

Blueprints for Training Information Bottlenecks for Collider Analyses

Friday 4 November 2022 11:30 (20 minutes)

Dimensionality reduction is a crucial aspect of data analysis in high energy physics, even if accompanied by information loss. Several methods, including histogram- and kernel-based analyses, are only computationally feasible for low-dimensional data. Furthermore, simulation models used in HEP can often only be validated for low-dimensional data. We provide several blueprints for using machine learning to create low-dimensional data representations (continuous event variables and discrete classification labels) for use in signal discovery and parameter estimation tasks. We also describe how to design the learned representation to facilitate a) searches with unknown model parameters and b) validation of simulation models in data control regions.

Primary authors: PEDRO, Kevin (Fermi National Accelerator Lab. (US)); SHYAMSUNDAR, Prasanth (Fermi National Accelerator Laboratory)

Presenter: SHYAMSUNDAR, Prasanth (Fermi National Accelerator Laboratory)

Session Classification: Interpretability