



Contribution ID: 670 Contribution code: **contribution ID 670**

Type: **Oral**

Hyperparameter Optimization of Data-Driven AI models on HPC Systems

Wednesday, 1 December 2021 17:20 (20 minutes)

In the European Center of Excellence in Exascale Computing “Research on AI- and Simulation-Based Engineering at Exascale” (CoE RAISE), researchers from science and industry develop novel, scalable Artificial Intelligence technologies towards Exascale. In this work, we leverage HPC resources to perform large scale hyperparameter optimization using distributed training on multiple compute nodes, each with multiple GPUs. This is part of CoE RAISE’s work on data-driven use-cases towards Exascale which leverages AI- and HPC cross-methods developed within the project.

Hyperparameter optimization of deep learning-based AI models is often compute resource intensive, partly due to the high cost of training a single hyperparameter configuration to completion, and partly because of the infinite set of possible hyperparameter combinations to evaluate. There is therefore a need for large scale, parallelizable and resource efficient hyperparameter search algorithms. We benchmark and compare different search algorithms for advanced hyperparameter optimization. The evaluated search algorithms, including Random Search, Hyperband and ASHA, are tested and compared in terms of both final accuracy and accuracy per compute resources spent.

As an example use-case, a graph neural network (GNN) model known as MLPF, which has been developed for the task of Machine Learned Particle-Flow reconstruction, acts as the base model for which hyperparameter optimization is performed. In essence, the MLPF algorithm combines information from tracks and calorimeter clusters to reconstruct charged and neutral hadron, electron, photon and muon candidates.

Further developments of AI models in CoE RAISE have the potential to greatly impact the field of High Energy Physics by efficiently processing the very large amounts of data that will be produced by particle detectors in the coming decades.

In addition, the large scale, parallelizable and resource efficient hyperparameter search algorithms are model agnostic in their nature and could be widely applicable in other sciences making use of AI, for instance in the use-cases of seismic imaging, remote sensing, defect-free additive manufacturing and sound engineering that are part of CoE RAISE WP4.

Significance

This is, to our knowledge, the first comparison of different hypertuning frameworks on HPC systems and should be relevant for anyone wanting to get the maximum performance out of their AI models while leveraging HPC resources.

References

Speaker time zone

Compatible with Europe

Primary authors: WULFF, Eric (CERN); GIRONE, Maria (CERN)

Presenter: WULFF, Eric (CERN)

Session Classification: Track 2: Data Analysis - Algorithms and Tools

Track Classification: Track 2: Data Analysis - Algorithms and Tools