Contribution ID: **26**   Type: **not specified**

Learning Symmetry-Independent Jet Representation via Jet-Based Joint Embedding Predictive Architecture

Tuesday 5 November 2024 14:30 (20 minutes)

This study introduces an approach to learning augmentation-independent jet representations using a Jet-based Joint Embedding Predictive Architecture (J-JEPA). This approach aims to predict various physical targets from an informative context, using target positions as joint information. We study several methods for defining the targets and context, including grouping subjets within a jet, and grouping jets within a full collision event. As an augmentation-free method, J-JEPA avoids introducing biases that could harm downstream tasks, which often require invariance under augmentations different from those used in pretraining. This augmentation-independent training enables versatile applications, offering a pathway toward a cross-task foundation model. J-JEPA has the potential to excel in various jet-based tasks such as jet classification, energy calibration, and anomaly detection. Moreover, as a self-supervised learning algorithm, J-JEPA pretraining does not require labeled datasets, which can be crucial with the impending dramatic increase in computational cost for HL-LHC simulation. The reduced dependency of J-JEPA on extensive labeled data allows learning physically rich representations from unlabeled data and fine-tuning the downstream models with only a small set of labeled samples. In a nutshell, J-JEPA provides a less biased, cost-effective, and efficient solution for learning jet representations.

Track

Tagging (Classification)

Co-authors: MOKHTAR, Farouk (Univ. of California San Diego (US)); LI, Haoyang (Univ. of California San Diego (US)); DUARTE, Javier Mauricio (Univ. of California San Diego (US)); Mr KATEL, Subash (UCSD); ZHAO, Zihan (Univ. of California San Diego (US))

Presenter: LI, Haoyang (Univ. of California San Diego (US))

Session Classification: Tagging