



MACHINE LEARNING FOR ASTROPHYSICS

2ND EDITION CATANIA, 8-12 JULY, 2024

Contribution ID: 235

Type: **not specified**

The TransformerPayne: Toward a Foundation Model for Stellar Spectra

Wednesday, 10 July 2024 10:20 (20 minutes)

Current and upcoming large-scale spectroscopic surveys are delivering an ever-increasing volume of high-quality stellar spectra. This wealth of data presents both opportunities and challenges in inferring the parameters of stellar atmospheres. Inference initially relied on polynomial interpolation across a regular spectral grid. This approach was later largely replaced by the use of emulators based on neural networks, which offer efficiency and flexibility. However, these often require extensive grids of stellar spectra and tend to reach a plateau in accuracy. To address this challenge, we introduce TransformerPayne: a wavelength-wise neural network-based emulator designed for the accurate and scalable emulation of stellar spectra. TransformerPayne significantly outperforms its predecessors, by achieving a remarkable reduction in mean absolute error, down to 0.001, when emulating normalized flux. This improvement significantly enhances the precision in inferring parameters of stellar atmospheres, as evidenced by refined Cramér-Rao bounds. Our findings demonstrate that TransformerPayne excels particularly in scenarios with spectral grids ranging from 1,000 to 10,000 spectra, achieving a 2 to 10 times improvement in emulation accuracy with respect to the Payne, multilayer perceptron-based model, measured by mean absolute and mean squared errors. Furthermore, by employing a two-stage training strategy - initially pre-training a base model on simplified physics models, then fine-tuning on more realistic spectra - we achieve a tenfold reduction in the sizes of spectral grids while maintaining emulator quality. Our comprehensive experiments not only establish a new benchmark in the emulation of stellar spectra but also highlight TransformerPayne's favorable scaling properties. Summarizing, these improvements facilitate more precise measurements of effective temperatures, surface gravities, and individual abundances, which are of great interest in astrophysics, while using much smaller spectral grids which can then rely on much more accurate and complex numerical models. Finally, the promising scaling properties open a prospect of a foundation model for stellar spectra crafted for astronomical research that can lead to transformative results similar to those in the domains of audio, images, and language.

Presenter: RÓŻAŃSKI, Tomasz

Session Classification: Time Domain / Multi-wavelength All-Sky Surveys