Contribution ID: **55**   Type: **Talk**

Towards an Astronomical Foundation Model for Stars with a Transformer-based Model

Tuesday, November 28, 2023 5:01 PM (15 minutes)

Rapid strides are currently being made in the field of artificial intelligence using Transformer-based models like Large Language Models (LLMs). The potential of these methods for creating a single, large, versatile model in astronomy has not yet been explored except for some uses of the basic component of Transformer – the attention mechanism. In this talk, we will talk about a framework for data-driven astronomy that uses the same core techniques and architecture as used by LLMs without involving natural language but floating point data directly. Using a variety of observations and labels of stars as an example, we have built a Transformer-based model and trained it in a self-supervised manner with cross-survey data sets to perform a variety of inference tasks. In particular, we have demonstrated that a single model can perform both discriminative and generative tasks even if the model was not trained or fine-tuned to do any specific task. For example, on the discriminative task of deriving stellar parameters from Gaia XP spectra, our model slightly outperforms an expertly trained XGBoost model in the same setting of inputs and outputs combination. But the same model can also generate Gaia XP spectra from stellar parameters, inpaint unobserved spectral regions, extract empirical stellar loci, and even determine the interstellar extinction curve. The framework allows us to train such foundation models on large cross-survey, multidomain astronomical data sets with a huge amount of missing data due to the different footprints of the surveys. This demonstrates that building and training a single foundation model without fine-tuning using data and parameters from multiple surveys to predict unmeasured observations and parameters is well within reach. Such 'Large Astronomy Models' trained on large quantities of observational data will play a large role in the analysis of current and future large surveys.

Primary authors: LEUNG, Henry (University of Toronto);   Prof. BOVY, Jo (University of Toronto)

Presenter: LEUNG, Henry (University of Toronto)

Session Classification: Contributed talks

Track Classification: New York