

# Towards Foundation Model for Stars with Transformer

Henry Leung

Leung & Bovy 2023 (arXiv:2308.10944)



UNIVERSITY OF  
TORONTO



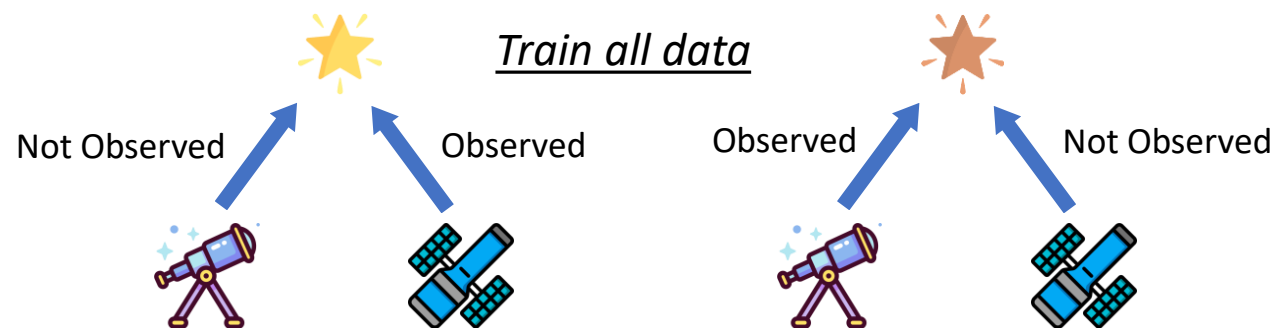
**DATA SCIENCES  
INSTITUTE**

# Motivation

- Foundation model for stars
  - Fine-tuning for downstream tasks
  - Trained with most available datasets
- Flexible model with Transformers

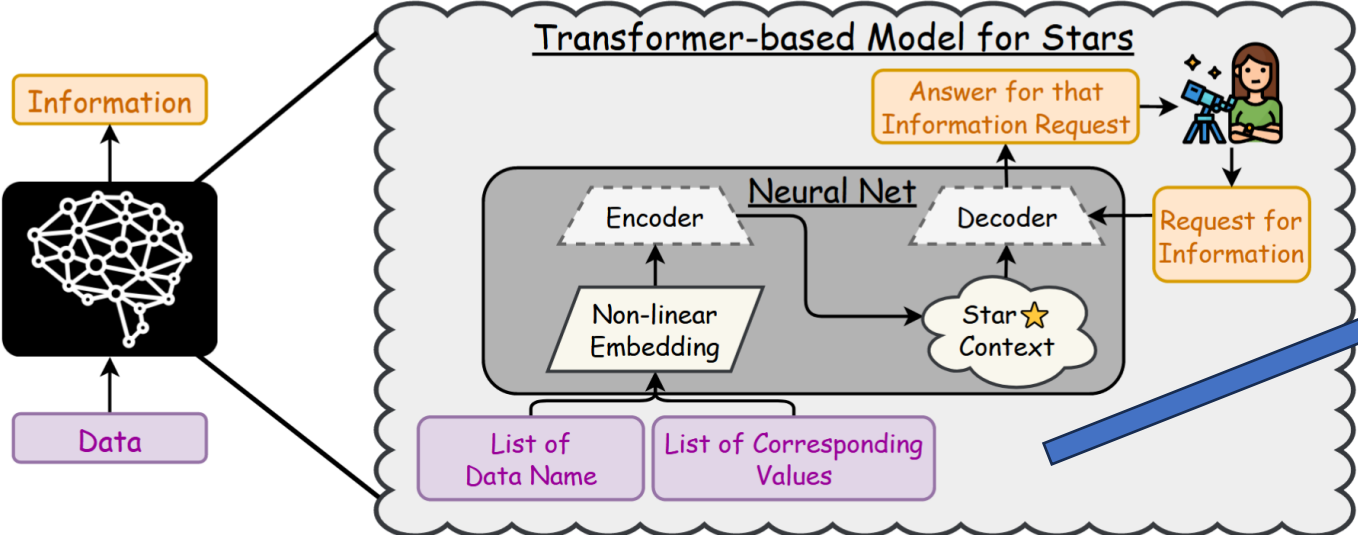


- Trained with heterogenous dataset at scale



# Model and Data

- A small proof-of-concept model
  - Based on Transformer encoder-decoder architecture (Vaswani et al. 2017)
  - Mostly Gaia XP spectra, APOGEE stellar parameters, 2MASS photometry



Embedding

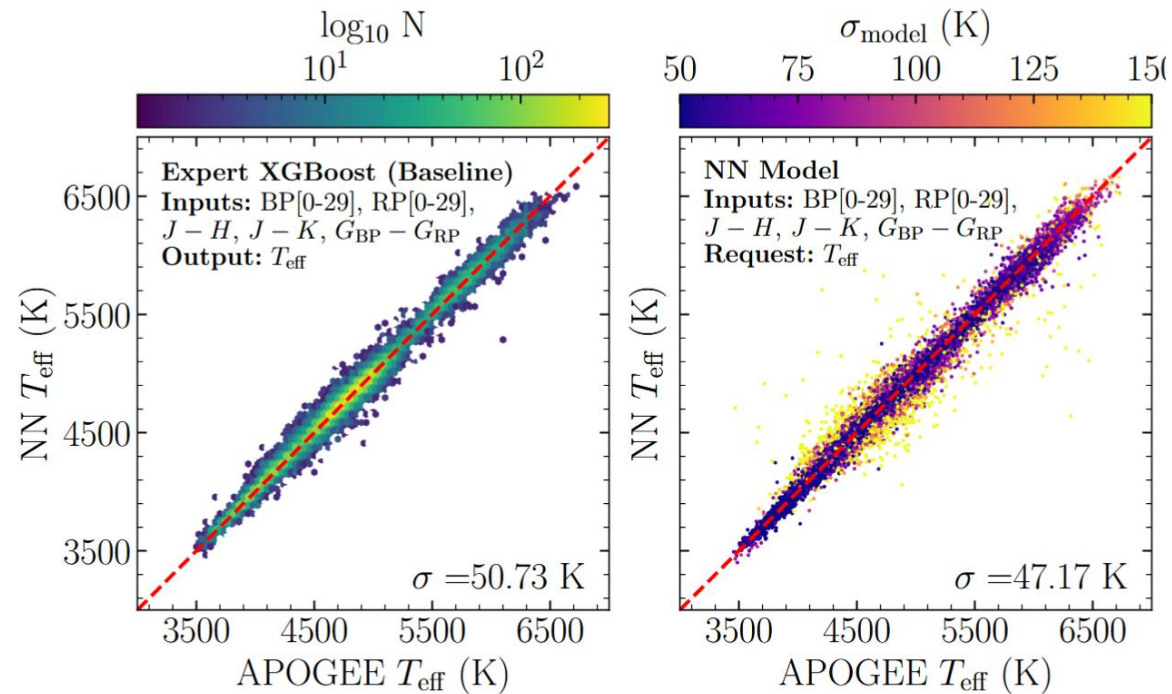
- Data is embedded by their name along with their corresponding value

$$y_x = f(w_x \cdot M_x) + w_{b,x}$$

Final Embedding      Observation type Embedding (e.g., teff)      Magnitude

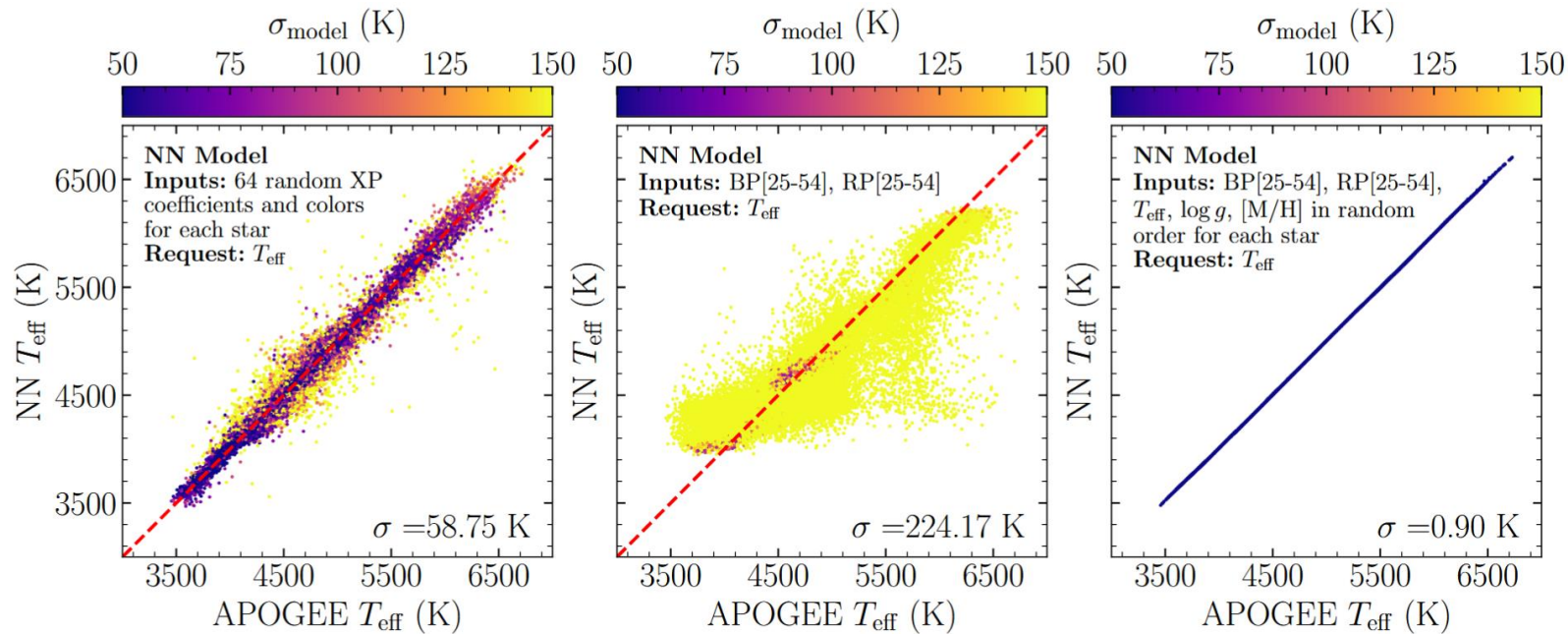
# Example: Discriminative Task

- Gaia XP spectra to Surface Temperature
  - Outperformed expert XGBoost (Zhang et al. 2023, Andrae et al. 2023)



# Example: Discriminative Task

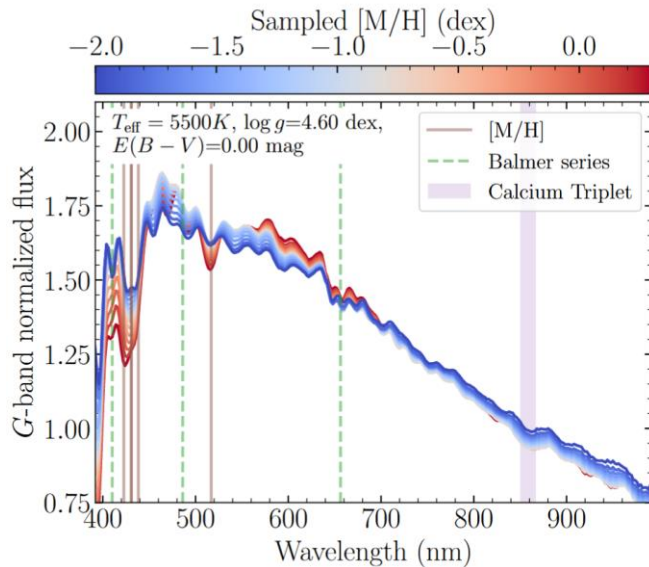
- Gaia XP spectra to Surface Temperature
  - Robust to random permutation in input
  - Reasonably uncertainty



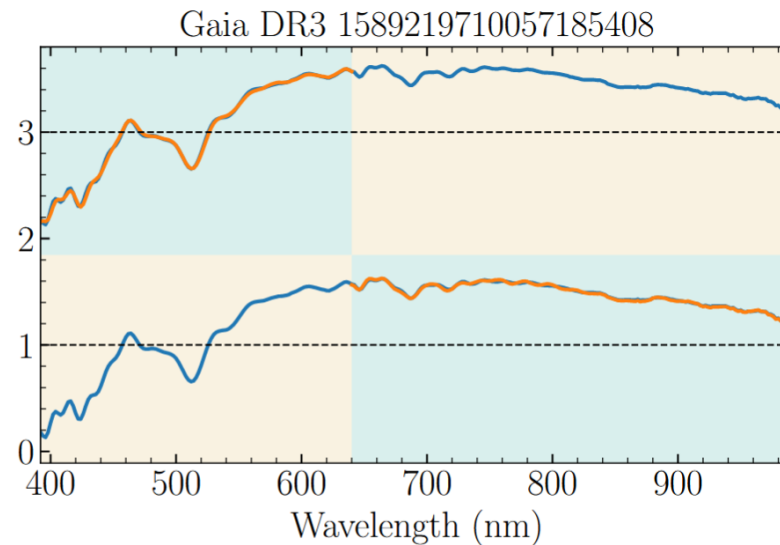
# Example: Generative Task

- Generative tasks with Gaia XP spectra

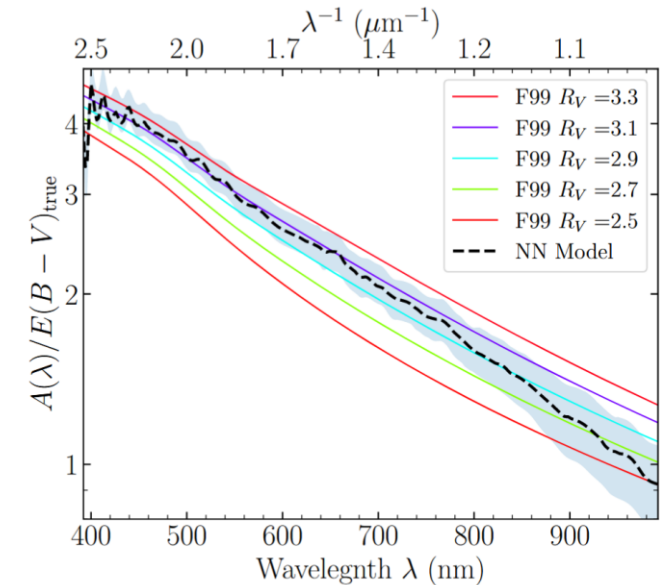
Stellar Parameters to  
Gaia XP Spectra



Portion of Gaia XP Spectra  
to another portion



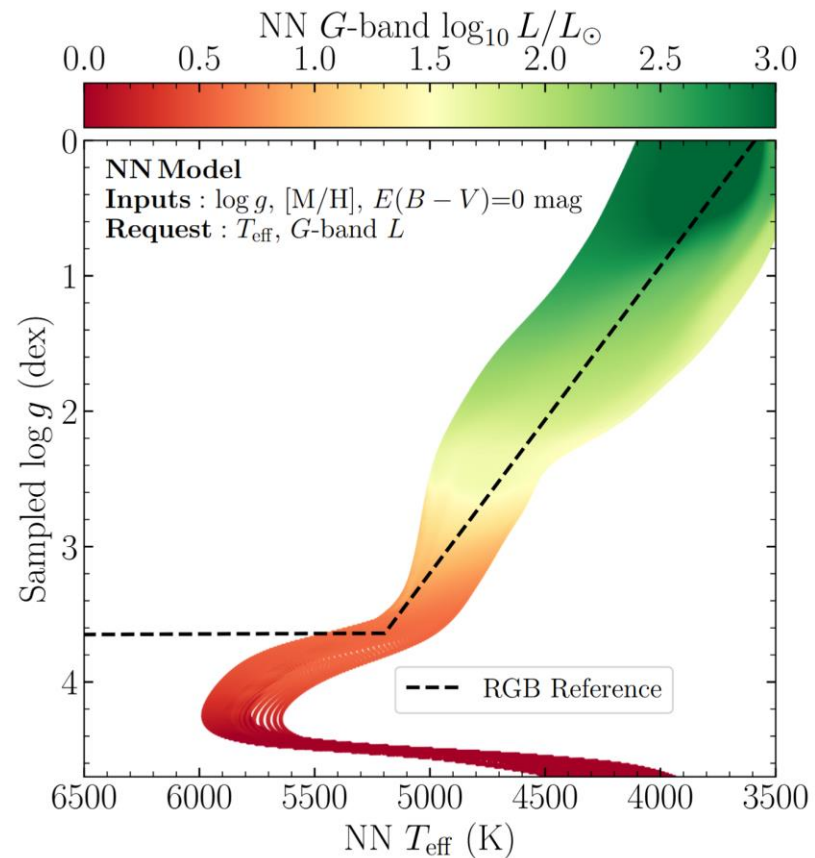
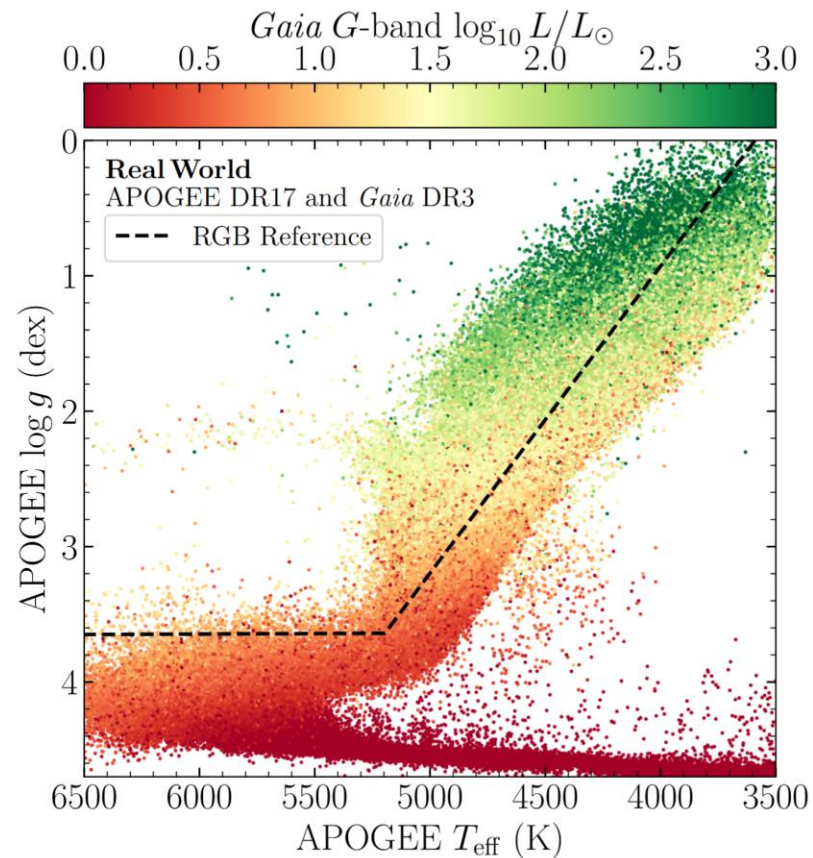
Infer interstellar extinction curve from  
generated Gaia XP Spectra





# Example: Generative Task

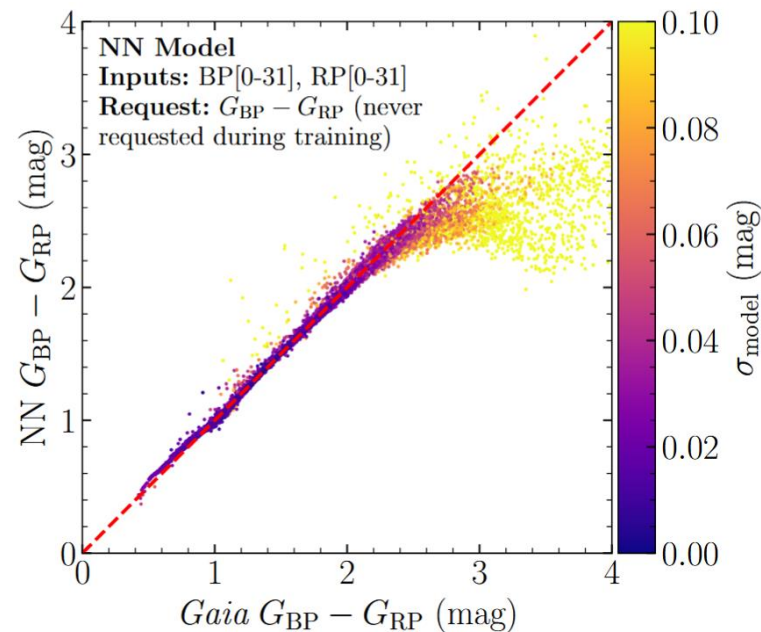
- Generating “isochrones” by sampling  $\log(g)$  and metallicity



# Embedding

- Gaia BP-RP color embedding only trained with the encoder, but never presented to decoder during training

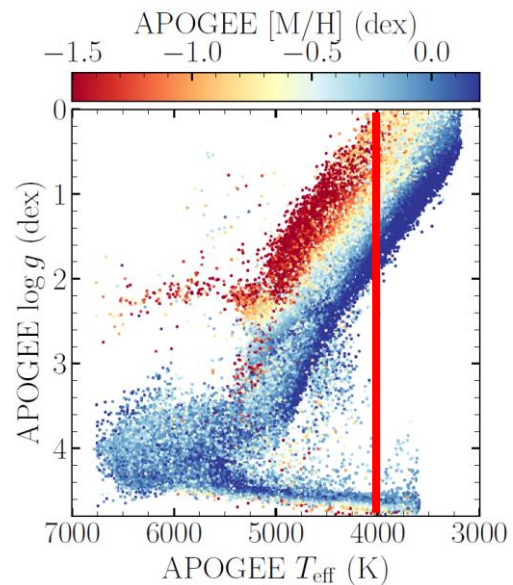
Gaia BP-RP color is computed from Gaia XP spectra which both encoder and decoder have seen during training





# Reliability issue

- Ambiguous questions
- Given a star with surface temperature 4000K, what should be its surface gravity?

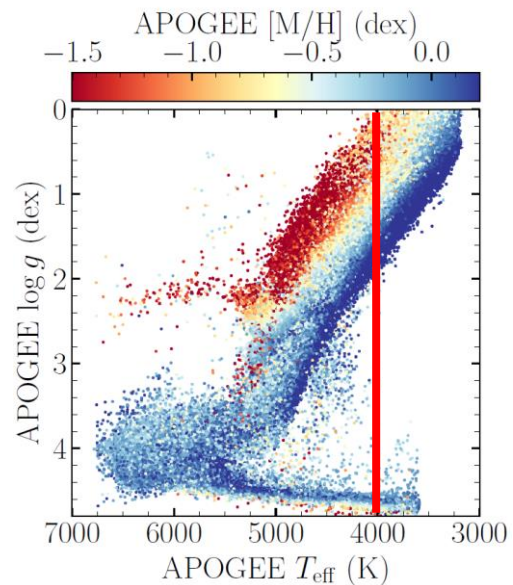


$\log(g)=3.0$  dex but  
uncertainty on  
 $\log(g)=2.0$  dex

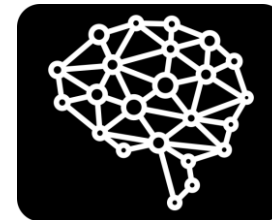


# Reliability issue

- Ambiguous questions
- Given a star with surface temperature 4000K with no extinction, what should be its surface gravity?

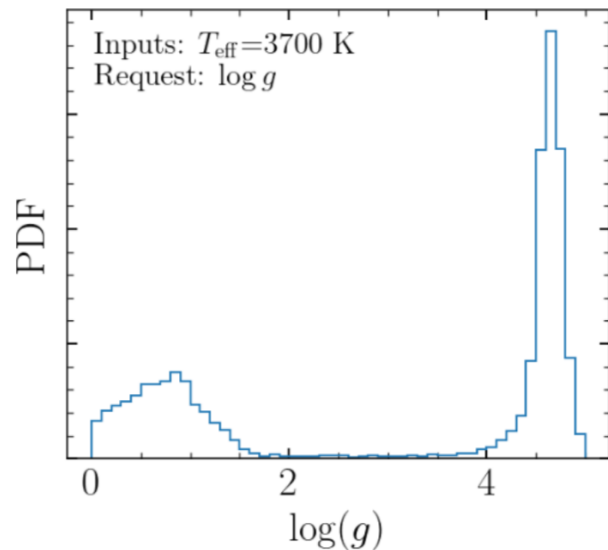


$\log(g)=4.6$  dex with  
only 0.1 dex on  
 $\log(g)$  uncertainty!



# Work-in-progress

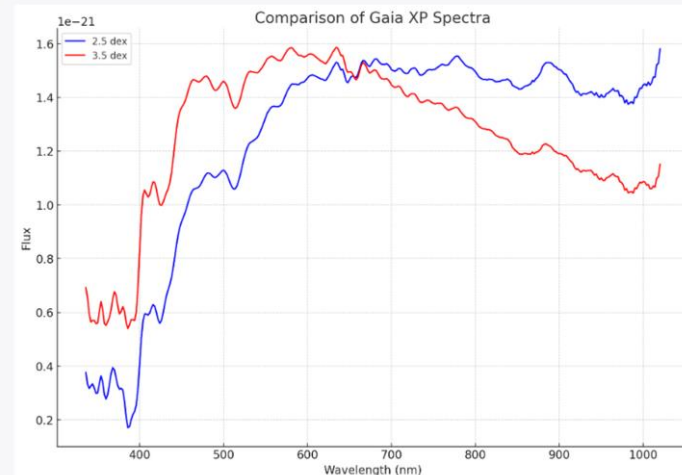
- Model will output PDF instead of a number with uncertainty from transformer-based model
- Natural language interface with Chat-GPT
  - Shoutout to Chat-Gaia by Nolan Koblischke



You

Given a star with surface gravity 2.5dex and another one with surface gravity 3.5dex, how their Gaia XP spectra look like in comparison?

Star-GPT



Here is the comparison of the Gaia XP spectra for two stars, one with a surface gravity of 2.5 dex (blue curve) and the other with 3.5 dex (red curve). The plot shows the flux as a function of wavelength for both stars, providing a visual representation of how their spectral characteristics differ.

# Conclusion

- A framework to adopt and adapt technologies behind LLMs to floating-point astronomical data
- A proof-of-concept that a single NN can perform a variety of inference tasks without fine-tuning
- Possibility of “Large Astronomy Model” in the future