Contribution ID: **23** Type: **Zoom**

Adversarial training for b-tagging algorithms in CMS

Thursday, 3 November 2022 09:40 (20 minutes)

Deep learning is a standard tool in high-energy physics, facilitating identification of physics objects. In particular, complex neural network architectures play a major role for jet flavor tagging. However, these methods are reliant on accurate simulations and a calibration is required to treat non-negligible performance differences with respect to data. In order to reduce residual disagreement between these two domains, adversarial methods are applied to close the generalization gap and to improve a classifier's robustness. Extensive studies have been carried out on a publicly accessible dataset with a fully connected neural network. Studying the impact of adversarial attacks on the inputs mimics the effect of systematic uncertainties. To mitigate this effect, an enhanced algorithm that adapts adversarial training is presented. We show that this strategy can also be applied to the DeepJet algorithm, a commonly used tagger at the CMS experiment. Due to the large number of inputs, small differences can add up to a considerable impact on performance. Utilizing the interplay of frameworks for sample creation, training, evaluation and scale factor derivation, we show that this mitigation strategy successfully improves agreement between data and simulated samples while maintaining a high performance. Thus, the introduction of the adversarial module is envisaged to become a useful ingredient for the upcoming generation of flavor tagging algorithms developed for Run3.

Primary authors: STEIN, Annika (Rheinisch Westfaelische Tech. Hoch. (DE)); COLLABORATION, CMS (CMS Experiment, CERN)

Presenters: STEIN, Annika (Rheinisch Westfaelische Tech. Hoch. (DE)); COLLABORATION, CMS (CMS Experiment, CERN)

Session Classification: Classification