

Speedup Techniques for Hyperparameter Optimization

Hyperband

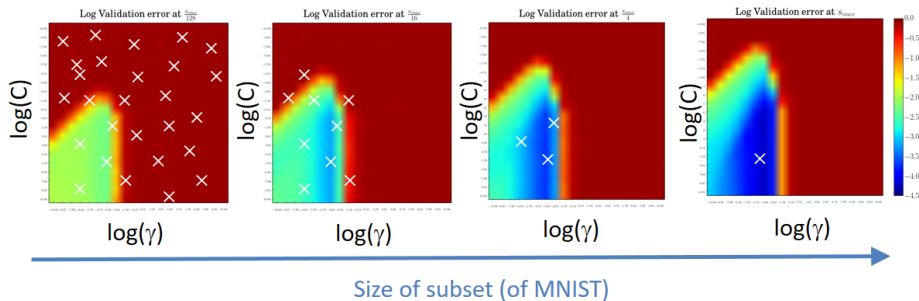
Bernd Bischl Frank Hutter Lars Kotthoff
Marius Lindauer Joaquin Vanschoren

A Simple Multi-Fidelity Algorithms: Successive Halving (SH)

[Jamieson and Talwalkar, AISTATS 2016]

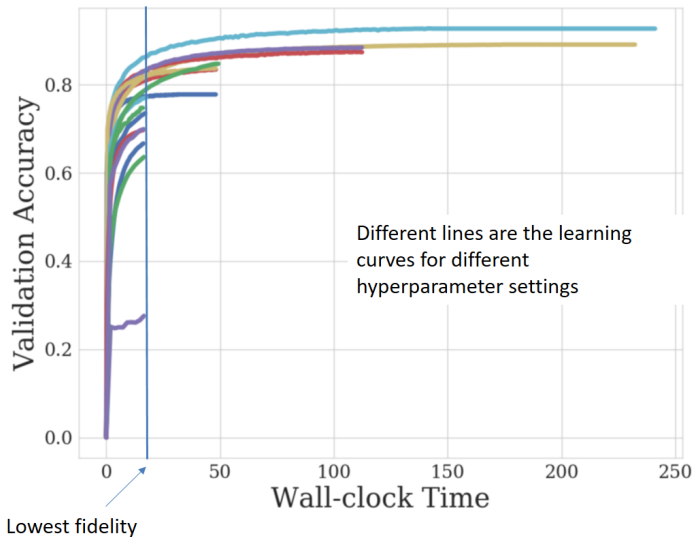
- A very simple algorithm:

- ▶ Sample N configurations uniformly at random & evaluate them on the cheapest fidelity
- ▶ Keep the best half (or third), move them to the next fidelity
- ▶ Iterate until the most expensive fidelity (= original expensive black box)



The Same SH Algorithm When the Fidelity is Runtime

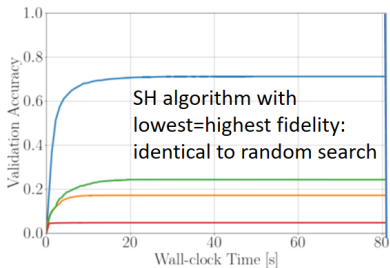
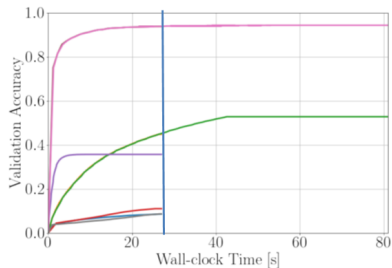
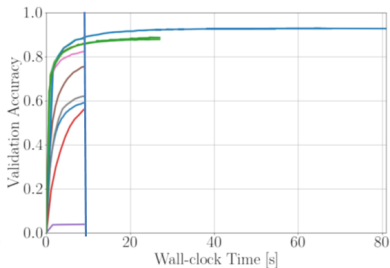
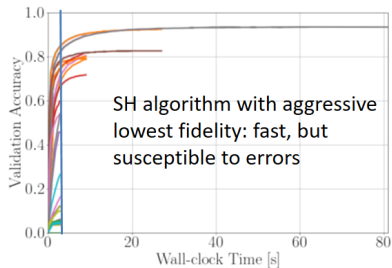
[Jamieson and Talwalkar, AISTATS 2016]



An Extension of SH with Theoretical Guarantees: Hyperband

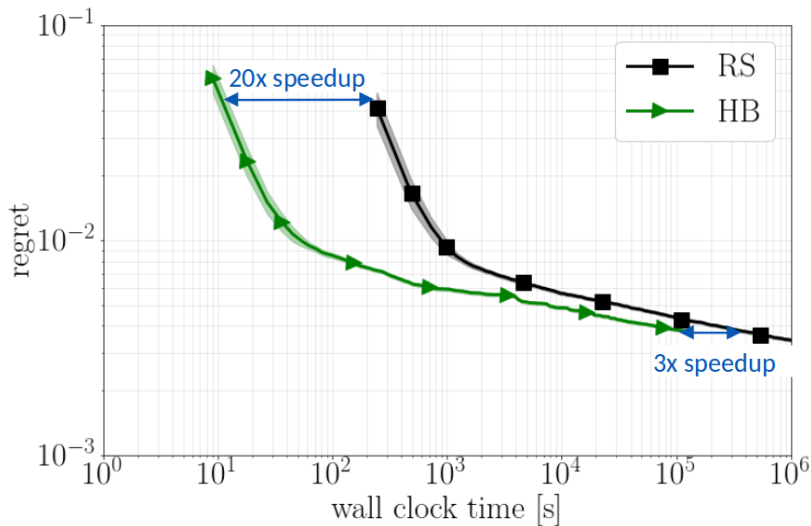
[Li et al., JMLR 2018]

- Main Idea:
hedge against errors in cheap approximations
- Algorithm:
run multiple copies of SH in parallel, starting at different cheapest fidelities



Empirical Evaluation: Hyperband vs. Random Search

[Falkner, Klein & Hutter, ICML 2018]



Questions to Answer for Yourself / Discuss with Friends

- **Discussion.** How do you think Hyperband would compare to successive halving using the most aggressive fidelity?
- **Discussion.** How slow is Hyperband in the worst case?