

Deployment of ML in changing environments

Tuesday, 4 October 2022 16:15 (15 minutes)

The deployment of fast ML models for on-detector inference is rapidly growing but faces key issues. One such issue is the difference between the training environment and the “real-world” environment in deployment giving unknown errors in inference. Examples of this include training a model on an abundance of well understood simulated data but deploying it on a real and imperfect detector or on a detector in which the performance changes over time that the ML model is unaware of. Various techniques can be employed to mitigate this issue including the use of uncertainty quantification to better understand the inference errors, retraining and redeploying models with new data or the use of continual learning where a model is continually updated with a stream of evolving data. This issue of deploying ML models in changing environments is presented as are the pros and cons of potential solutions.

Primary author: RADBURN-SMITH, Benjamin (Imperial College (GB))

Co-authors: TAPPER, Alex (Imperial College London); BROWN, Christopher Edward (Imperial College (GB)); BARBONE, Marco; KOMM, Matthias (Deutsches Elektronen-Synchrotron (DE)); SUMMERS, Sioni Paris (CERN)

Presenters: RADBURN-SMITH, Benjamin (Imperial College (GB)); BROWN, Christopher Edward (Imperial College (GB)); BARBONE, Marco

Session Classification: Contributed Talks