Contribution ID: **119**   Type: **Poster (from contributed talk)**

Masked particle modelling - Poster

Thursday 1 February 2024 17:00 (5 minutes)

The Bert pretraining paradigm has proven to be highly effective in many domains including natural language processing, image processing and biology. To apply the Bert paradigm the data needs to be described as a set of tokens, and each token needs to be labelled. To date the Bert paradigm has not been explored in the context of HEP. The samples that form the data used in HEP can be described as a set of particles (tokens) where each particle is represented as a continuous vector. We explore different approaches for discretising/labelling particles such that the Bert pretraining can be performed and demonstrate the utility of the resulting pretrained models on common downstream HEP tasks.

Would you like to be considered for an oral presentation?

Authors: RAINE, Johnny (Universite de Geneve (CH));   HEINRICH, Lukas Alexander (Technische Universitat Munchen (DE));   Mr LEIGH, Matthew (University of Geneva);   KAGAN, Michael (SLAC National Accelerator Laboratory (US));   Prof. OSADCHY, Rita (University of Haifa);   KLEIN, Samuel Byrne (Universite de Geneve (CH));   GOLLING, Tobias (Universite de Geneve (CH))

Presenter: KLEIN, Samuel Byrne (Universite de Geneve (CH))

Session Classification: Poster Session

Track Classification: 1 ML for object identification and reconstruction