ID de Contribution: **89** Type: **Non spécifié**

A new method for the data driven estimation of background using GAN

mercredi 28 septembre 2022 14:20 (20 minutes)

In this contribution we present a novel data driven method for the estimation of background by generating a new misidentified object using generative adversarial networks (GAN). In High Energy Physics, characterizing signal hypothesis requires distinguishing its signature from a large number of background processes with similar final states. Machine learning (ML) classification algorithms are widely used to obtain the optimal separation between signal and background processes. However, constraints on the storage and computing power restrict the generated sample size, which may cause Monte Carlo (MC) simulation of events with large population weights. The training of ML algorithms with these limited size samples may degrade the classification performance.

Many of these background events come from processes similar to signal process, but with one (or more) misidentified object (such as γ + jets background process for the H $\rightarrow \gamma\gamma$ analysis). Therefore, a good description of these background processes is important for the sensitivity of the analyses. In this presentation, a novel technique using a GAN based object simulation is presented. A conditional GAN algorithm is used to simulate a new misidentified object for the data events that cannot pass the object identification criteria. Hence, this large sample of filtered out data events can be used in the description of the processes with misidentified objects. We demonstrate that using conditional GAN algorithm, the observables of the generated object retains the correlations with other features.

Auteurs principaux: COUDERC, Fabrice (CEA); MALCLES, Julie (IRFU, CEA-Saclay); SAHIN, Mehmet Ozgur (CEA-Saclay); LOHEZIC, Victor (IRFU (CEA) / Université Paris-Saclay)

Orateur: LOHEZIC, Victor (IRFU (CEA) / Université Paris-Saclay)

Classification de Session: Wed afternoon