



Contribution ID: 59

Type: **Lightning Talk**

Accommodating Transformer in the FastML era of HEP-EX

Monday 25 September 2023 18:20 (5 minutes)

Recent years have witnessed the enormous success of the transformer models in various research fields including Natural Language Processing, Computational Vision as well as natural science territory. In the HEP community, models with transformer backbones have shown their power in jet tagging tasks. However, despite the impressive performance, transformer-based models are often large and computationally heavy, resulting in low inference speeds. In this talk, I will discuss the preliminary results of our effort in accelerating the transformer models in the context of FastML.

Primary author: QIAN, Sitian (Peking University (CN))

Presenter: QIAN, Sitian (Peking University (CN))

Session Classification: Contributed Talks

Track Classification: Contributed Talks