Scaling ML: model selection and workflow

Proposed Milestones

1. **Identify target ML models in collaboration with experiments**
2. Port, train, and run at least two target models on two different HPC systems
3. Compare two data parallel training solutions for at least one target model
4. Compare two hyperparameter optimization tools on at least one target model
5. Setting up a prototype Inference as-a-service platform on at least one DOE HPC system

Task list sign up sheet

Email list: CCE-SML@anl.gov

First meeting in January

ML models

- **Cosmic Frontier**
    - Cosmological simulations, DES adversarial domain adaptation (no info yet)
    - LSST image processing (no info yet)
- **Energy Frontier**
    - Tracking with IaaS
    - Jet tagging with particle based transformers
- **Intensity Frontier**
    - DUNE reconstruction (training and IaaS)

Workflow Scaling Training

- Choose an HPC to train on
  - Examples: Aurora, Perlmutter
  - Has the model already been trained on an HPC? Maybe try a different HPC
- Train on one node/GPU
  - Find out if environment is adequate for training model (e.g., are python packages easy to find)
  - Get training data to HPC
  - Benchmark computational performance
- Train on multiple GPUs on the same node
  - Identify issues
  - Measure computational performance and compare to single GPU
- Train across multiple nodes and multiple GPUs
  - Identify issues and compare performance
- HPO

Workflow for Inference as a Service

- Choose one HPC and/or other IaaS provider
  - Understand/address security and workflow synchronization issues.
- Test IaaS locally on a hybrid CPU/GPU node (as an offloading mechanism)
  - Compare to direct offloading through ONNX or similar
- Test IaaS across HPC interconnect
  - Study latency/throughput scaling vs # of client and server nodes
- Test IaaS over WAN
  - Same as above, monitor WAN latency and bandwidth looking for saturation limit