Contribution ID: **47**   Type: **not specified**

HEP ML Lab —An end-to-end framework for signal vs background analysis in high energy physics

Thursday 9 November 2023 17:35 (1 minute)

We have developed an end-to-end data analysis framework, HEP ML Lab (HML), based on Python for signal-background analysis in high-energy physics research. It offers essential interfaces and shortcuts for event generation, dataset creation, and method application.

With the HML API, a large volume of collision events can be generated in sequence under different settings. The representations module enables easy conversion of event data into input formats required by various methodologies. The API also includes three categories of analysis methods: cut-based analysis, multivariate analysis, and neural networks, to cater to diverse needs. Coupled with built-in metric parameters, users can preliminarily assess the performance of different analytical methods while using them.

While the high-energy physics research community has already explored several frameworks that integrate data and analysis methods, we advocate for integrating the entire end-to-end process into a single framework. By offering a unified style of programming interface, it reduces the need for researchers to switch between different software and frameworks. This not only simplifies and clarifies the research process, but also facilitates the reproduction of previous research results, leading to more persuasive conclusions.

To demonstrate the convenience and effectiveness of HML, we provide a case study that differentiates between Z jets and QCD jets. We provide benchmark testing for the three built-in methods and ultimately export shareable datasets and model checkpoints.

Author: LI, Jing (Dalian University of Technology, Liaoning, China)

Co-author: Prof. SUN, Hao (Dalian University of Technology, Liaoning, China)

Presenter: LI, Jing (Dalian University of Technology, Liaoning, China)

Session Classification: Remote Discussion