Contribution ID: **661** Contribution code: **contribution ID 661** Type: **Poster**

Generative models uncertainty estimation

In recent years fully-parametric fast simulation methods based on generative models have been proposed for a variety of high-energy physics detectors. By their nature, the quality of data-driven models degrades in the regions of the phase space where the data are sparse. Since machine-learning models are hard to analyze from the physical principles, the commonly used testing procedures are performed in a data-driven way and can't be reliably used in such regions. In our talk we propose three methods to estimate the uncertainty of generative models inside and outside of the training phase space region, along with data-driven calibration techniques. Test of the proposed methods on the LHCb RICH fast simulation is also presented.

Significance

Scientific applications of machine learning are complicated by the rigor required by the scientific method. It is not enough just to have a black-box model that outputs some answers. We need to know the uncertainty of those predictions. This problem is well-studied for classification and regression and not so for generative models. The latter are becoming a popular choice for fast simulation in HEP, at least from ACAT 2017. In our work, we present novel algorithms for estimating the uncertainty of generative models. The need for them is particularly acute in precision analyses, such as the CKM angles or Lepton Flavour Universality tests. Understanding of systematic uncertainty there is an absolute necessity and is the only way to completely incorporate fast simulation techniques into those analyses to solve the computing requirements problems.

References

Our work builds on those two papers, which cover non-generative models:
https://papers.nips.cc/paper/2017/file/9ef2ed4b7fd2c810847ffa5fa85bce38-Paper.pdf
http://proceedings.mlr.press/v48/gal16.html

The LHCb RICH fast simulation has been presented at ACAT-2019:
https://iopscience.iop.org/article/10.1088/1742-6596/1525/1/012097/meta

Speaker time zone

Compatible with Europe

Primary authors: Ms SHISHIGINA, Agata (HSE University); Mr CHIMPOESH, Constantine (HSE University); ANDERLINI, Lucio (Universita e INFN, Firenze (IT)); Dr KAZEEV, Nikita (Yandex School of Data Analysis (RU))

Presenter: Dr KAZEEV, Nikita (Yandex School of Data Analysis (RU))

Session Classification: Posters: Orange

Track Classification: Track 2: Data Analysis - Algorithms and Tools