Contribution ID: **36**   Type: **Oral**

hls4ml: deploying deep learning on FPGAs for L1 trigger and Data Acquisition

Thursday, 9 November 2023 16:20 (15 minutes)

Machine learning is becoming ubiquitous across HEP. There is great potential to improve trigger and DAQ performance, and potentially in other real-time controls applications. However, the exploration of such techniques within the field in low latency/power FPGAs has just begun. We present hls4ml, a user-friendly software, based on High-Level Synthesis (HLS), designed to deploy network architectures on FPGAs. As a case study, we use hls4ml for boosted-jet tagging with deep networks at the LHC. We map out resource usage and latency versus network architectures, to identify the typical problem complexity that hls4ml could deal with. We discuss current applications in HEP experiments and future applications. We also report on recent progress in the past year on newer neural network architectures such as binary and ternary networks, large convolutional neural networks, support for QONNX, graph neural networks and transformer neural network types.

Early Career

Yes

Primary author: LIU, Mia

Presenter: LIU, Mia

Session Classification: RDC5

Track Classification: RDC Parallel Sessions: RDC5: Trigger and DAQ