Machine Learning in ATLAS
Resource Requirements

Walter Hopkins & Daniel Murnane

ATLAS ML Forum Co-conveners

October 2023

ATLAS EXPERIMENT

ATLAS ML Forum Challenges & Opportunities

- 2 meetings held to identify high-priority challenges in physics and CP groups
  - [March 9th 2023]() and [March 16 2023]()
  - Input from 6 physics & 7 CP groups
- A large and diverse set of challenges and possible solutions were raised
- [A document is being drafted]() that compiles all of these topics

Word cloud of the 13 ML Challenges & Opportunities contributions

ATLAS ML Forum Challenges & Opportunities

Broadly, **challenges fall into two categories**:

- Physics/CP based: decreasing uncertainties, reduction of number of uncertainties, follow-through on implementing ML methods is needed
- ML based: need better documentation, access to computing resources, etc.

The **opportunities** are to improve the impact, understanding and access to ML across ATLAS physics and CP groups

We then gathered voting of importance of these challenges and opportunities


Importance Votes
Opportunities & Solutions
Challenges
Data Usage & Availability
Availability of Existing Tools
Casual Advice on ML Techniques
Formal Recommenda...
Central Discussion Location
Construction of Missing Tools
Human Resources...
Common...
Quality Documentation
Data Model Sharing
Hardware Investment
Personpower
Publication Outreach...
Collaboration
Formal Processes & Onboarding
Feedback Loops & Discussion
Dissemination of Open...

ML Infrastructure C&Os

- Approximately 30% of interest is relevant to ML hardware, software & tooling
- Can be mostly summarized as **GPU access** & **data access**

ATLAS EXPERIMENT

Hardware: GPUs

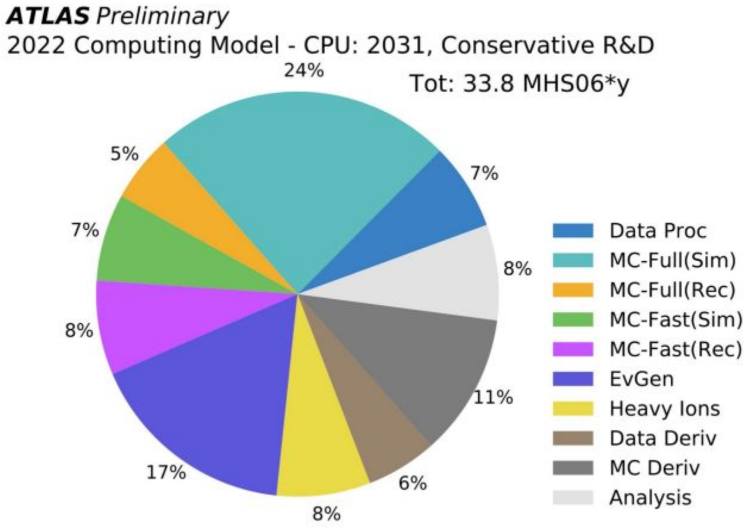
R&D Model Training & Inference

- Assuming all O(20) CP and analysis groups are eventually developing large-parameter, memory-hungry models such as Transformers
- HL-LHC offline reconstruction problems typically require O(days) to train per GPU
- A CP/analysis group may be experimenting with O(10) models in tandem
- In the case of needing to perform large training on a single model, could consume in burst O(10) GPUs
- We could predict a max load of O(200) GPUs across ATLAS CP/analysis groups for model training
- These need not all be premium GPUs, but could be split across (according to budget)
    - O(16) A100 - premium for large model training
    - O(32) V100 - decent for training and fast inference
    - O(64) T4 - decent for inference and simple model training

Hardware: GPUs

Production Inference

- Consider that the bulk of MC simulation and reconstruction could be targeted by GPU-based ML solutions, and some part of analysis and derivations could be accelerated with GPU ML
- We could estimate that O(50%) of the ATLAS computing model could be accelerated by GPU-based ML by 2031



Hardware: Disk Space

- Analyses in Run 3 could use up to 100 TB when considering systematics when optimizing selections for searches and measurements.
  - There are probably going to be 20-50 analyses per Physics Analysis group (there are ten groups)
- CP groups may even need more space since they could use lower-level information (e.g., calorimeter deposits, tracks, etc).
- Example reconstruction use-case: Graph neural network-based offline track reconstruction consumes O(1Pb) of raw and post-processed data for training
  - Estimate that there are O(10) such hit-level datasets (each for different production type - ttbar, Z', ...) for training

Software

- Common ML tools such as Tensorflow, Pytorch, Lightning, scikit-learn, matplotlib, numpy, etc are necessary
- These tools are already available through LCG views, though typically through a single monolithic “ML” view
- More flexibility might be good, e.g., multiple specific views, or using conda or lmod

Infrastructure

- SWAN and Jupyter notebook work quite well.
- Having machines with shorter run time limitations (e.g., only jobs that will take 1 hour) and that can be accessed interactively would help with testing
- Long run time queues (multiple days) are needed for optimizations and large scale training
    - Checkpointing does solve this

Typical Use Case: Trigger-Track-Tag

- Foresee a convergence to a typical pattern for HL-LHC ATLAS reconstruction in trigger (e.g. Event Filter), tracking (ITk), tagging (FTAG)
- Circa 2023, O(1m) transformer or GNN applied to tagging and tracking/triggering respectively, run on point clouds of size O(1k) and O(100k) respectively
- Training of order O(1 week) for dedicated model
- Inference of order O(1 second) for event reconstruction (tagging or tracking)

Typical Use Case: Trigger-Track-Tag

- Circa 2027, expect the following patterns:
- Scaling up of transformer/GNN size; Observe clear improvement of model performance with model size
- Fixed GPU size
- Therefore, scaling of training parallelism
- In inference, will need to perform model reduction, therefore expect need for hardware that can handle typical optimizations (e.g. custom precisions)
- More and more combination of stages (POC: trackless triggering/tagging), therefore further training memory needed

ML Typical Patterns

- Anecdotally, vast majority of training and development on institutional resources (university/lab/national clusters)
- Some US ATLAS collaboration with Google, TPU usage; this was not pursued aggressively
- General convergence on Pytorch for training (tracking/tagging/triggering)
- General convergence on Onnx for inference (official support within Athena reco framework)
- Most simulation is still classical - not yet obvious how much will move to ML-based

CERN ML Infrastructure Wishlist

1. Well-documented access to GPUs, especially for students at institutes that are not well-resourced - even priority access to low-GPU institutes
2. However, importantly, we don't want to lose regular grid resources in the hunt for ML infrastructure - can *usually* obtain these within institutes
3. Open Data and accessibility of internal data: Lower barrier to entry, for example for a student coming from a different experiment

Conclusion

- In ATLAS, almost all CP & analysis groups are using shallow ML, and many starting to use transformer+GNN models
- Currently, models are relatively small, and hardware appears to be sufficient
- Hardware + tools are most lacking for student onboarding at low-GPU institutes
- Documentation very lacking for CERN-wide / grid GPU training + inference usage
- A common solution to hardware and data sharing *for ML purposes* (i.e. ML-specific data formats?) is clearly needed