

AI4EIC — Live Notes — 08/23/2022

AI/ML Sector of the EICUG Software Working Group



<https://eic.ai> <https://indico.bnl.gov/event/16605/>

Special Meeting: AI/ML in streaming readout • August 24, 9-11 a.m. ET / 3-5 p.m. CET



Markus Diefenthaler
Jefferson Lab



Vincenzo Lomonaco
U. of Pisa

AI/ML in streaming readout Continual Learning

Speaker: Markus Diefenthaler (Jefferson Lab)

Notes

Questions/comments

- David Lawrence:
 - Comment: INDRA/ASTRA adwin; add AI for experimental control in SRO, corrections to the data with calibration data, along with correct data before we take it — complement INDRA/ASTRA (features / metadata from slow-control)
 - Comment: Work with AI generally you need lots of data to do with it; varied data; what kind of commissioning you need; variation ; historic data; if you do not have much variation hard to learn — commissioning need to be considered
 - Follow-up: MD: agree; this should be extended also to accelerator (near real-time — feedback loops; info on controls' experimentation)
 - Tanja: sensitivity to environment, adding: no smooth behavior; unexpected behavior;
- Vincenzo: what are the main bottlenecks in this vision?

- MD: largely inspired by LHCb; they have physics-driven(“candles”)/informed ML-approaches — a bit of a challenge for EIC
 - MD: there is no need to do it near real-time; but it would be desirable; what kind of biases are we introducing? How do we deal with biases?
 - MD: finding/establishing a common knowledge — community challenge
 - MD: understanding UQ

Speaker: Vincenzo Lomonaco (U. of Pisa)

Notes

Questions/comments

- Wouter: Prototyping → how to get it in a production environment? Open source approaches in that direction?
 - VL: application on smart device, DL toolkits do not support learning on device — strong and general efforts / improve support ; training on such devices, production-ready environment is a real challenge for many ML applications — focus for many companies is inference; training is a different game;
- Thomas Britton: Lot of it focuses on classification; any interest in doing regression (not strictly classification)?
 - VL: classification is perhaps for simplicity since we have benchmark
- CFanelli: CLAIR Consortium
 - VL:
- CFanelli: stability and plasticity in distributed environment
 - VL:
 - WD: Federated learning is likely something we'll need to explore for training models for identification or regression in our production chains. The model of just running large campaigns and analyzing them is not automated or responsive enough.
- TB: