

# BART-L: Towards a Foundation Model for Theoretical High-Energy Physics

*Monday 8 December 2025 15:30 (30 minutes)*

Transformers excel at natural language processing, raising the question of whether they can also learn the mathematical structures underlying particle physics. We introduce BART-L, a transformer-based model trained to generate particle physics Lagrangians from field content and symmetry information. Trained on Lagrangians consistent with the Standard Model gauge group  $SU(3) \times SU(2) \times U(1)$ , BART-L achieves over 90% accuracy for expressions involving up to six matter fields. Embedding analyses indicate that the model internalizes concepts such as group representations and conjugation operations, despite not being explicitly trained for them. We further study its out-of-distribution performance to identify architectural factors limiting generalization. This framework provides an early indication of what a foundation model for theoretical high-energy physics might look like, along with its potential capabilities and inherent limitations.

**Authors:** CAMARGO-MOLINA, Eliel (Uppsala University); KOAY, Yong Sheng

**Co-authors:** ENBERG, Rikard; MORETTI, Stefano (Science and Technology Facilities Council STFC (GB))

**Presenter:** KOAY, Yong Sheng