Supporting Research at the Intersection of Physics and Machine Learning

Savannah Thais (Princeton University)
sthais@princeton.edu

Abstract

Machine Learning (ML) is becoming an integral part of physics research: many critical HEP algorithms for triggering, reconstruction, and analysis rely on ML and there are entire conferences and summer schools dedicated to this intersection. Moreover, the unique constraints of physics experiments and the ability to exploit symmetries inherent in physics data have made the field of physics-informed ML a vibrant sub-field of Computer Science (CS) research. However, despite the relevance and importance of this research, pursuing a career at the intersection of these fields remains tenuous and undefined endeavor. This LoI aims to begin the discussion on how we can better support academic and career development at the intersection of physics and ML and ensure we can continue to benefit from and contribute to state-of-the-art ML.

1 Current Environment

HEP research has long required advanced, cutting-edge computing techniques, and physicists have historically contributed to the development of these methods. Over the past decades, there have been great advances in the data-processing power of ML algorithms and these methods have quickly been adopted by physicists to address the unique timing, memory, and latency constraints of HEP experiments [1].

Physics is also impacting ML research: the constraints of HEP experiments and known symmetries of physical systems create a rich environment for the development of novel ML (see for example [2]). There are even entire conferences and workshops dedicated to this intersection including the Microsoft Physics $\cap$ ML lecture series [3] and the ML and the Physical Sciences workshop at NeurIPS [4].

However, despite the demonstrated criticality and vibrancy of this research, from a physics background the path to a sustainable research career at the intersection of physics and ML is unclear at best. In fact, many researchers seem to regard ML research and applications as minor side projects[1].

There are very few physics department supported courses on ML [2] and interested students are often forced to convince CS departments to allow them into their graduate courses (which typically do not discuss physics-informed ML). Furthermore, at many universities collaborative interactions between Physics and CS departments are very limited, which in turn limits the opportunities for physicists to engage with cutting-edge ML research and develop partnerships with industry researchers. [3]. As a physics graduate student at a US institute it is essentially impossible to focus one's thesis research on ML as most departments do not acknowledge this as 'physics work' (despite the fact that these algorithms are in many cases critical to the functioning of physics experiments and data processing). Perhaps most egregiously, although there is now some support for this work at the post-doc or research staff level (see IRIS-HEP[6] or the new IAIFI[7]), early-career researchers are often actively advised against applying to these positions as they constitute 'career death sentences'.

2 Proposals for Future Work

It is clear that these issues must be addressed in order to best support and advance our field, research, and researchers. An important first step is to work towards a cultural shift around the perception of what it means to 'be a physicist' and what work is considered as contributions to physics. We must value technical and software contributions in the same manner we value analysis work and foster a supportive and educational environment for graduate students and early-career researchers interested in ML and physics. It benefits no-one to discourage individuals from pursuing work in this vibrant and promising field.

Concretely, this necessitates developing cross-department courses, workshops, and research collaborations. Graduate programs should approach technical and computing skills with the same rigor they do traditional physics courses. Departments and advisors should be encouraged to allow students and post-docs to work across disciplines and explore opportunities for industrial partnerships or internships. This process should also include a reevaluation of requirements for physics graduate theses.

The research landscapes of both physics and ML are shifting, and by adjusting to these changes we can ensure a vibrant future for both data-intensive physics and physics-informed ML.

[1] Due to the lack of community statistics on this topic, the evidence presented here is primarily anecdotal

[2] although [5] is a good example of how such a course might work

[3] NYU Physics collaborations with the Center for Data Science is an excellent example of cross-department collaboration

References

[1] Dan Guest, Kyle Cranmer, and Daniel Whiteson. "Deep Learning and Its Application to LHC Physics". In: *Annual Review of Nuclear and Particle Science* 68.1 (Oct. 2018), pp. 161–181. ISSN: 1545-4134. DOI: 10.1146/annurev-nucl-101917-021019. URL: http://dx.doi.org/10.1146/annurev-nucl-101917-021019.

[2] Miles Cranmer et al. *Lagrangian Neural Networks.* 2020. arXiv: 2003.04630 [cs.LG].

[3] Microsoft. *Physics Meets ML Lecture Series.* URL: http://physicsmeetsml.org/.

[4] et al Baydin Altim Gunes. *Machine Learning and the Physical Sciences Workshop.* URL: https://ml4physicalsciences.github.io/2020/.

[5] Pankaj Mehta. *Machine Learning for Physicists.* URL: http://physics.bu.edu/~pankajm/PY895-ML.html.

[6] *Institute for Research and Innovation in Software for High-Energy Physics.* URL: https://iris-hep.org/.

[7] .