



Large scale hyperparameter searches

# High Performance Deep Learning

Eric Marcus & Jonas Teuwen



# Overview

## Hyperparameter searching

- Goal

## Building blocks

- Hydra + Lightning
- Slurm-based clusters & Submitit
- MLFlow

# Hyperparameter searching

## Goal

- As little labor as possible: change only 1 line of code – the hyperparameters
- Track our experiments convenient and live
- Automated searches that optimize a metric
- Scalability to large compute



# Hyperparameter searching

## Goal

- Example code:

<https://github.com/NKI-AI/hyperparameter-search-template>

# Hyperparameter searching

## Code outline

- Set up a selection of hyperparameters (in bash script) and sbatch it
- 1: Reserve CPU-node – the hyperparameter master node
- 2: Run *mlflow* server
- 3: Determine how many nodes should be spawned
- 4: Automatically spawn nodes, each with a set of hyperparameters
- 5: All nodes report their results to the master node's mlflow
- We can sit back and relax as the results roll in

# Hyperparameter searching

mlflow
Experiments
Models
GitHub Docs

Experiments + <

Search Experiments

- hyperparam\_rhpc\_test 🔗 🗑️
- test\_experiment 🔗 🗑️
- bash\_test\_experiment 🔗 🗑️
- hyperparam\_mnist\_test 🔗 🗑️
- hyperparam\_mnist\_b... 🔗 🗑️

## hyperparam\_mnist\_bayesian 📄

[Share](#)

Experiment ID: 4

▶ Description [Edit](#)

🔄 Refresh
Compare
Delete
Download CSV 📄
↓ Start Time ▾
All time ▾

☰
🗒️
⚙️ Columns
Only show differences 
🔍 metrics.rmse < 1 and params.model = "tree"
Search
Filter
Clear

Showing 4 matching runs

|                          |              |          |          |              | Metrics > |           |            | Parameters >    |                |                |
|--------------------------|--------------|----------|----------|--------------|-----------|-----------|------------|-----------------|----------------|----------------|
| <input type="checkbox"/> | ↓ Start Time | Duration | User     | Source       | epoch     | train/acc | train/loss | model/net/lin1_ | model/params/t | model/params/t |
| <input type="checkbox"/> | 🟢 1 day ago  | 1.1min   | e.marcus | 📄 _submit.py | 19        | 0.994     | 0.021      | 512             | 603658         | 603658         |
| <input type="checkbox"/> | 🟢 1 day ago  | 38.6s    | e.marcus | 📄 _submit.py | 9         | 0.992     | 0.027      | 512             | 603658         | 603658         |
| <input type="checkbox"/> | 🟢 1 day ago  | 1.2min   | e.marcus | 📄 _submit.py | 19        | 0.994     | 0.02       | 256             | 336650         | 336650         |
| <input type="checkbox"/> | 🟢 1 day ago  | 36.6s    | e.marcus | 📄 _submit.py | 9         | 0.993     | 0.024      | 256             | 336650         | 336650         |

[Load more](#)

# Building blocks

## Hydra + Lightning

- Configuration framework
- Keep track of your configurations
- Options: *structured* config or yaml-based or validation-scheme
- Every config can be overridden in the command-line as well:

```
python train.py trainer.gpus=1
```



# Building blocks

## Hydra + Lightning

- Sweepers
- Automatic hyperparameter searching
- Optimizes one or more metrics



# Building blocks

## Slurm-based clusters & SubmitIt

- Slurm – Simple Linux Utility for Resource Management
- Schedules jobs on available compute
- Keeps track of the queue
  
- We want to automatically make slurm do the job submissions
- SubmitIt does those submissions

# Building blocks

## Slurm-based clusters & SubmitIt

- *Example: resources (squeue)*  
after a submission

| JOBID      | PARTITION | NAME     | USER    | ST | TIME  | NODES | NODELIST(REASON) |
|------------|-----------|----------|---------|----|-------|-------|------------------|
| 8817301_26 | gpu_titan | train_li | nki_eri | R  | 1:32  | 1     | r34n5            |
| 8817301_25 | gpu_titan | train_li | nki_eri | R  | 3:44  | 1     | r34n6            |
| 8817301_23 | gpu_titan | train_li | nki_eri | R  | 4:14  | 1     | r34n6            |
| 8817301_24 | gpu_titan | train_li | nki_eri | R  | 4:14  | 1     | r34n6            |
| 8817301_19 | gpu_titan | train_li | nki_eri | R  | 34:32 | 1     | r28n2            |
| 8817301_20 | gpu_titan | train_li | nki_eri | R  | 34:32 | 1     | r28n2            |
| 8817301_17 | gpu_titan | train_li | nki_eri | R  | 35:19 | 1     | r28n4            |
| 8817301_18 | gpu_titan | train_li | nki_eri | R  | 35:19 | 1     | r28n4            |
| 8817301_22 | gpu_titan | train_li | nki_eri | R  | 34:19 | 1     | r28n4            |
| 8817301_21 | gpu_titan | train_li | nki_eri | R  | 34:32 | 1     | r28n2            |
| 8817301_0  | gpu_titan | train_li | nki_eri | R  | 45:54 | 1     | r28n3            |
| 8817301_1  | gpu_titan | train_li | nki_eri | R  | 45:54 | 1     | r28n3            |
| 8817301_2  | gpu_titan | train_li | nki_eri | R  | 45:54 | 1     | r29n2            |
| 8817301_3  | gpu_titan | train_li | nki_eri | R  | 45:54 | 1     | r29n2            |
| 8817301_4  | gpu_titan | train_li | nki_eri | R  | 45:54 | 1     | r35n4            |
| 8817301_5  | gpu_titan | train_li | nki_eri | R  | 45:54 | 1     | r35n4            |
| 8817301_6  | gpu_titan | train_li | nki_eri | R  | 45:54 | 1     | r35n4            |
| 8817301_7  | gpu_titan | train_li | nki_eri | R  | 45:54 | 1     | r36n2            |
| 8817301_8  | gpu_titan | train_li | nki_eri | R  | 45:54 | 1     | r36n2            |
| 8817301_9  | gpu_titan | train_li | nki_eri | R  | 45:54 | 1     | r36n2            |
| 8817301_10 | gpu_titan | train_li | nki_eri | R  | 45:54 | 1     | r36n3            |
| 8817301_11 | gpu_titan | train_li | nki_eri | R  | 45:54 | 1     | r36n3            |
| 8817301_12 | gpu_titan | train_li | nki_eri | R  | 45:54 | 1     | r36n4            |
| 8817301_13 | gpu_titan | train_li | nki_eri | R  | 45:54 | 1     | r36n4            |
| 8817301_14 | gpu_titan | train_li | nki_eri | R  | 45:54 | 1     | r36n4            |
| 8817301_15 | gpu_titan | train_li | nki_eri | R  | 45:54 | 1     | r36n5            |
| 8817301_16 | gpu_titan | train_li | nki_eri | R  | 45:54 | 1     | r36n5            |
| 8817299    | shared    | clavicle | nki_eri | R  | 48:59 | 1     | r15n1            |



# Building blocks

## MLFlow

- Logging
- Ideal for hyperparameter experiments
- Comparing runs
- Saving model checkpoints



# Recap

## Goal

- As little labor as possible: change only 1 line of code – the hyperparameters
- Track our experiments convenient and live
- Automated searches that optimize a metric
- Scalability to large compute