

Machine learning at CERN: ATLAS, LHCb, and more

Steven Schramm*, on behalf of the ATLAS and LHCb Collaborations

Université de Genève (CH)

E-mail: steven.schramm@cern.ch

The use of machine learning is increasing at the LHC experiments including both the ATLAS and LHCb collaborations, in terms of the number of users, the breadth of applications, and the set of different techniques under study. While traditionally applied in the context of improving the final analysis selection for a given physics result, machine learning is now also being applied in many other places, including object reconstruction, object calibration, object identification, simulation, and automation. The variety of machine learning tools being used is also expanding, and many areas are benefiting from the use of deep learning methods. It is expected that this growth in machine learning within particle physics will continue, as the large and rapidly increasing datasets provide the perfect environment to develop and refine new machine learning algorithms which can maximally exploit the complex data.

*39th International Conference on High Energy Physics
4-11 July 2018
Seoul, South Korea*

*Speaker.

1. Introduction

The usage of Machine Learning (ML) in particle physics is a rapidly growing field, and both the ATLAS [1] and LHCb [2] collaborations at the LHC are actively investigating the benefits of ML. While ML has historically been used in final analysis event selection to improve the sensitivity to a given process, it is increasingly being used for lower-level tasks. The set of available ML tools is also increasing, including both the traditionally used Boosted Decision Trees (BDTs) and Neural Networks (NNs), as well as new approaches exploiting *deep learning*: Deep Neural Networks (DNNs), Convolutional Neural Networks (CNNs), Recursive Neural Networks (RNNs), and more.

2. Object reconstruction, calibration, and identification

ML is increasingly studied in the context of object reconstruction, either to improve the reconstruction performance or to reduce the amount of time it takes to reconstruct an event. Charged-particle track reconstruction, a complex pattern-matching task, is one example. LHCb has shown that simple NNs can quickly reject fake tracks while retaining a very high signal efficiency [3], and thus uses NNs in the trigger to reduce the combinatoric complexity and allow for the usage of offline precision track reconstruction algorithms in a real-time environment, as shown in Figure 1.

ML is also being used to improve object calibrations. Figure 2 shows that ATLAS has found that Boosted Regression Trees (BRTs) provide significantly improved hadronic tau p_T resolution, especially at low p_T where the resolution is large and is most relevant to typical analyses [4].

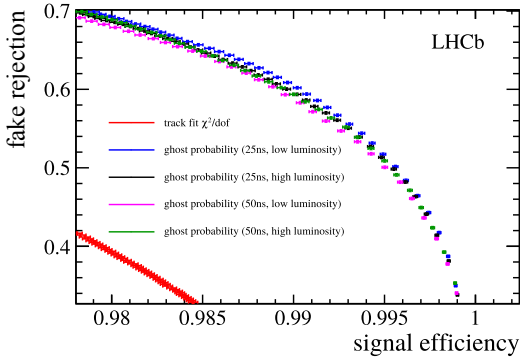


Figure 1: Neural networks efficiently suppress fake tracks in the LHCb trigger, reducing combinatorics compared to the nominal χ^2 fit and allowing for offline track reconstruction algorithms to be used [3].

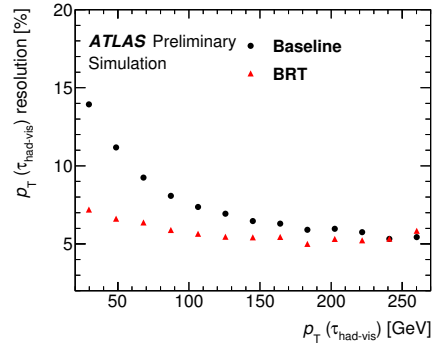


Figure 2: ATLAS usage of a Boosted Regression Tree (BRT) significantly reduces the hadronic tau p_T relative resolution with respect to the baseline approach, especially at low p_T [4].

Object identification significantly benefits from ML, and ATLAS has shown that ML can reduce the background to hadronically decaying top-quarks by a factor of 2 compared to cut-based classifiers [5]. The same result shows that BDTs and DNNs have nearly identical performance in this situation, where the ML classifier uses only *high-level variables* (properties of the object under study). In contrast, Figure 3 shows that b -tagging using RNNs benefits from *low-level variables* (inputs to the object under study) [6]. Using the same variables as before leads to improved performance, while RNNs also facilitate adding more variables, further improving the light-jet rejection.

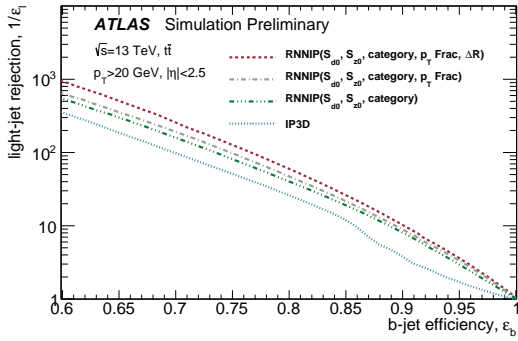


Figure 3: Recurrent Neural Networks (RNNs) exploit correlations among input variables to improve light-jet rejection compared to the ATLAS baseline (blue dots) when using the same variables (green dash-multi-dots) or with more variables (others) [6].

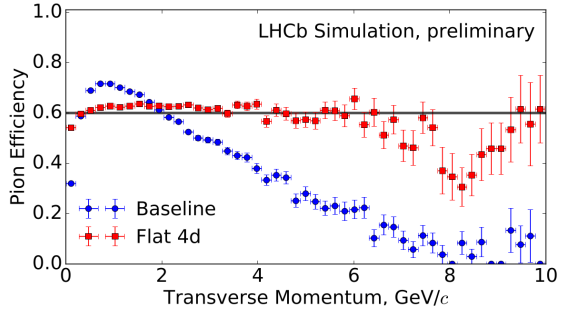


Figure 4: The uniform boosting approach is used in LHCb to simultaneously remove undesired correlations in multiple dimensions, as seen by comparing the nominal (blue) and uniformly boosted (red) distributions for pion identification [7].

In addition to object classification, it is often important to identify objects in a way that reduces the dependence on other variable(s) of interest, thus avoiding shaping relevant distributions. LHCb trains BDTs using uniform boosting to reweight the training events in a way that the correlation is simultaneously removed in multiple dimensions [7]. The resulting classifier in Figure 4 has reduced the p_T dependence of both kaon (not shown) and pion identification efficiencies.

3. Simulation and automation

One rapidly growing application of ML to particle physics is the simulation of showers of particles in calorimeters, where full simulations take a substantial fraction of the total computing resource requirements of modern particle physics experiments. As such, there is substantial interest in the possibility of using ML algorithms to replace this expensive computation with parametrized solutions. Variational Auto-Encoders (VAEs) and Generative Adversarial Networks (GANs) are commonly used, as they can stochastically generate roughly physical distributions once adequately trained. Such techniques are still being improved upon, while first demonstrations are promising.

ML can also be used for anomaly detection of data quality monitoring, quickly and effectively identifying whether a run is of good or bad quality. By automatically handling most cases, experts time can instead be focused on understanding the more difficult to classify runs.

4. Physics analysis applications

ML has historically been applied to physics analysis in order to improve the sensitivity beyond traditional cut-based approaches. The first ATLAS observation of the Higgs boson made use of two NNs [8], which were necessary to push the sensitivity above the threshold of 5σ to claim the observation of a new particle. Similarly, the first observation of pentaquarks by LHCb made use of ML tools [9]. While these are two high profile cases, they are far from the only examples, indeed the number of searches and measurements using ML in the analysis selection continues to grow.

The first ATLAS result for the evidence of $H \rightarrow \tau\tau$ production published both cut- and BDT-based analyses, giving a rare public insight into the benefit of ML. The use of a BDT increased the expected Higgs boson sensitivity from 2.5σ to 3.3σ , and the observed sensitivity from 3.2σ to 4.5σ [10]. BDTs were therefore important to pass the expected sensitivity evidence threshold.

LHCb similarly made use of BDTs to claim the first evidence of the rare $B_s \rightarrow \mu\mu$ decay mode [11]. In addition to such analysis selections, roughly two thirds of physics analyses in LHCb rely on ML at a more fundamental level through the use of bonsai BDT triggers [12]. Bonsai BDTs binarize the nodes for faster evaluation; this comes with a small associated performance loss compared to the full BDT, but performs better than cut-based selections.

5. Summary

Machine learning is an increasingly important aspect of modern particle physics. While originally machine learning was primarily confined to optimizing the final analysis selection, it is now commonly used for much more fundamental tasks. These include object reconstruction, calibration, and identification, as well as simulation and automation. As the particle physics community gathers an increasingly large dataset, the potential benefits of machine learning grow. The ATLAS and LHCb collaborations are actively investigating both new applications and improvements to existing methods, and are poised to maximally exploit their data in the years to come.

References

- [1] ATLAS collaboration, *ATLAS Experiment at the CERN Large Hadron Collider*, *JINST* **3** (2008) 003.
- [2] LHCb collaboration, *The LHCb Detector at the LHC*, *JINST* **3** (2008) 005.
- [3] M. De Cian, S. Farry, P. Seyfert and S. Stahl, *Fast neural-net based fake track rejection in the LHCb reconstruction*, *LHCb-PUB-2017-011* (2017) .
- [4] ATLAS collaboration, *Measurement of the tau lepton reconstruction and identification performance in the ATLAS experiment using pp collisions at $\sqrt{s} = 13$ TeV*, *ATLAS-CONF-2017-029* (2017) .
- [5] ATLAS collaboration, *Performance of Top Quark and W Boson Tagging in Run 2 with ATLAS*, *ATLAS-CONF-2017-064* (2017) .
- [6] ATLAS collaboration, *Identification of Jets Containing b-Hadrons with Recurrent Neural Networks at the ATLAS Experiment*, *ATL-PHYS-PUB-2017-003* (2017) .
- [7] Talk by M. Martinez at the 2nd IML workshop, indico.cern.ch/event/668017/contributions/2947017/.
- [8] ATLAS collaboration, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, *Phys. Lett.* **B716** (2012) 1–29.
- [9] LHCb collaboration, *Observation of $J/\psi p$ Resonances Consistent with Pentaquark States in $\Lambda_b^0 \rightarrow J/\psi K^- p$ Decays*, *Phys. Rev. Lett.* **115** (2015) 072001.
- [10] ATLAS collaboration, *Evidence for the Higgs-boson Yukawa coupling to tau leptons with the ATLAS detector*, *JHEP* **04** (2015) 117.
- [11] LHCb collaboration, *First Evidence for the Decay $B_s^0 \rightarrow \mu^+ \mu^-$* , *Phys. Rev. Lett.* **110** (2013) 021801.
- [12] V. V. Gligorov and M. Williams, *Efficient, reliable and fast high-level triggering using a bonsai boosted decision tree*, *JINST* **8** (2013) P02013.