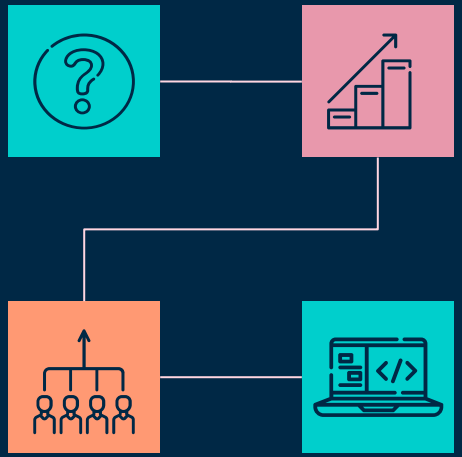
Optimizing Asynchronous Successive Halving (ASH) for Hyperparameter Tuning

Evan and Roy

OUTLINE

PROBLEM

What are hyperparameters and their relevance



MODEL AND DATA

Examining ASH and the Iris dataset

WHY PARALLEL?

Motivation for parallelization and optimization

IMPLEMENTATION

How we plan on parallelizing the app

THE PROBLEM

What are hyperparameters?

- ML models use them to influence **predictive accuracy** and **output**
- Examples include:
  - Learning rate
  - Dropout rate
  - Activation function

Why are they important?

- Key to **performance improvement** and catering to **hypersensitive** ML models
- Current techniques include:
  - **Manual Tuning**
  - **Brute Force**: Grid Search and Random Search

CMU ML Blog

MODEL AND DATA

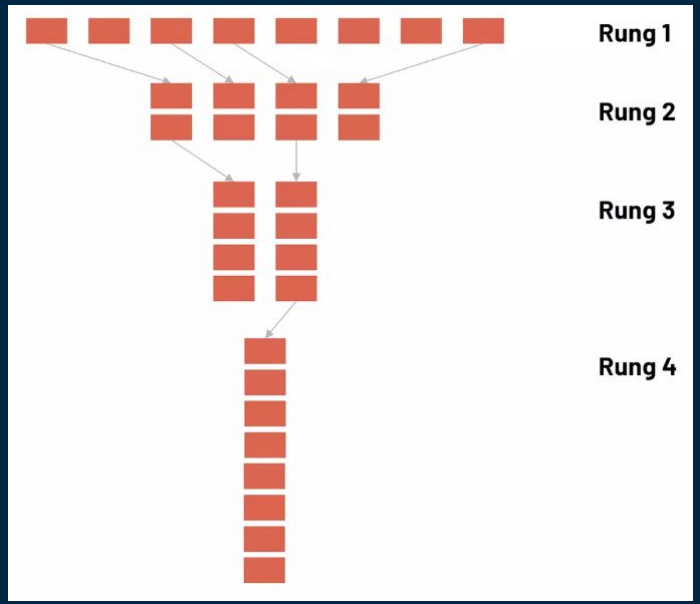
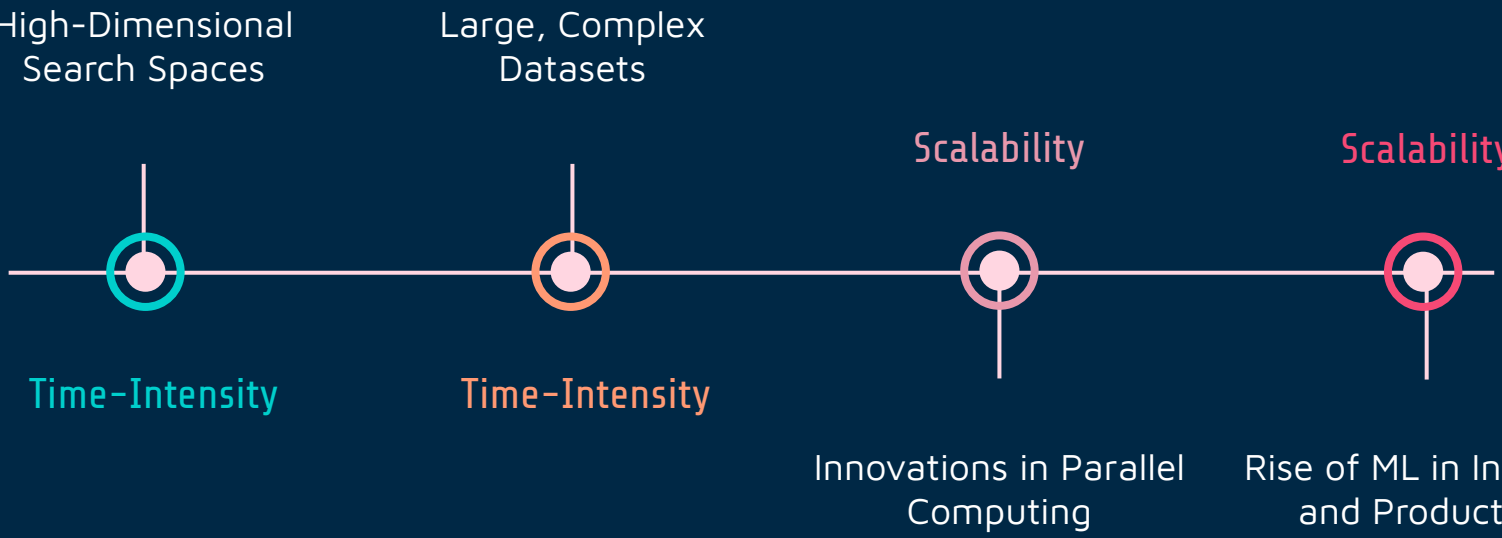


Figure 1. Job allocation and promotion in Successive Halving

- SEQUENTIAL BASELINE: SUCCESSIVE HALVING (SH)
  - Uniform resource allocation across candidates per rung
- PARALLELIZATION MECHANISM (ASH)
  - Training multiple candidate models simultaneously
- CONTEXT OF THE MODEL
  - Optimizing an ML model's hyperparameters
- LIMITATIONS TO THE MODEL
  - Robustness, computational cost
- DATA SOURCE AND TRANSFORMATION
  - Iris Dataset – transformed from API to C++ vectors

Source: Li, L. (2019, December 19). Massively parallel hyperparameter optimization. ML@CMU | Carnegie Mellon University. Retrieved March 22, 2023, from https://blog.ml.cmu.edu/2018/12/12/massively-parallel-hyperparameter-optimization/

WHY PARALLEL?



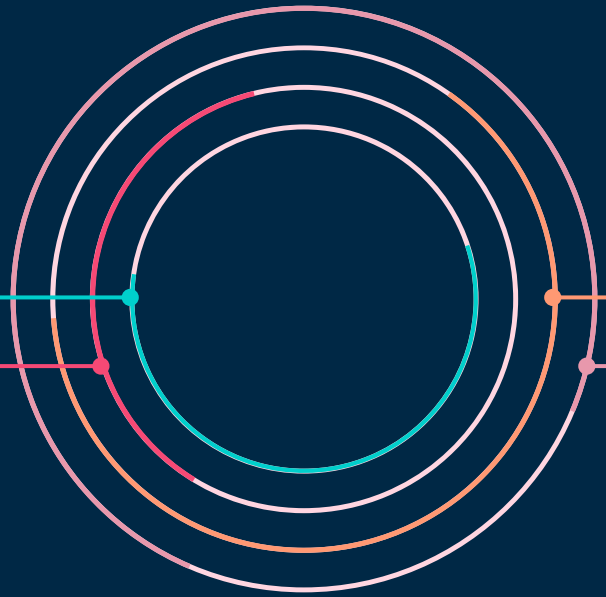
IMPLEMENTATION PLAN

Shared memory – **multi-threading** the training process for individual candidate models

openMP

openMPI

Simultaneously running jobs and enables **message-passing** between multiple processors



Concurrently calculate gradients for batch gradient descent when training a model

Why openMP?

Why openMPI?

Desynchronize jobs and **communicate** which hyperparameter configurations are to be promoted to next rung

WORKS CITED

Li, L. (2019, December 19). Massively parallel hyperparameter optimization. ML@CMU | Carnegie Mellon University. Retrieved March 22, 2023, from https://blog.ml.cmu.edu/2018/12/12/massively-parallel-hyperparameter-optimization/

Li, L., Jamieson, K., Rostamizadeh, A., Gonina, E., Ben-Tzur, J., Hardt, M., Recht, B., Talwalkar., A. (2020, March 15). A System For Massively Parallel Hyperparameter Tuning. Retrieved March 22, 2023, from https://arxiv.org/pdf/1810.05934.pdf