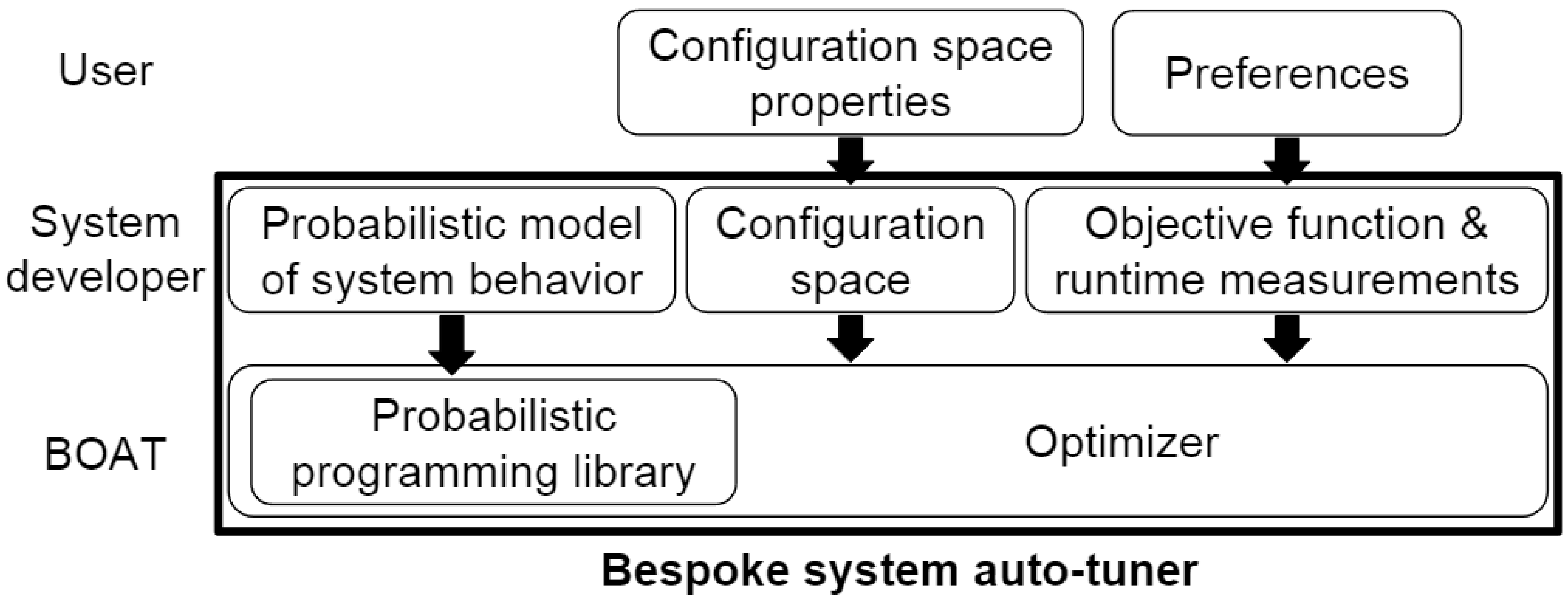


Figure 3: Flow of data when using BOAT

Design of BOAT Framework

- Semi-parametric models

- DAG models

Evaluation

The evaluation focuses
on quantifying two properties:

- The benefits of auto-tuning

- The need for a bespoke auto-tuner.

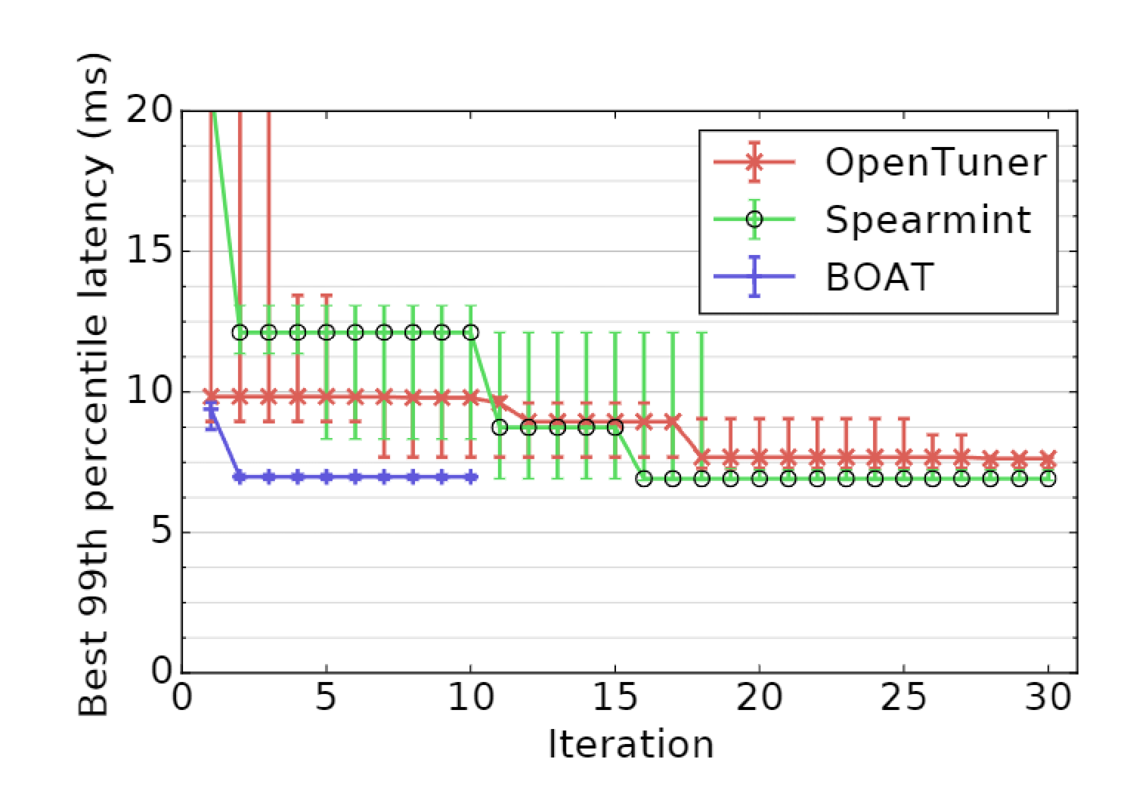
in two case studies:

- Garbage collection

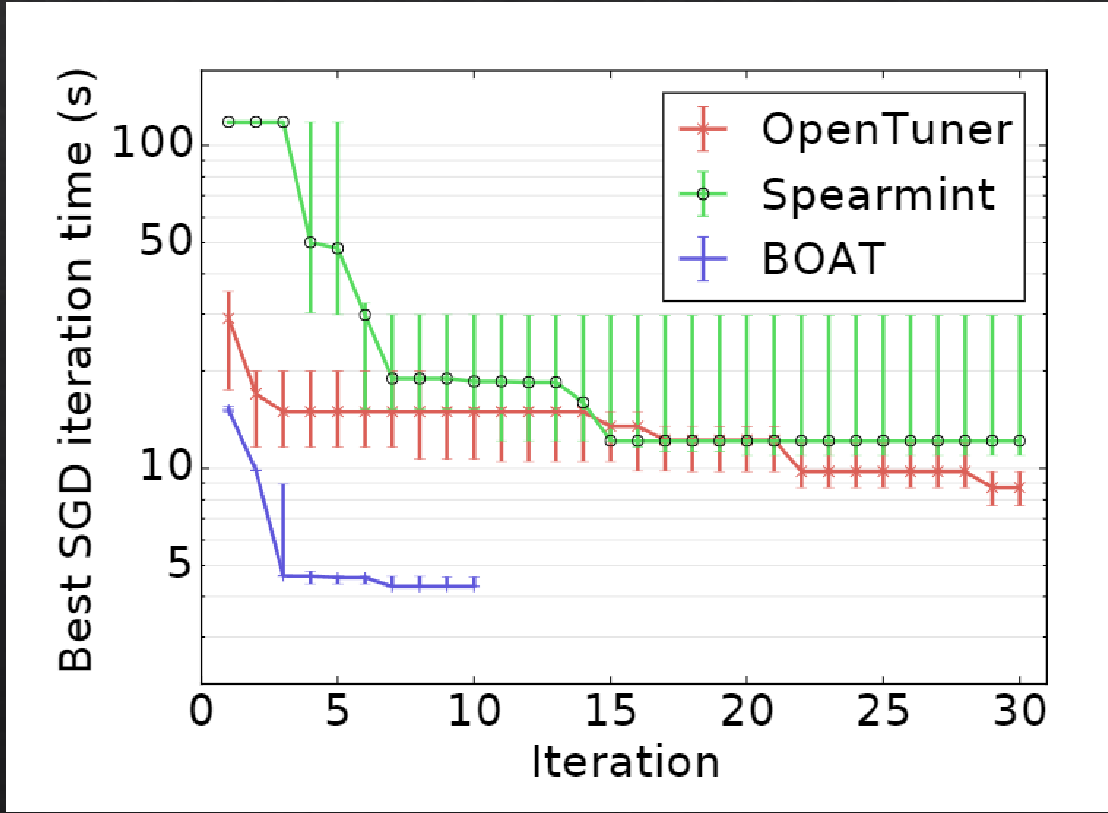
- Neural networks

Evaluation

Comparison with other auto-tuners in two case studies



Garbage collection case study



Neural network case study

Pros

- Handling Complex Configuration Spaces
- Reduced Number of Iterations for Convergence
- Global Performance Portability

Cons

- Complexity in Building Probabilistic Models
- Potential Failure in Capturing Objective Function Landscape

Reverences

- Dalibard, V., Schaarschmidt, M., & Yoneki, E. (2017). BOAT: Building Auto-Tuners with Structured Bayesian Optimization.
- Dalibard, V. (2017). A framework to build bespoke auto-tuners with structured Bayesian optimisation.
- Wang, W. (2020). Bayesian Optimization Concept Explained in Layman Terms.
- Snoek, J., Larochelle, H., & Adams, R. P. (2012). Practical Bayesian Optimization of Machine Learning Algorithms.

Questions?