Contribution ID: **176**   Type: **Oral presentation**

Machine Learning for Sub-Microsecond Edge Data Processing

Tuesday 23 April 2024 14:10 (20 minutes)

The improvements in artificial intelligence, particularly the many flavours of Machine Learning (ML), add a powerful and versatile tool to data acquisition (DAQ) strategies. However, large and deep neural networks remain memory and compute intensive, limiting their usability at the edge. One of the most important aspect of integrating ML in a DAQ system is determining when and where integrating a machine learning algorithms will be most beneficial and how to minimize the model size without losing the precision and accuracy required for a scientific application.

We designed and tested a DAQ integrating ML for the CookieBox, a detector capable of sampling an x-ray shot from LCLS-II and determining its suitability for an experiment. By moving the analysis to the edge, this detector can now be used to veto x-ray shots in real-time and reduce the overall data collected for an experiment. The analysis is completed in 0.4 us with an accuracy of 84~% using a few watts of power.

In this talk, we will share insight on the choices made to minimize the ML footprint, implement the data preprocessing and optimize the latency. In addition, we will present the validation methods we used and how our edge ML compares to offline methods.

Minioral

Yes

IEEE Member

Yes

Are you a student?

No

Author: Prof. CORBEIL THERRIEN, Audrey (Université de Sherbrooke)

Co-authors: Mr GOUIN-FERLAND, Berthié (Université de Sherbrooke); Mr RAHIMIFAR, Mohammad Mehdi (Université de Sherbrooke); Mr WINGERING, Quentin (Université de Sherbrooke); Dr COFFEE, Ryan (SLAC National Accelerator Laboratory)

Presenter: Prof. CORBEIL THERRIEN, Audrey (Université de Sherbrooke)

Session Classification: Oral presentations

Track Classification: AI, Machine Learning, Real Time Simulation, Intelligent Signal Processing