

A multi-task Large Language Model for jets

lunedì 29 luglio 2024 17:40 (20 minuti)

Multi-head attention based Transformers have taken the world by storm, given their outstanding capacity of learning accurate representations of diverse types of data. Famous examples include Large Language Models, such as ChatGPT, and Vision Transformers, like BEiT, for image generation. In this talk, we take these major technological advancements to the realm of jet physics. By creating a discrete version of jet constituents, we let an Auto-regressive Transformer network learn the 'language' of jet substructures. We demonstrate that our Transformer model learns highly accurate representations of different types of jets, including precise predictions of their multiplicity, while providing explicit density estimation. Moreover, we show that the Transformer model can be used for a variety of tasks that involve both jet tagging and generation. Finally, we discuss how a pre-trained Transformer can be used as a baseline for fine-tuned models created for specific tasks for which data may be scarce.

Autori principali: MÜCK, Alexander (RWTH Aachen); REYES-GONZALEZ, Humberto (RWTH Aachen); KRÄMER, Michael (RWTH Aachen)

Relatore: REYES-GONZALEZ, Humberto (RWTH Aachen)

Classifica Sessioni: Novel Techniques

Classificazione della track: Novel Techniques