Masked particle modelling

Wednesday, 1 November 2023 09:40 (10 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.

Brainstorming idea [title]

Optimising search strategies

Brainstorming idea [abstract]

Given a fixed search strategy for new physics that is RECASTABLE, active learning allows us to set exclusion boundaries in the full phase space of a given signal model. How could we automate follow up analyses with different selections etc that might contain new physics? eg) what data should we look at? How can the search strategy be tweaked? ...

Primary 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)); KLEIN, Samuel Byrne (Universite de Geneve (CH)); GOLLING, Tobias (Universite de Geneve (CH))

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

Session Classification: Young Scientist Forum