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
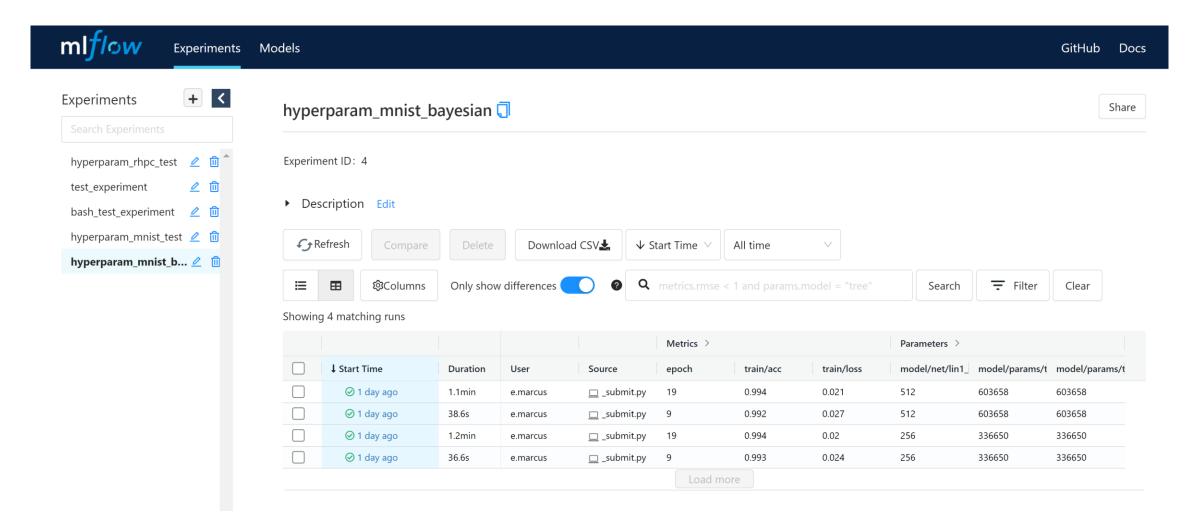- Scalability to large compute

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

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

UNIVERSITY OF AMSTERDAM

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