AI for astronomical spectroscopy: current challenges and prospects for the WST

Nils Candebat

INAF- Observatory of Arcetri

MALSPEC : Machine Learning for Spectroscopy

Germano Sacco, Francesco Belfiore, Laura Magrini, Stefano Zibetti

nils.candebat@inaf.it

Painting in a watercolour style of a stunning view of a spiral galaxy, likely captured by a space telescope. The galaxy is centrally located and appears to be a bright, luminous core surrounded by swirling arms of gas and dust. The arms are composed of various shades of blue, white, and even hints of orange, indicating the presence of different elements and temperatures within the galaxy. The background is a deep, dark space, filled with numerous stars of varying brightness. The stars are scattered throughout the image, creating a sense of depth and vastness. The overall composition of the image highlights the intricate beauty and complexity of the galaxy, showcasing its spiral structure and

Generated using DeepSeek Janus P

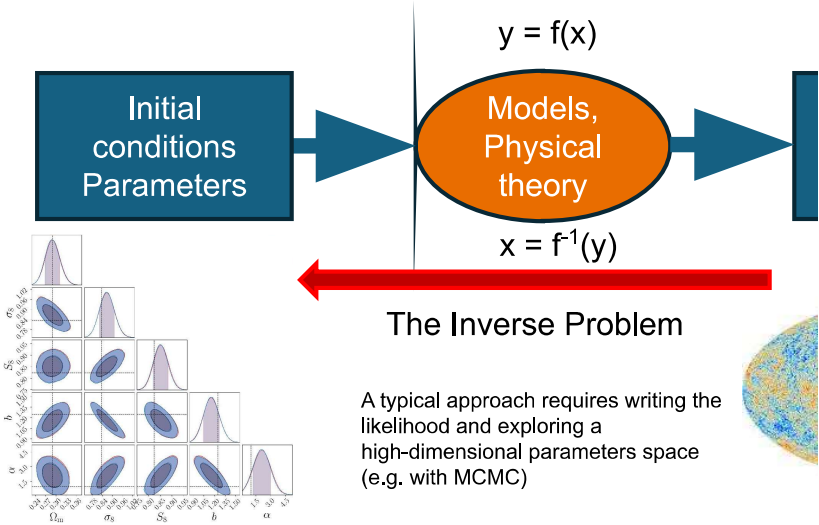
Neural Network in one slide

Universal Approximation Theorem:
Let $f : K \to \mathbb{R}$ be a continuous function, where K is a compact subset of $\mathbb{R}^n$. For any $\epsilon > 0$, there exists a feedforward neural network $\hat{f}(x)$ with a single hidden layer and a finite number of neurons such that:

$$|f(x) - \hat{f}(x)| < \epsilon \quad \forall x \in K$$

where $\hat{f}(x)$ is the output of the neural network. The network uses a non-constant, bounded, and continuous activation function σ.

Solving the inverse probl



cINN applied to stars

Drawing in an art deco style of a stunning view of the cosmos, showcasing a vast expanse of space filled with numerous stars and celestial bodies. The background is a deep blue, representing outer space, with stars of various sizes and colours scattered throughout. Some stars are bright white, while others are fainter and orange-brown, indicating different temperatures and compositions. In the centre of the image, there is a dense cluster of stars, with some forming a bright, glowing nucleus surrounded by smaller stars. This central region appears to be a galaxy or a nebula, with its intricate structure and glowing core being particularly prominent. The image also shows some nebulous formations, which are wispy clouds of gas and dust, adding to the complexity and beauty of the cosmic scene. These nebulous regions are scattered across the image, contributing to the overall sense of depth and vastness. @Hubble

Generated using DeepSeek Janus P

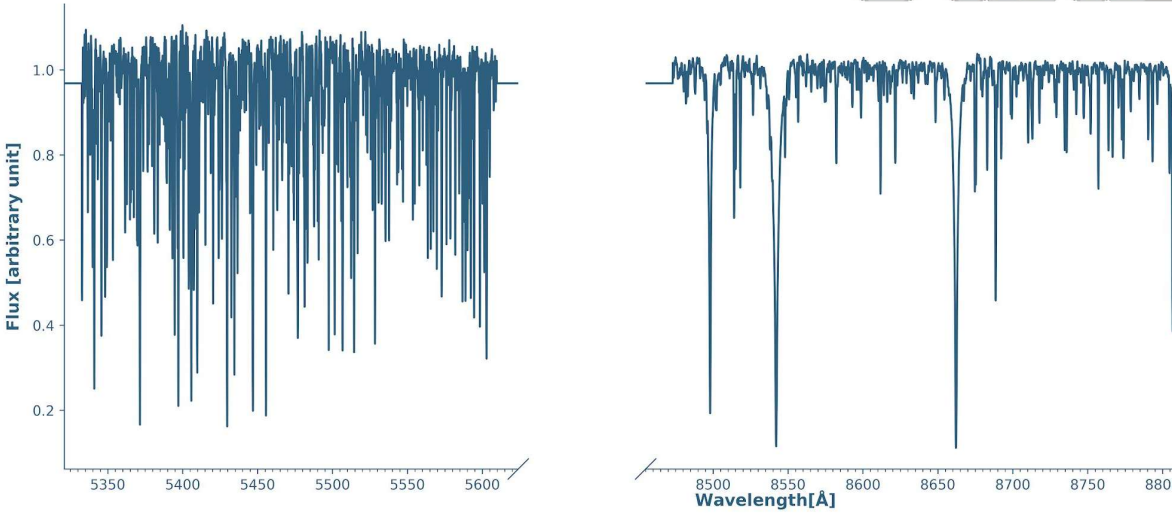
Dataset

=> Spectra: GAIA-ESO Survey (GES), *Randich et al. 2022*

- 52,841 spectra from GIRAFFE@VLT
- spectral range 5330-5610Å & 8480-8980Å
- 17,000 pixels per spectrum

=> 9 Parameters (targets) derived from multiple classical pipelines *Hourihane et al. (2022):*

a) *Temperature effective - TEFF*
b) *Surface gravity - log(g)*
c) *Metallicity abundance - [Fe/H]*
d) *Aluminium abundance - [Al/H]*
e) *Magnesium abundance - [Mg/H]*
f) *Calcium abundance - [Ca/H]*
g) *Nickel abundance - [Ni/H]*
h) *Titanium abundance - [Ti/H]*
i) *Silicon abundance - [Si/H]*



Dataset distribution:

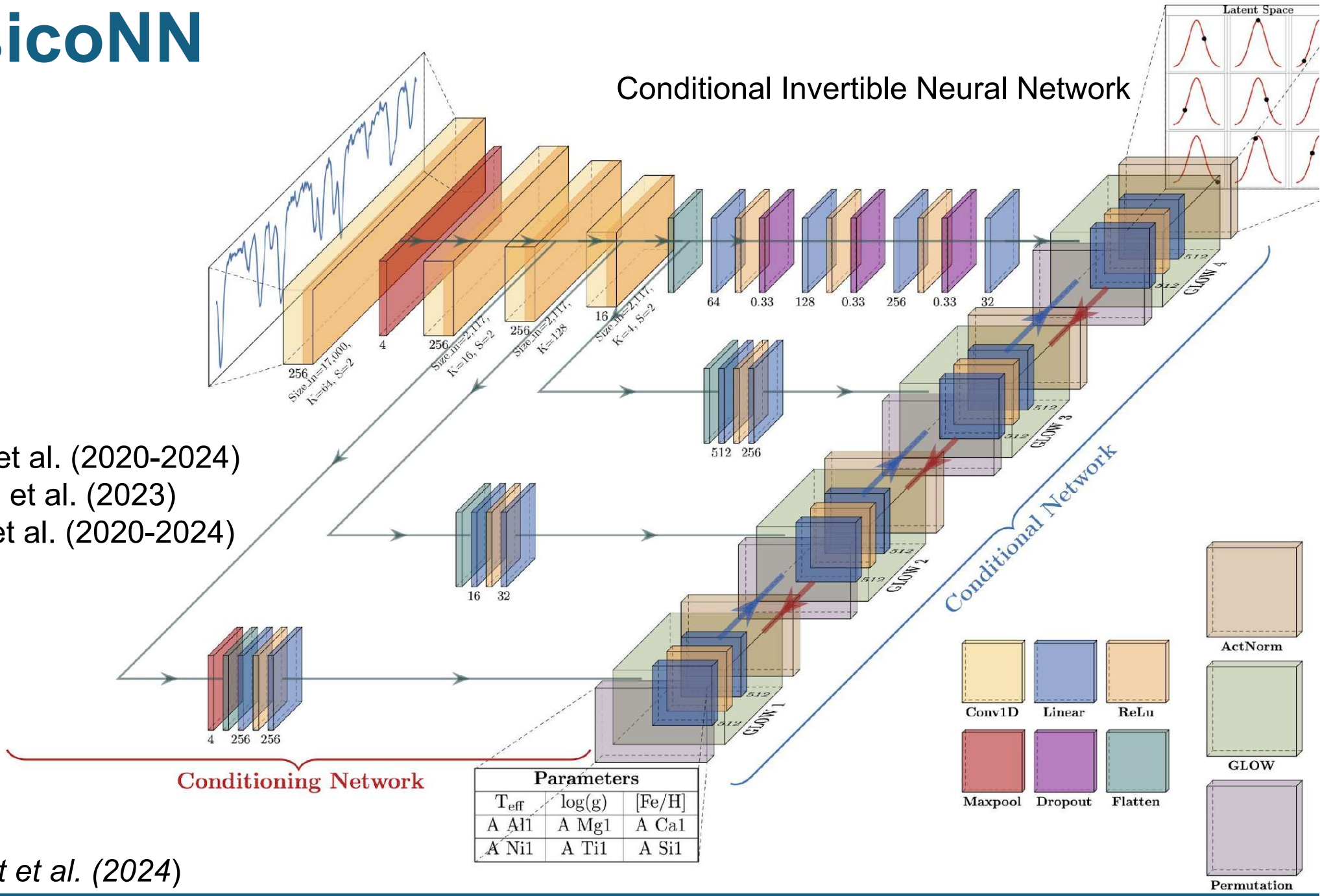
- GIRAFFE HR10-HR21 : all stars [52,841 stars]
- Training-Test : high SNR(>25) and 9 params [6,96
- Noisy : low SNR (<25) and 9 params [2,613 stars]

OssicoNN

Conditional Invertible Neural Network

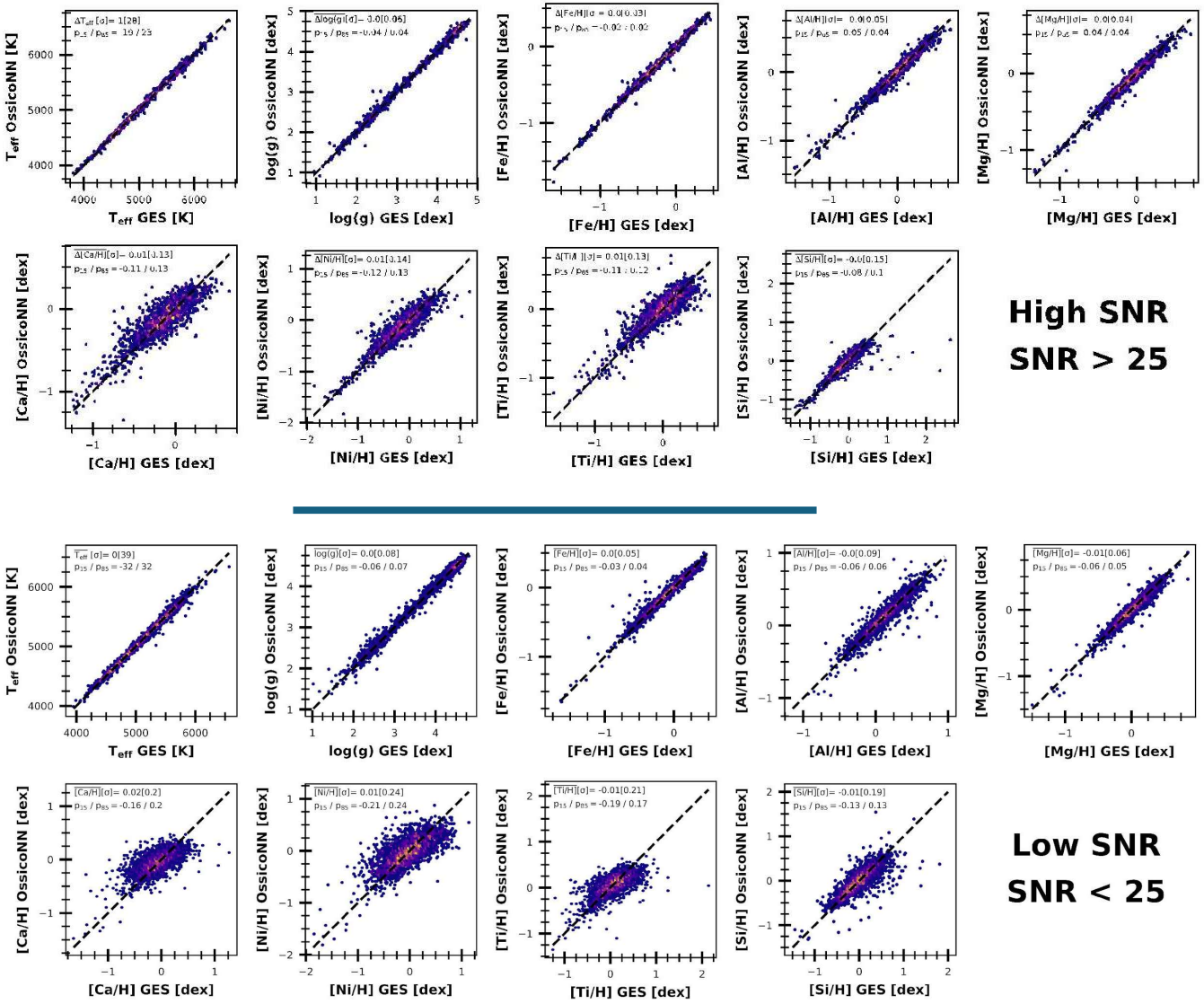
cINN :

- Kang et al. (2020-2024)
- Zhang et al. (2023)
- Ksoll et al. (2020-2024)
- …



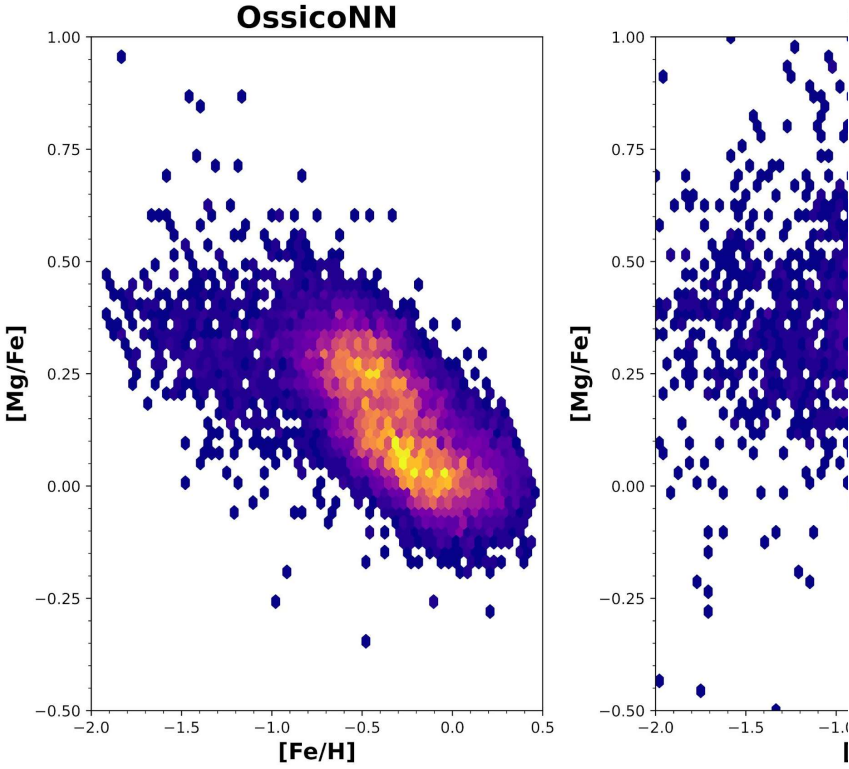
Candebat et al. (2024)

Results



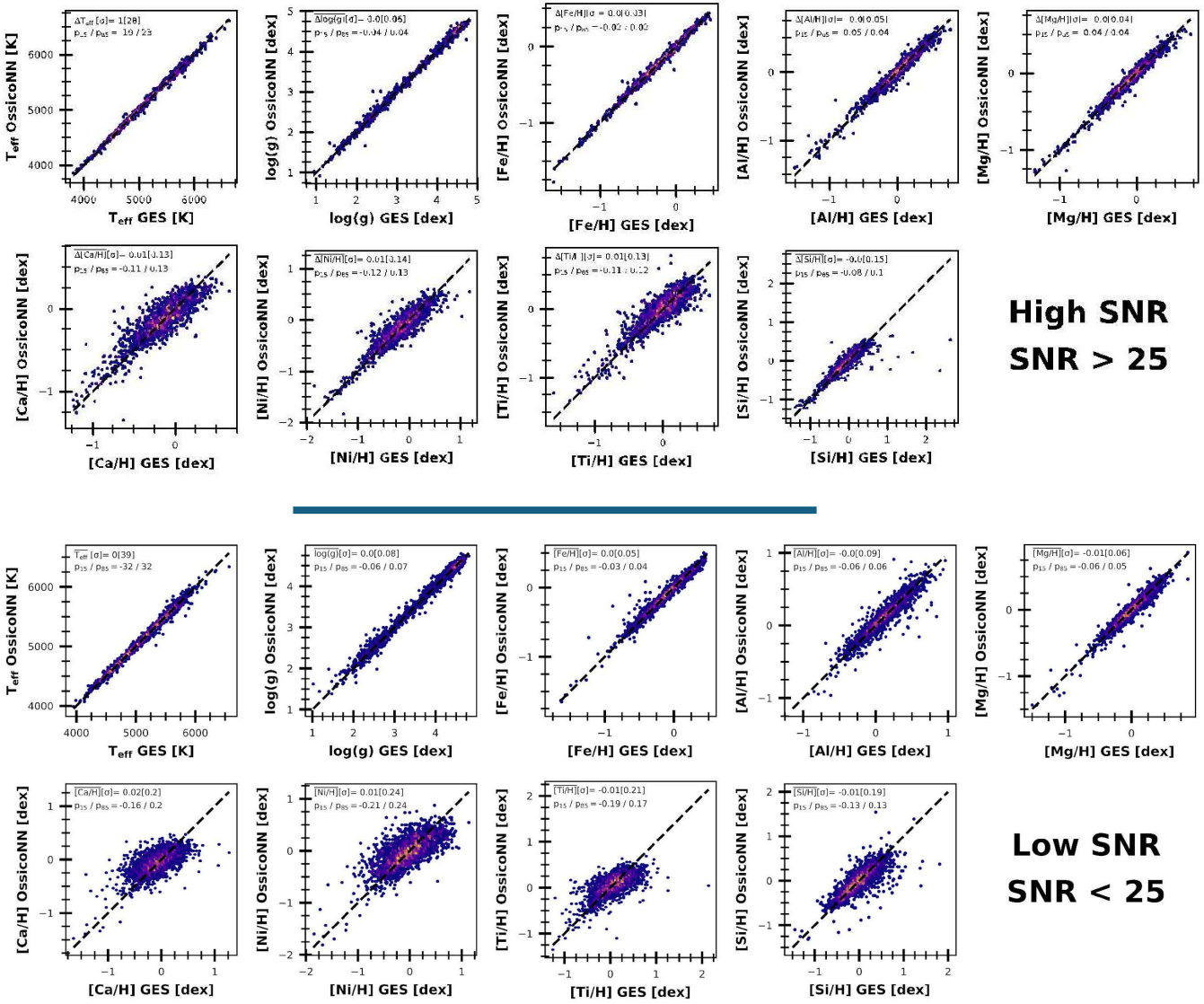
High SNR
SNR > 25

Low SNR
SNR < 25



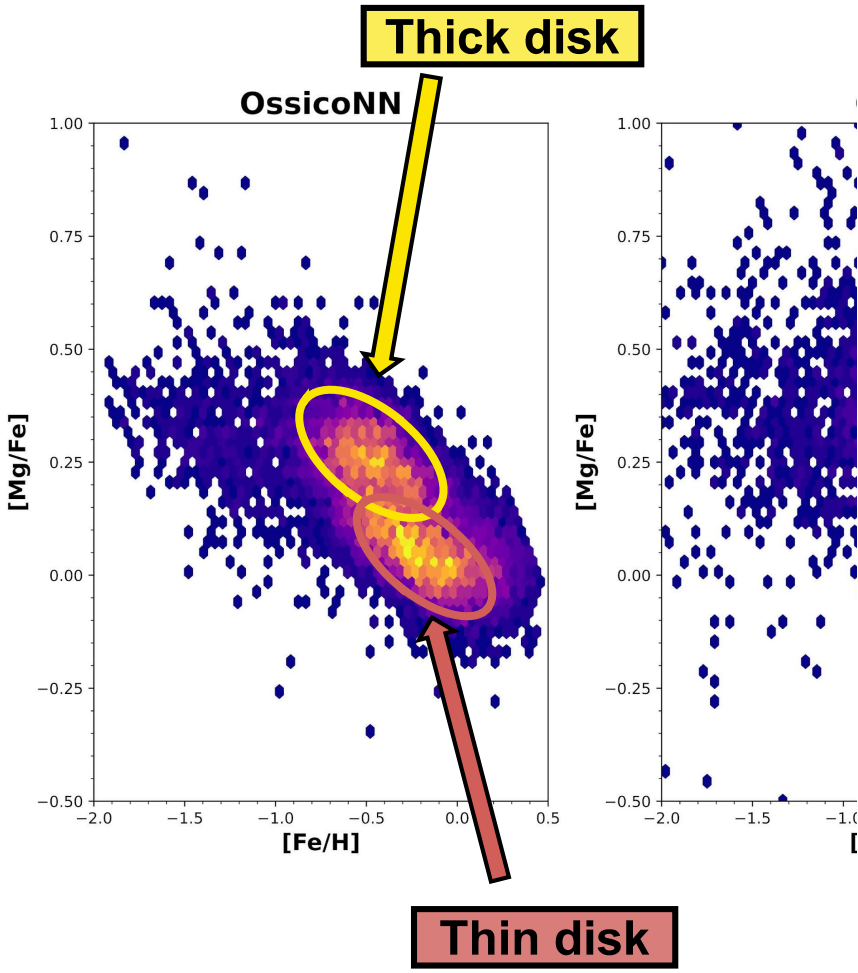
Candebat et al. (2024)

Results



High SNR
SNR > 25

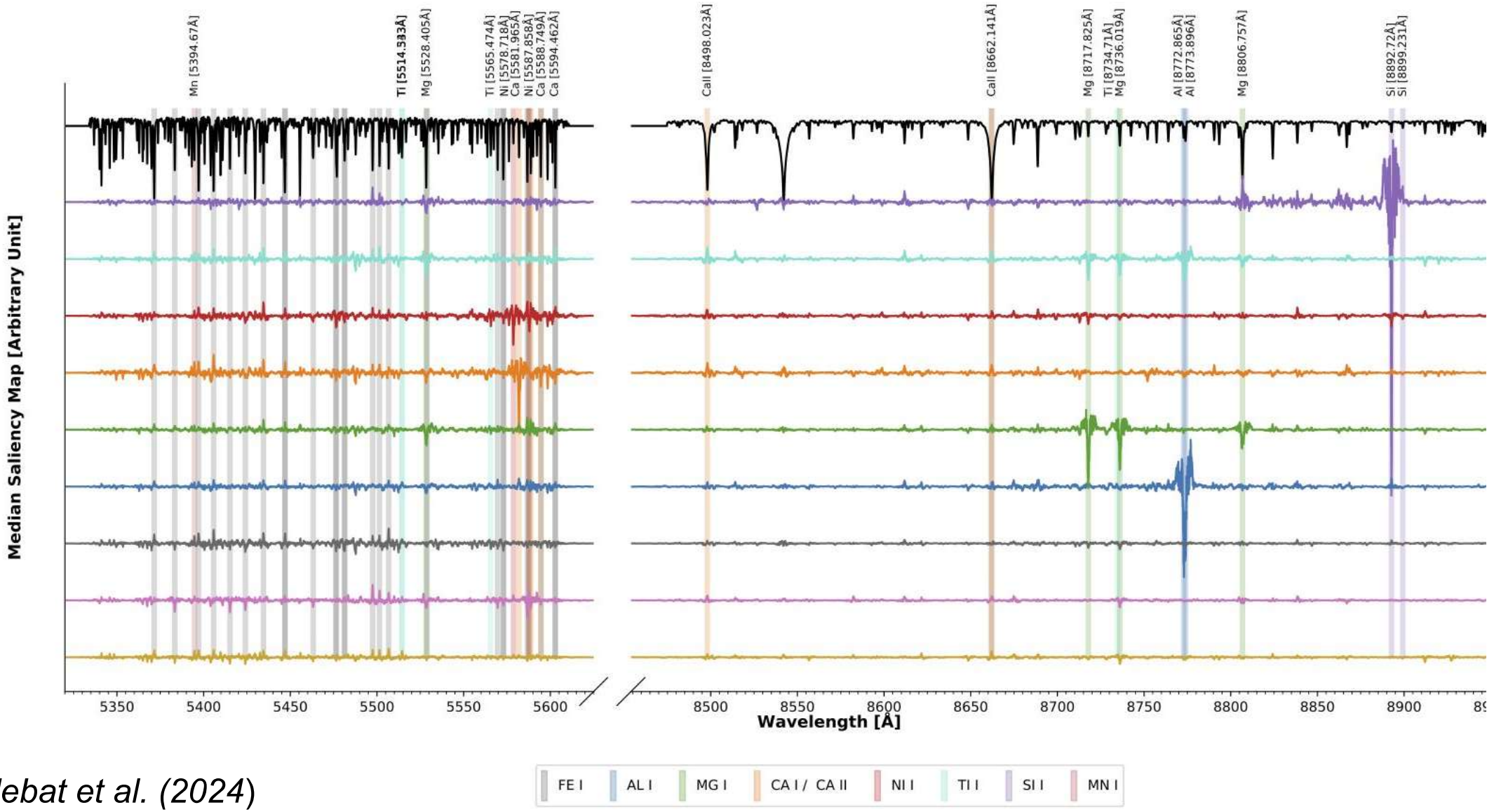
Low SNR
SNR < 25



Candebat et al. (2024)

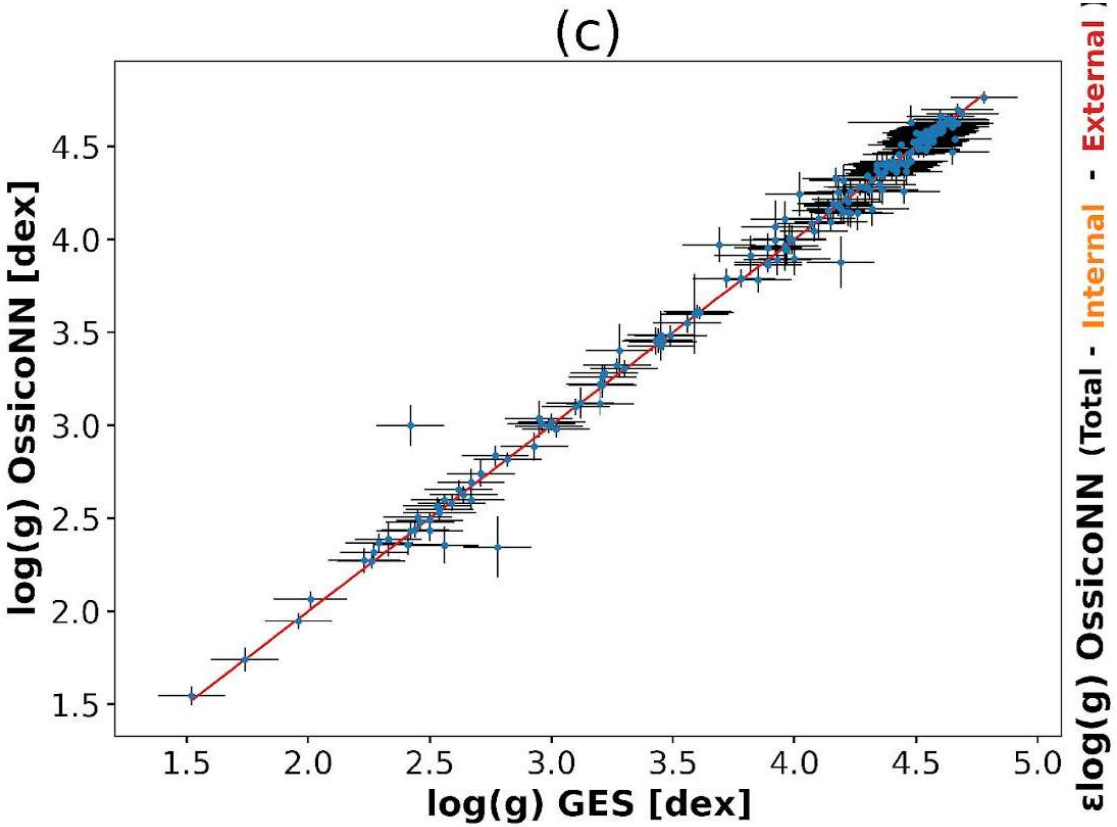
Alignment Validation: Saliency map

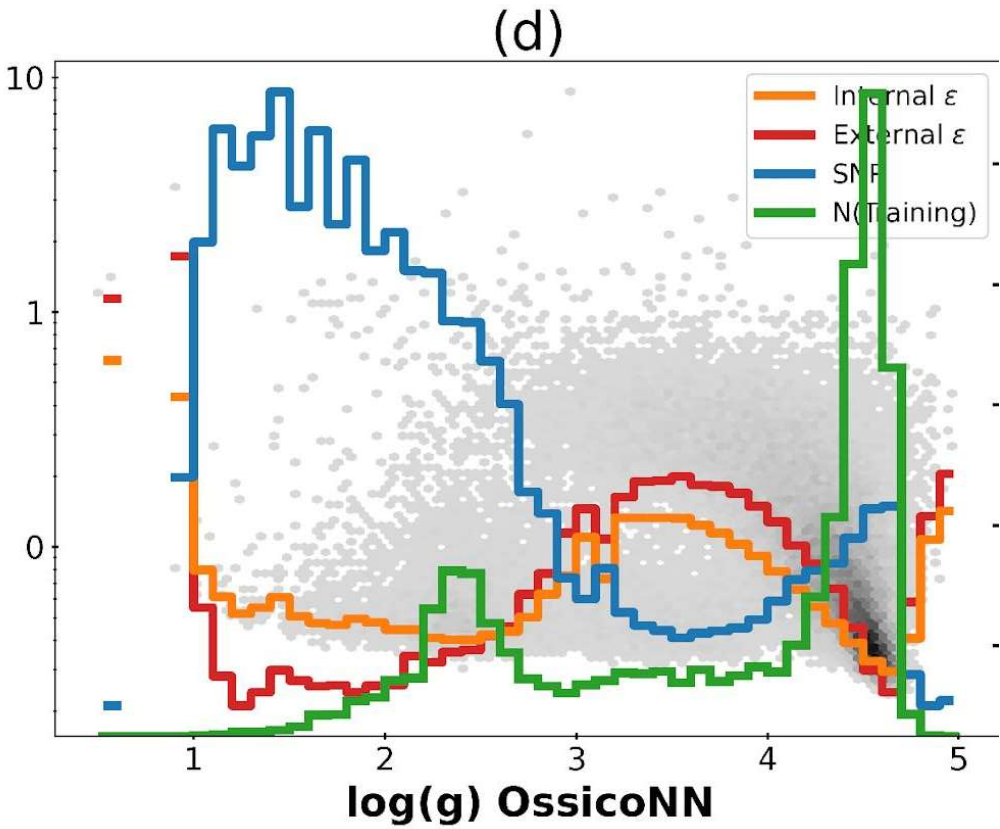
Saliency Map: technique to visualizing the input features that contribute the most to its output prediction



Candebat et al. (2024)

Uncertainty





Candebat et al. (2024)

WST MALSPEC

Large Language model Foundation Model

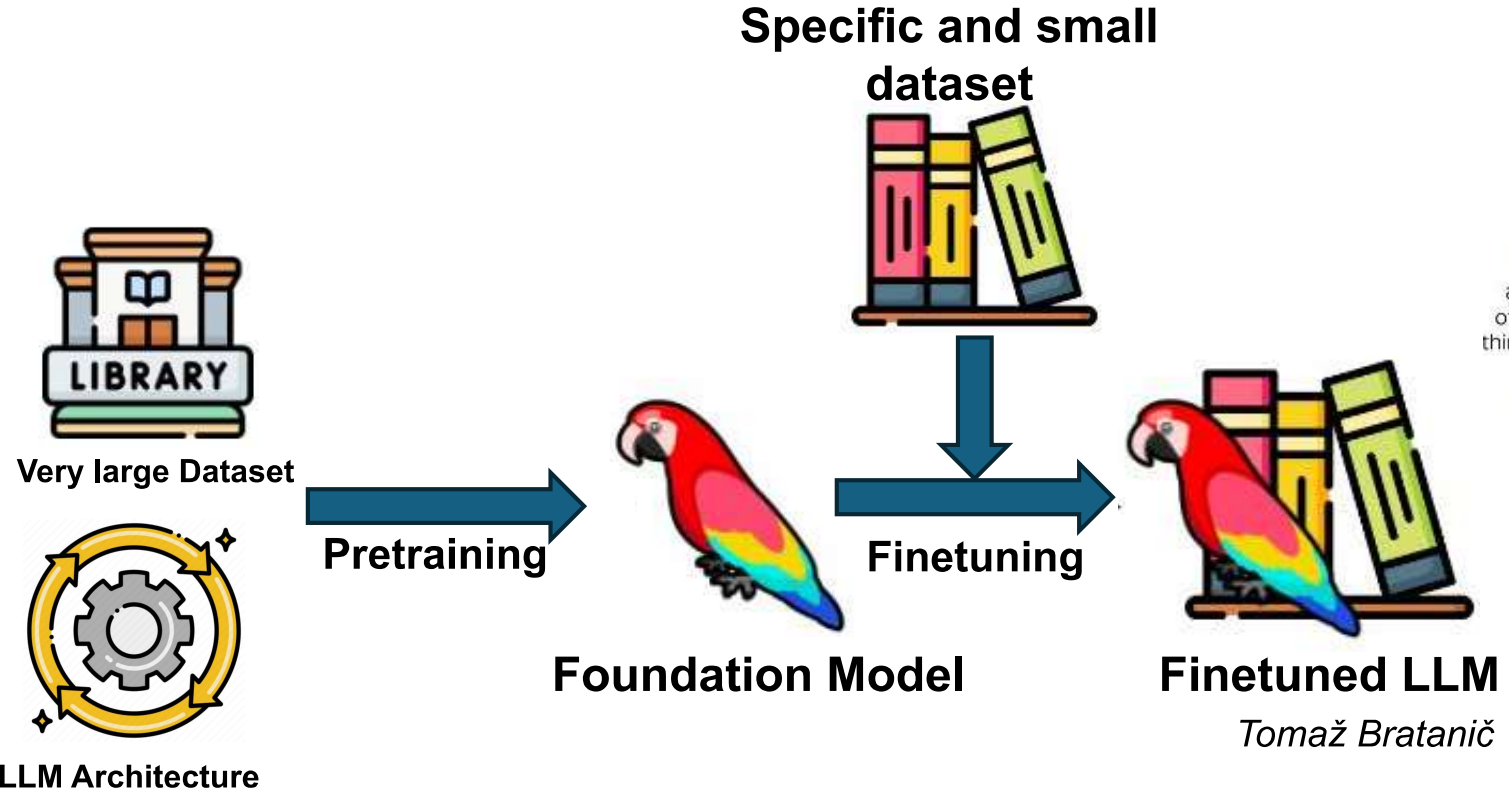
Drawing in an art deco style of a collection of galaxies captured in a deep space photograph. The galaxies are of varying shapes and sizes, with some appearing as spiral nebulae and others as elliptical or irregular shapes. The background is filled with numerous small stars, creating a dense and vibrant cosmic scene. The galaxies are illuminated in different colors, with some displaying a mix of blues, purples, and reds, indicating the presence of various gases and star formations. The overall composition suggests a rich and diverse galaxy cluster, showcasing the intricate beauty of the universe.

Generated using DeepSeek Janus F

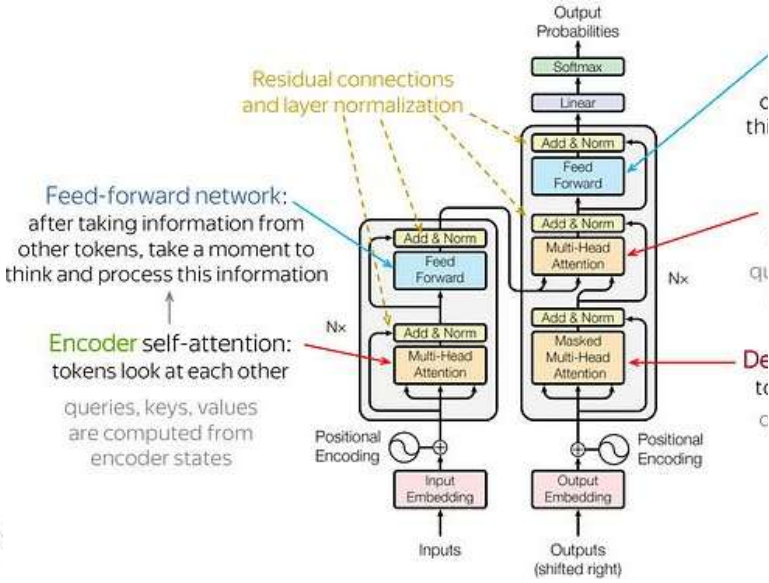
AI & cINN Applied to stars **LLM**

Idea

Deepseek Mistral Qwen Perplexity Llama Claude Hugging Face Gemini Op



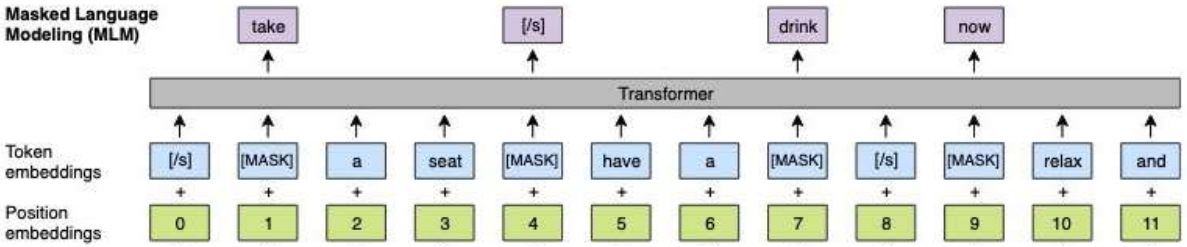
Tomaž Bratanič



Transformers – Self Atter

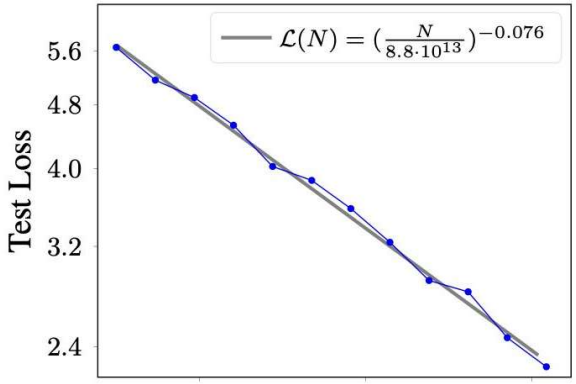
Why (1/2): self supervised/ scalling law

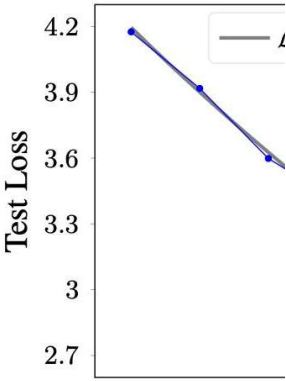
1- Self-Supervised Learning pre-training



Lample & Conneau 2019

2 - Scalling Law





Compared to supervised method (first slides) there is no need to label the data with classical method

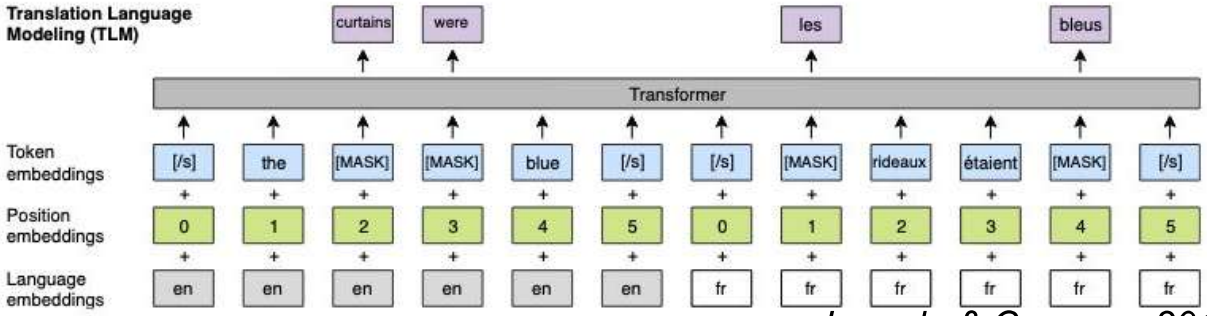
=> No physics to assum
=> No need to first create label, you can use a larger dataset, important when label are hard to derive (ex: non LTE physics)

There is therorically no limit or slowing i

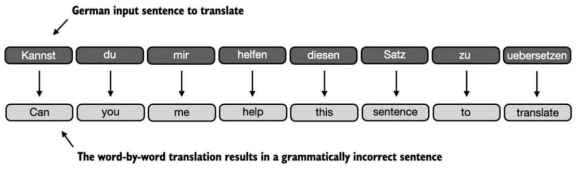
=>Each night of WST will improve th

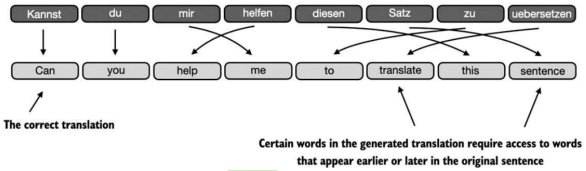
Why (2/2): Generalization / Versatility

3 - Cross-Domain Generalization



Lample & Conneau 2019





❌ Supervised training

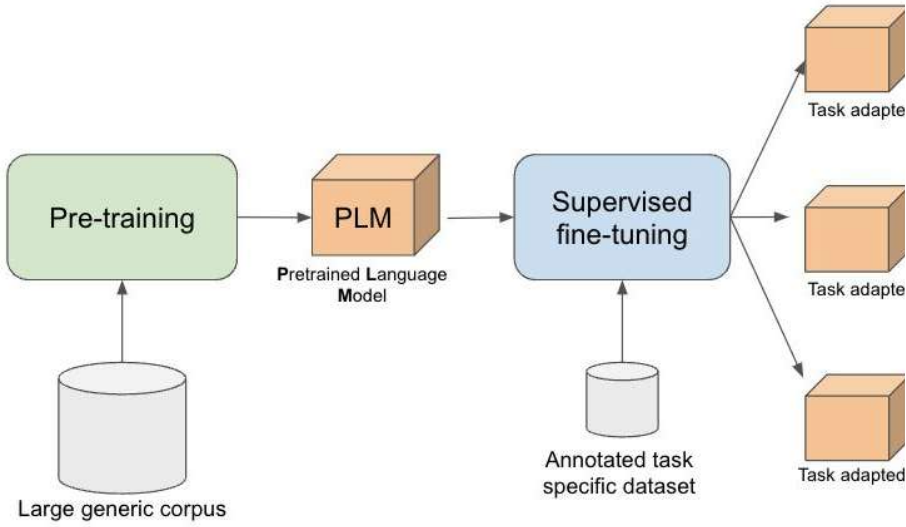
✅ LLM

Sebastian Raschka

Due to self-attention mechanism and tokenisation, LLM learn universal patterns, not memorizing sequences and are permutation-invariant reasoner (e.g., *"redshift = 0.5"* and *"z = 0.5"*)

=>LLM can be trained on diverse observational data spanning cosmic scales (galaxy, stars…)
=>LLM can be trained on diverse instruments (eg: WST and MOONS)

4- Versatility



After pre-training you can finetune for very di
with very few labels (100-1000)
Eg: create WST-like spectra, classify galaxy,

=> Foundation models arrive pre-trained a
deploy, enabling diverse teams to implem
AI capabilities with minimal technical ove

Foundation Model in Astro

‘Chat GPT’ like

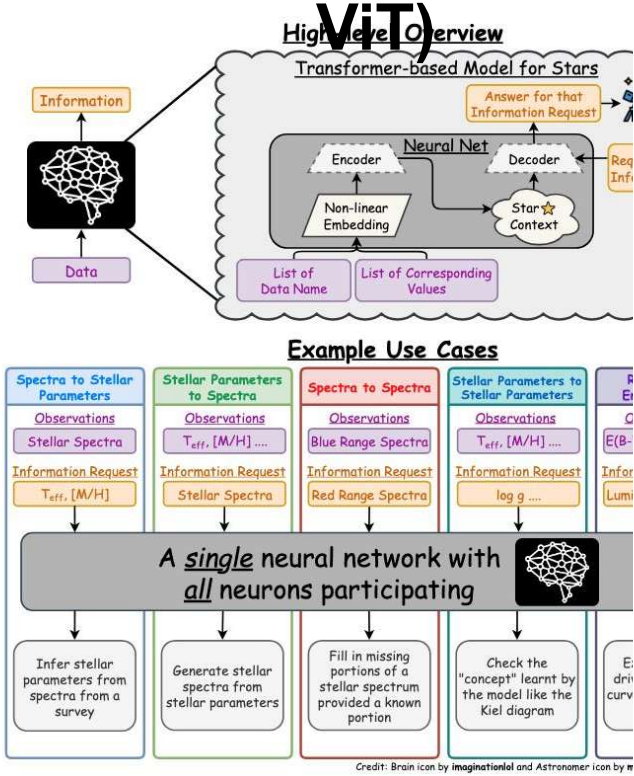
Nguyen et al.*(2023)*

AstroLlama

Cyuca & Ting 2023,

Science data processing

ViT)

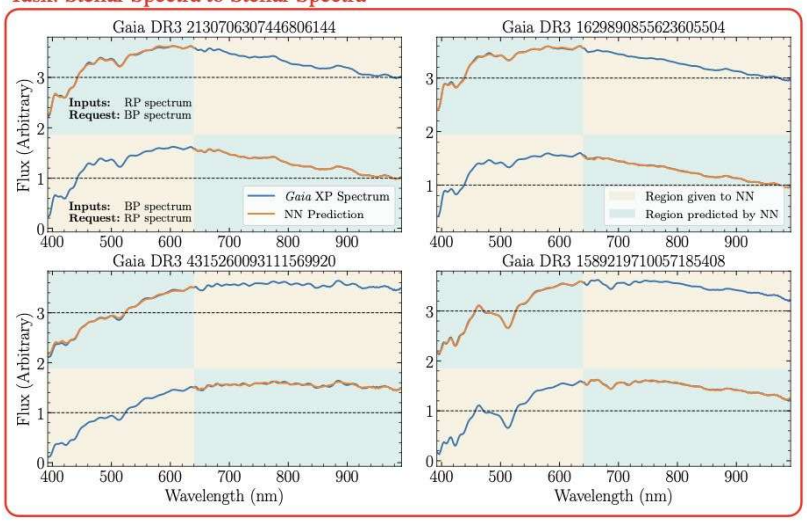


Lung & Bov

Foundation Model
Stars with a
Transformer-based

LLM for Astro

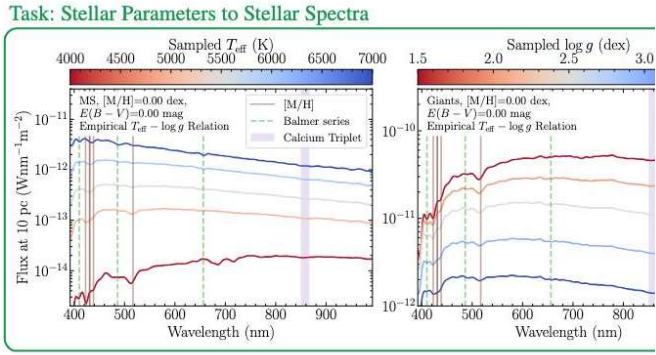
Task: Stellar Spectra to Stellar Spectra

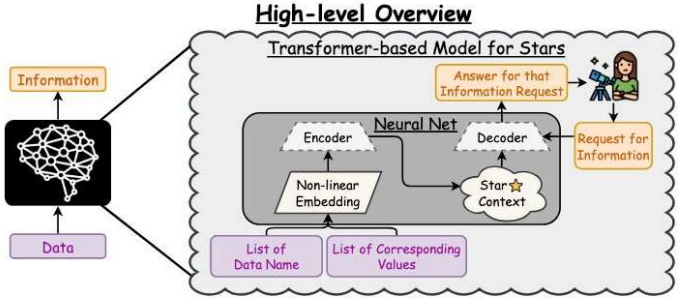


Dataset:

110 *Gaia* XP coefficients,
$G_{BP} - G_{RP}$, $J - H$, J -K,
T_{eff}, log$g$, [M/H],
logarithmic$E(B-V)_C$,
and *Gaia*G-band pseudo-luminosity

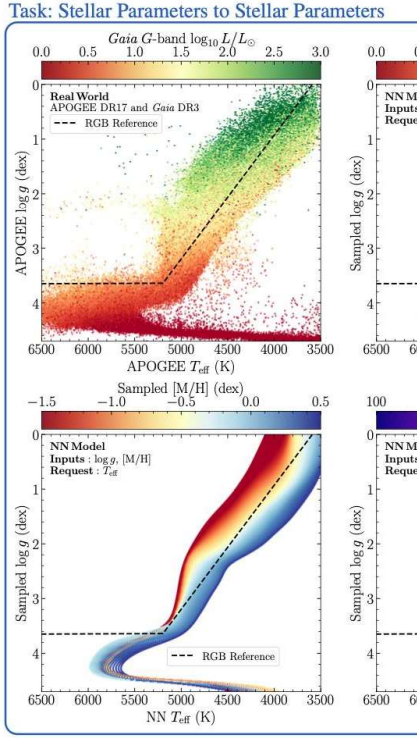
Task: Stellar Parameters to Stellar Spectra



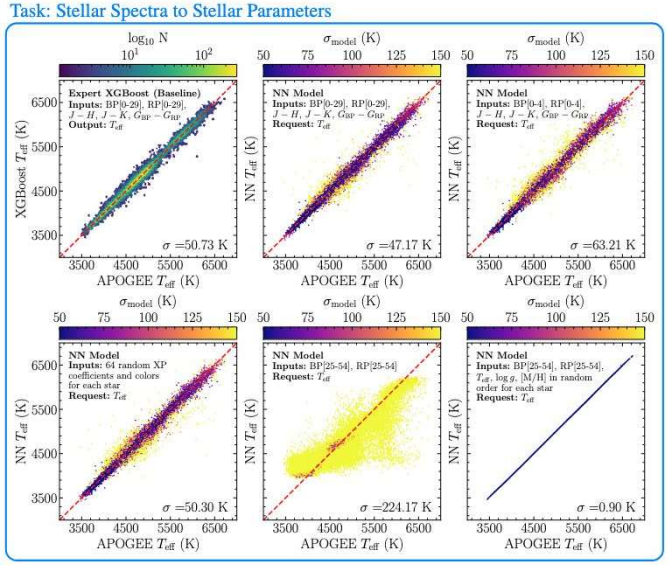


Lung & Bovy (2023)

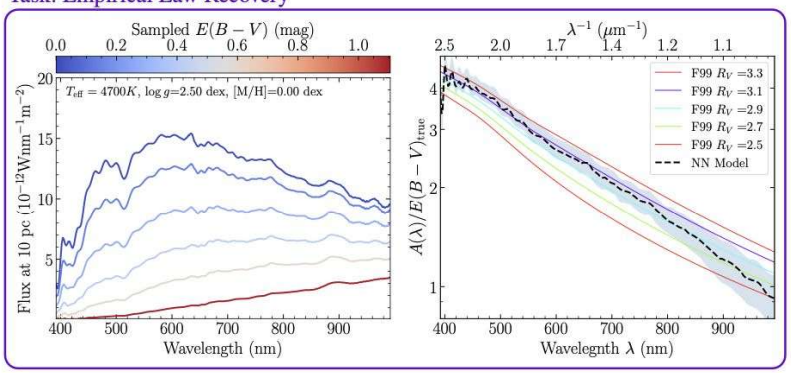
Task: Stellar Parameters to Stellar Parameters



Task: Stellar Spectra to Stellar Parameters

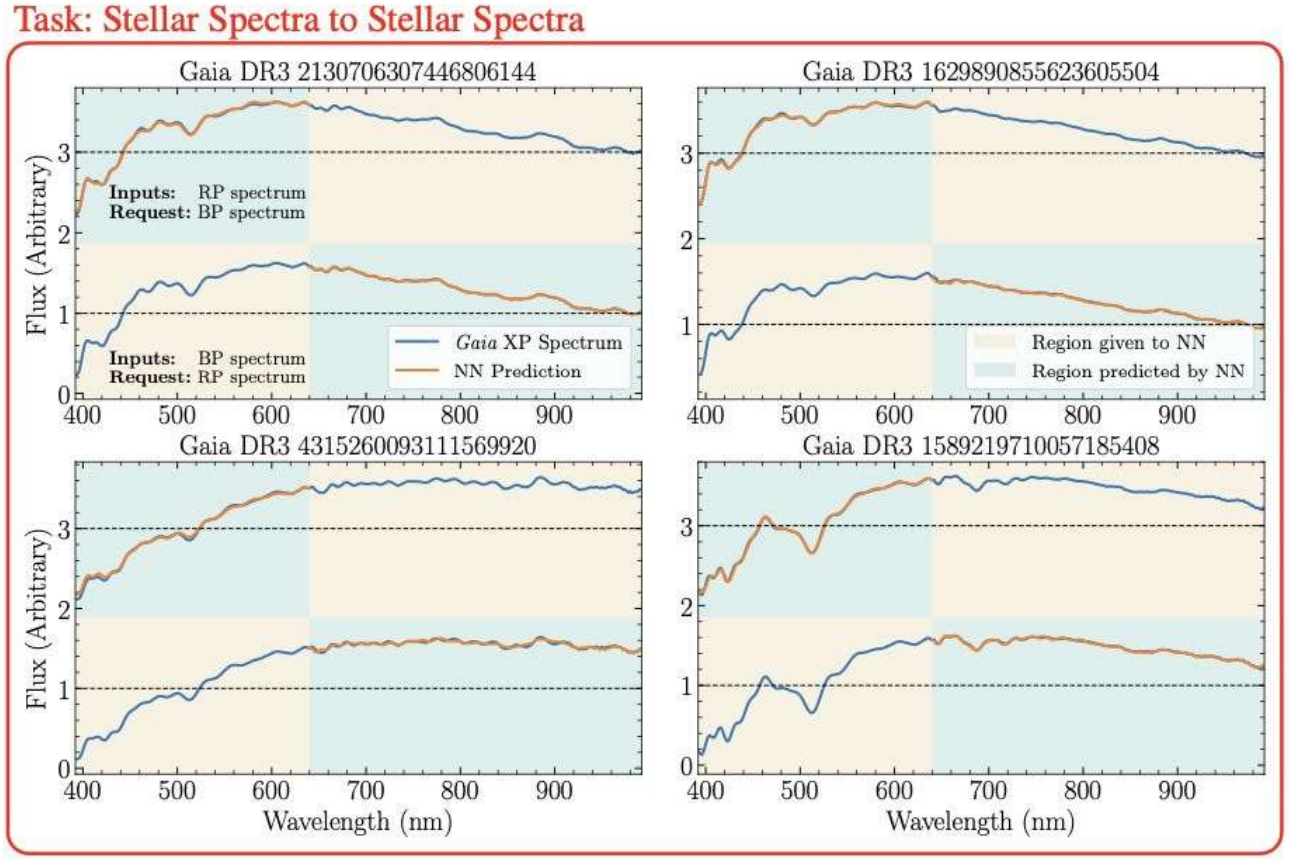


Task: Empirical Law Recovery

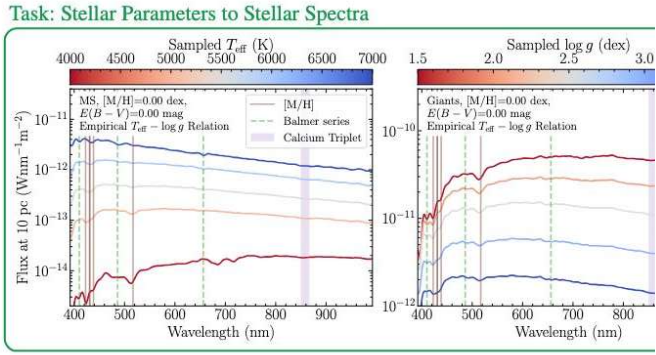


AI & cINN | Applied to stars | **LLM**

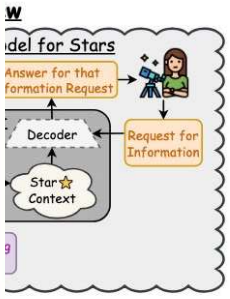
LLM for Astro



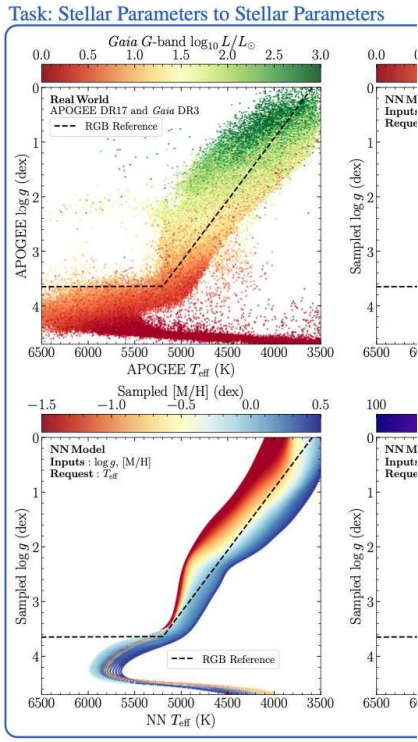
nts,

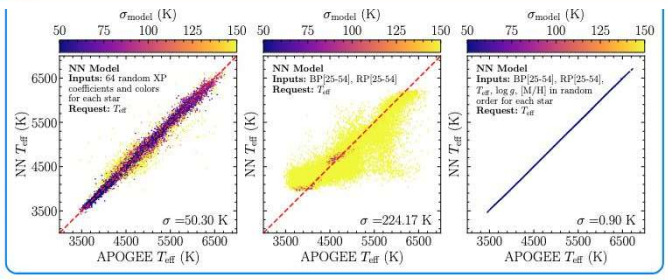


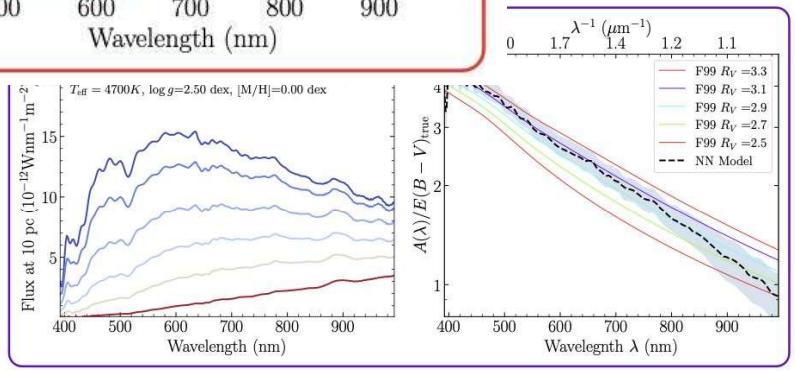
udo-luminosity



Bovy (2023)



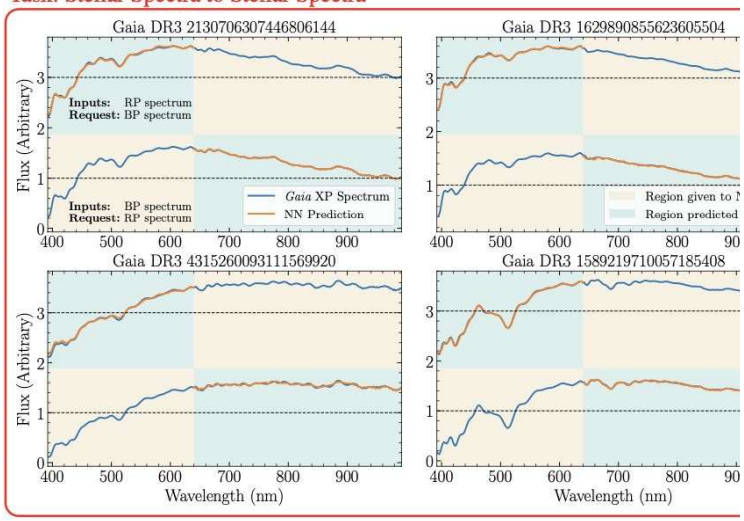




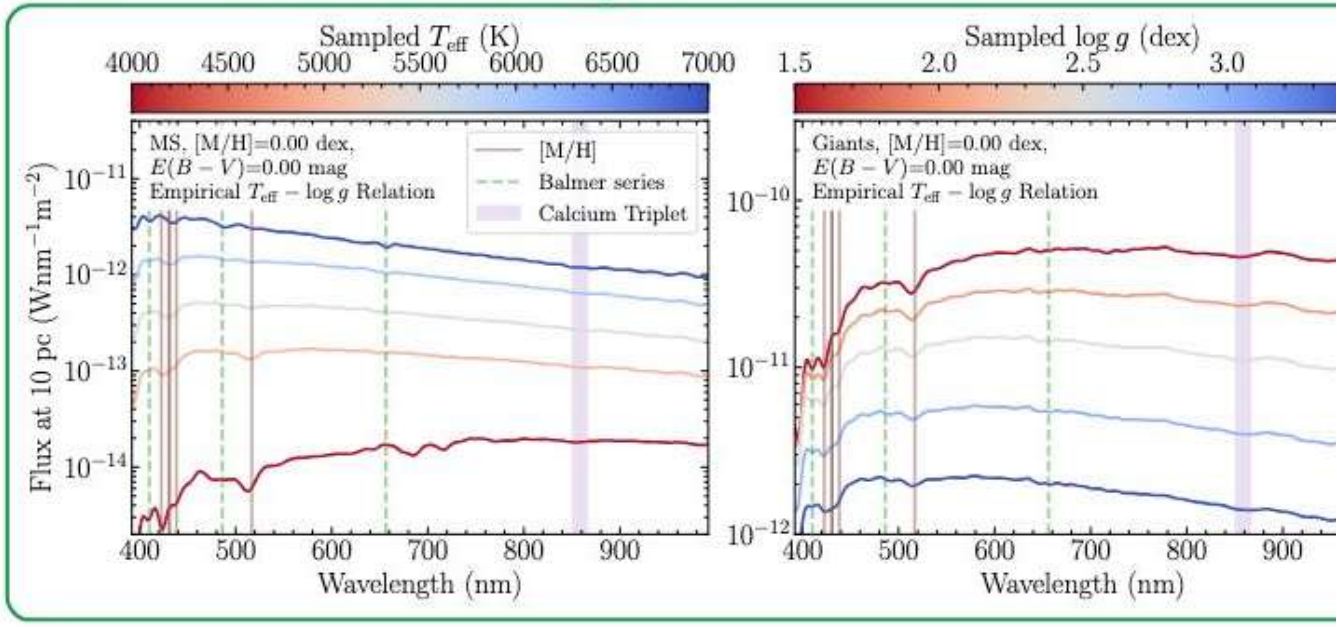
AI & cINN Applied to stars **LLM**

LLM for Astro

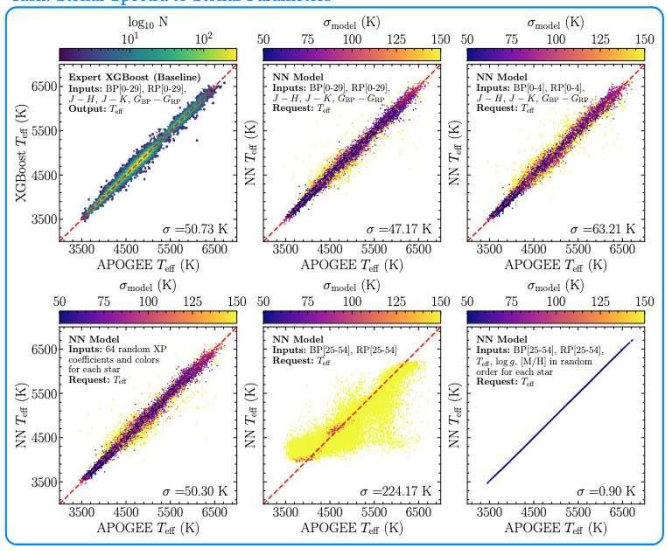
Task: Stellar Spectra to Stellar Spectra



Task: Stellar Parameters to Stellar Spectra

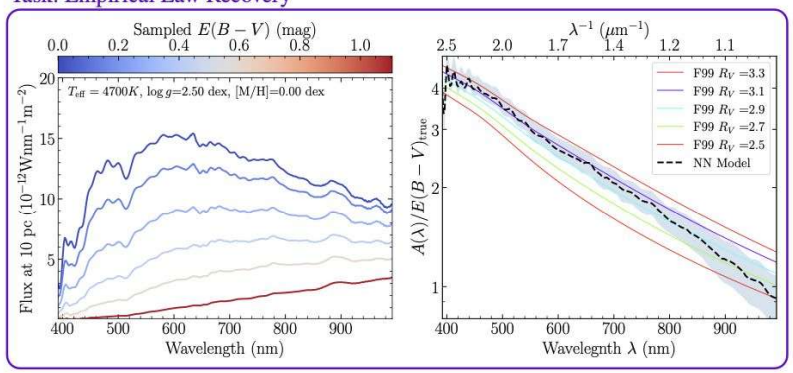


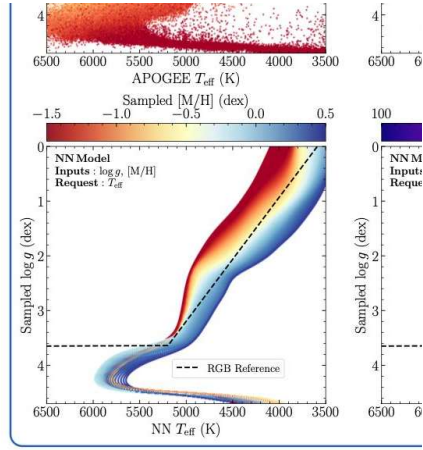
Task: Stellar Spectra to Stellar Parameters



Lung & Bovy (2023)

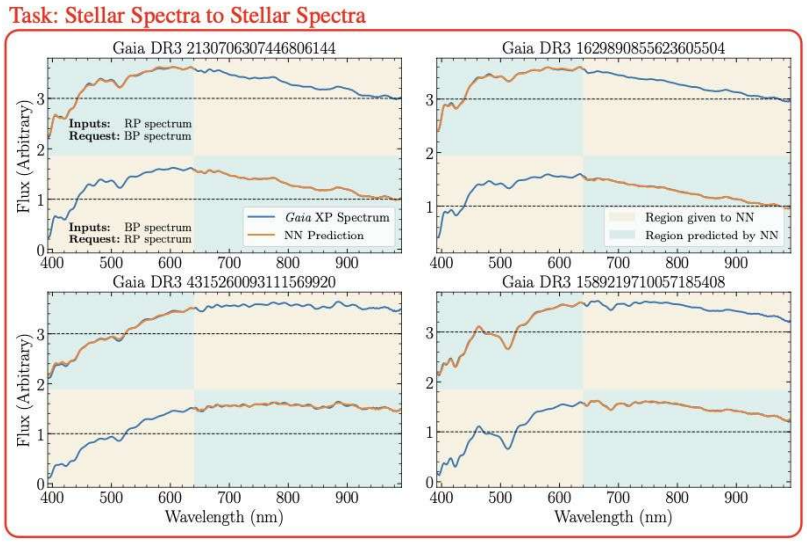
Task: Empirical Law Recovery

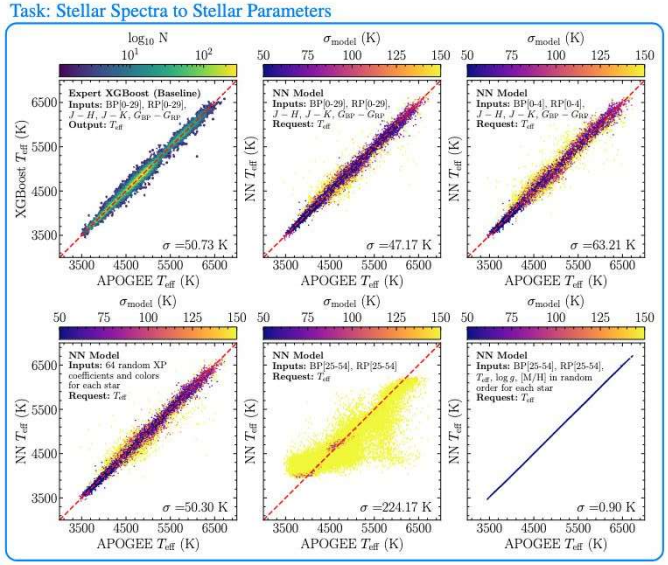




AI & cINN | Applied to stars | **LLM**

LLM for Astro





Dataset:

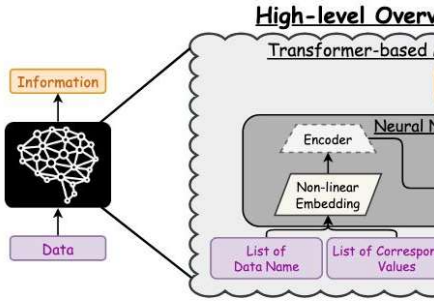
110 *Gaia* XP coefficie

$G_{BP} - G_{RP}$, $J - H$, J -K

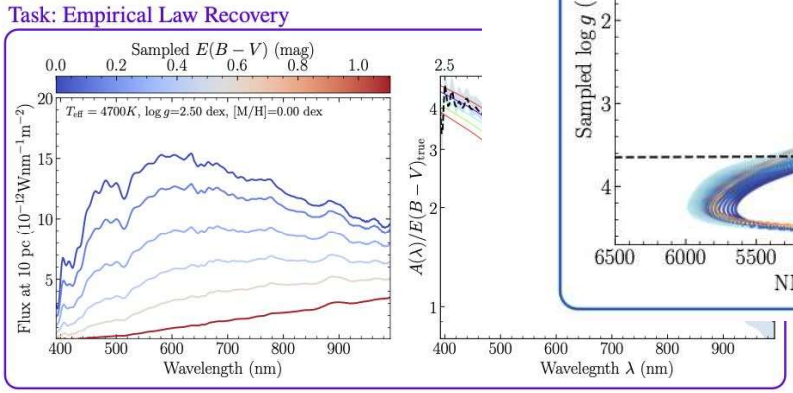
$T_{\rm eff}$, log$g$, [M/H],

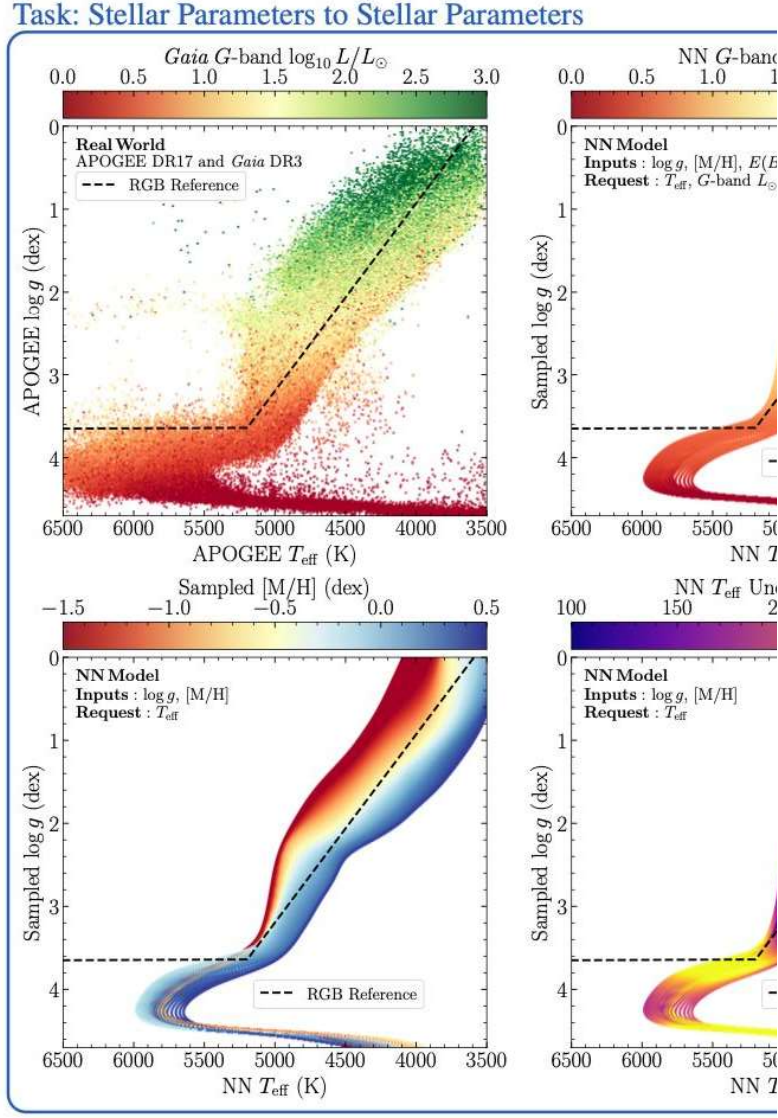
logarithmic$E(B-V)_C$,

and *Gaia*G-band pse

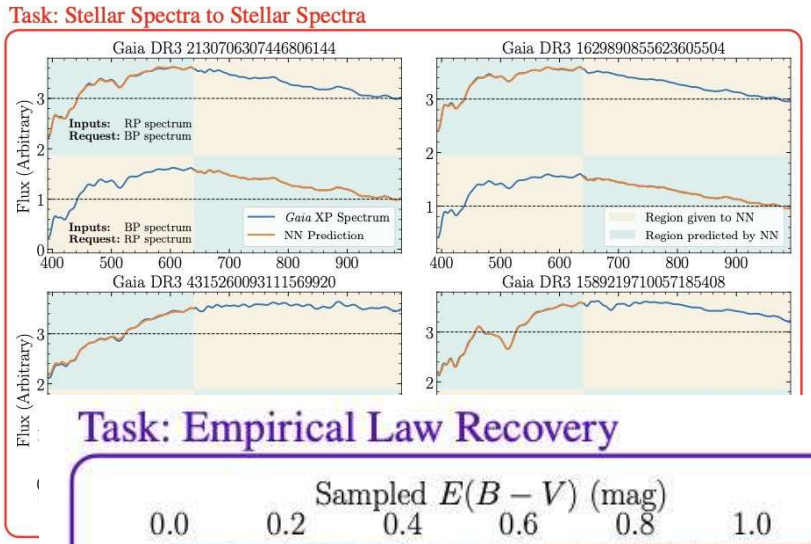


Lung





LLM for Astro



Dataset:

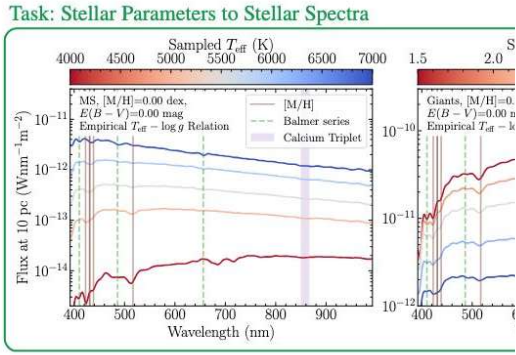
Spectrum: 110 *Gaia* XP coefficients,

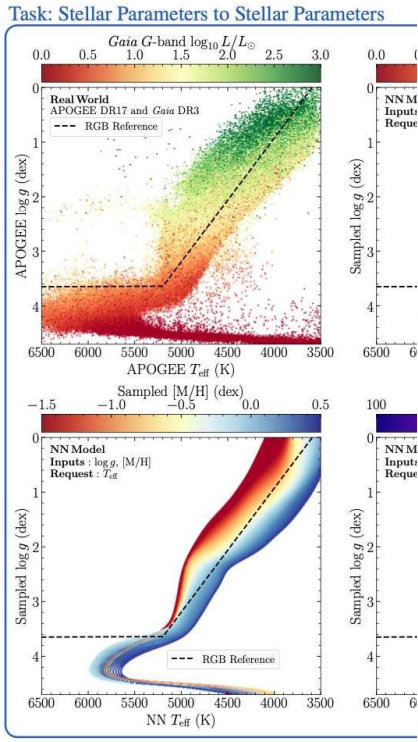
$G_{BP} - G_{RP}$, $J - H$, $J - K$,

T_{eff}, $\log g$, [M/H],

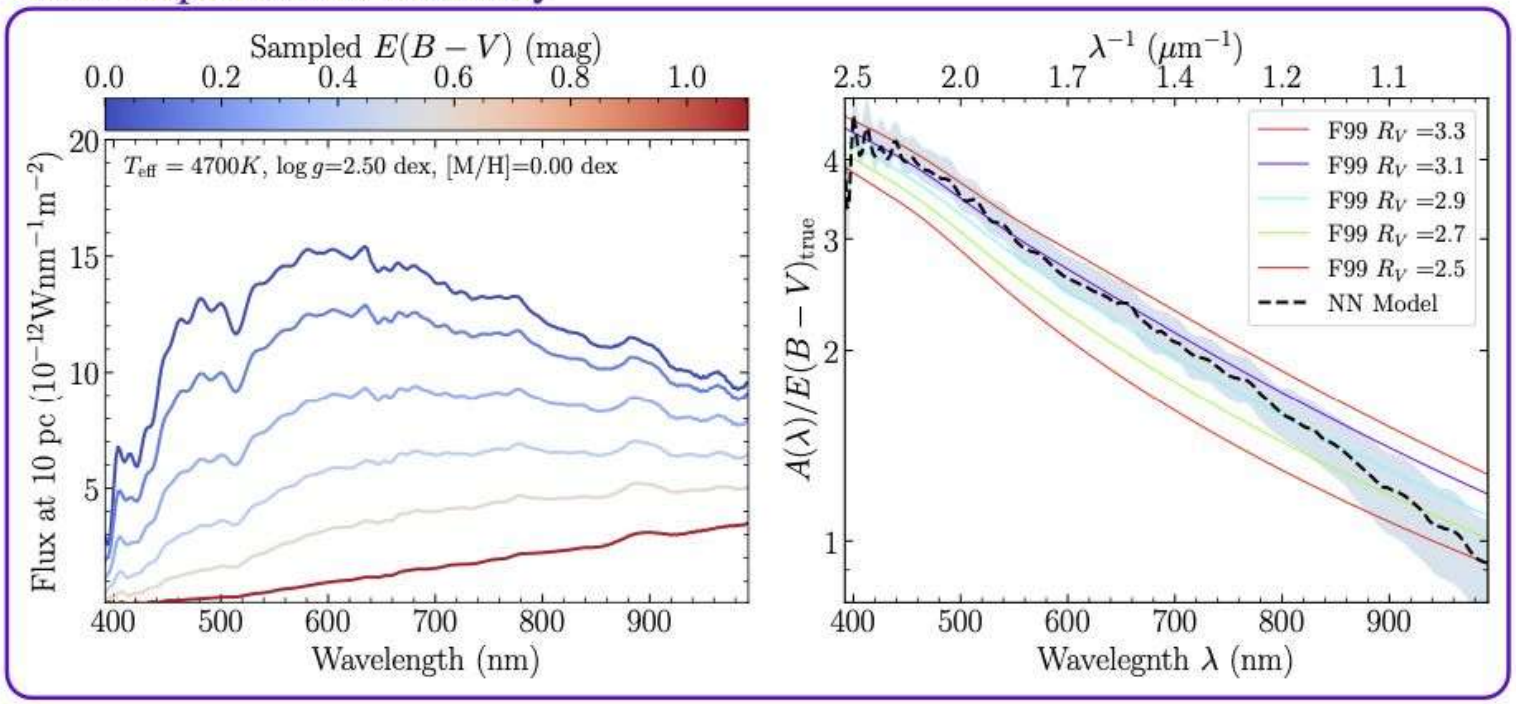
Logarithmic $E(B-V)_C$,

Gaia G-band pseudo-luminosity

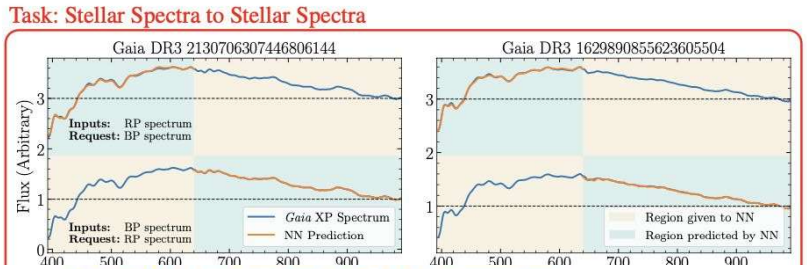




Task: Empirical Law Recovery

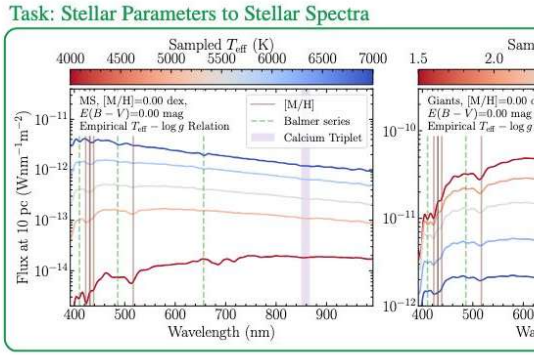


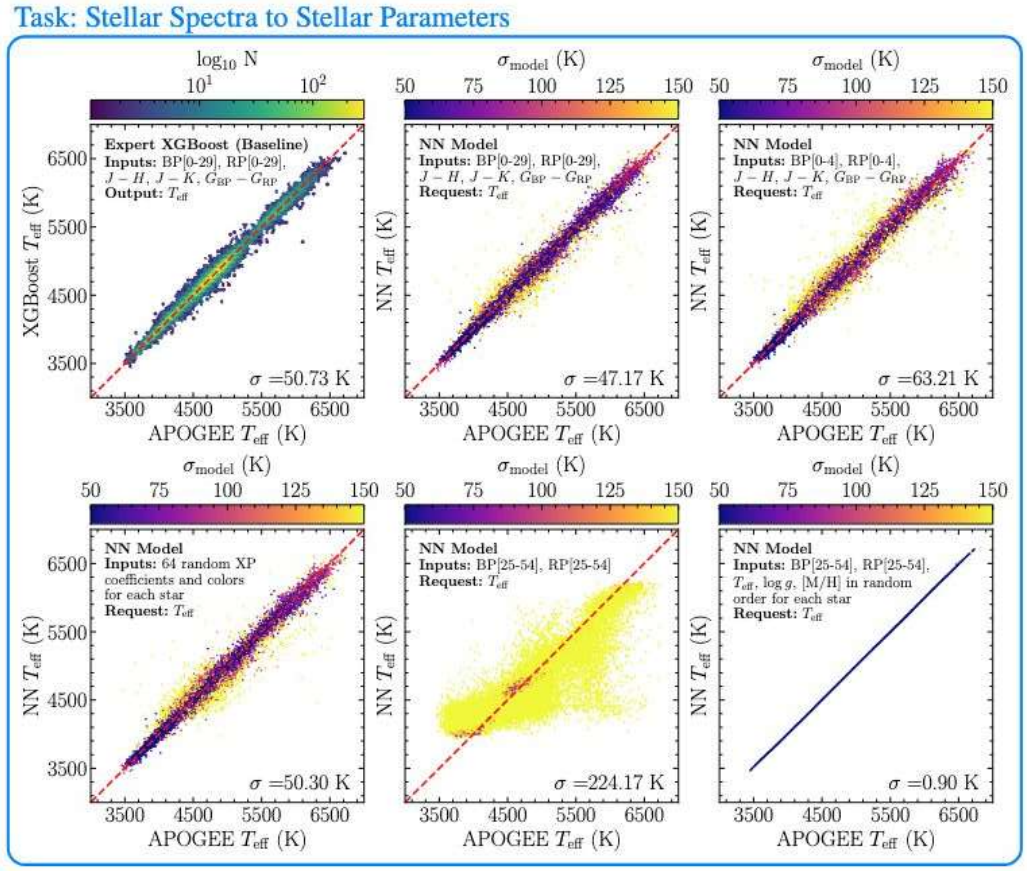
LLM for Astro

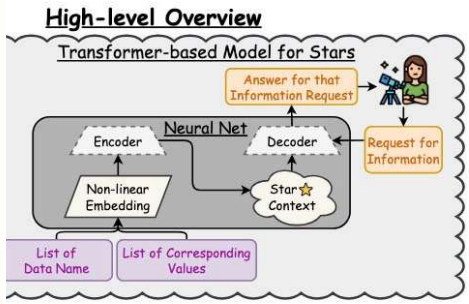


Dataset:

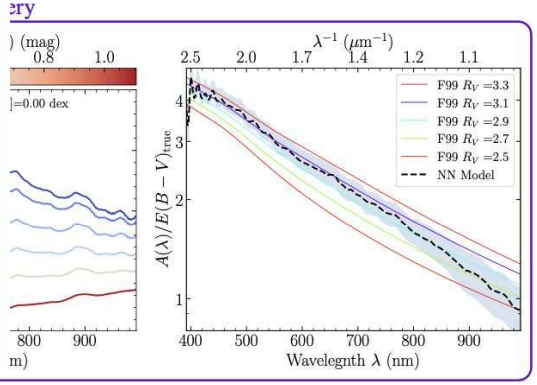
110 *Gaia* XP coefficients,
$G_{BP} - G_{RP}$, $J - H$, $J - K$,
T_{eff}, log$g$, [M/H],
logarithmic$E(B-V)_C$,
G-band pseudo-luminosity

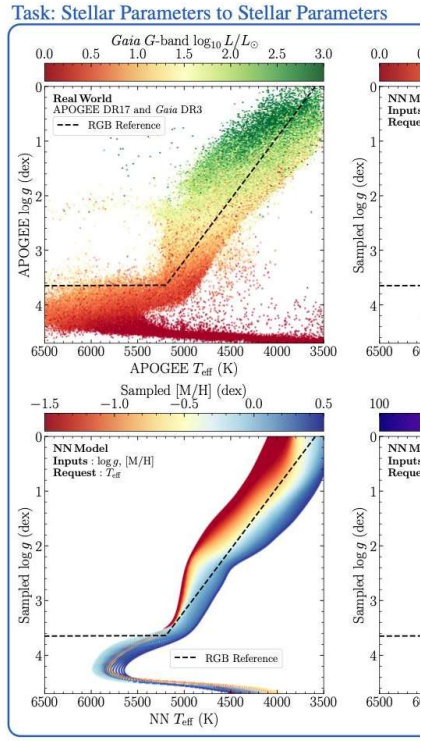




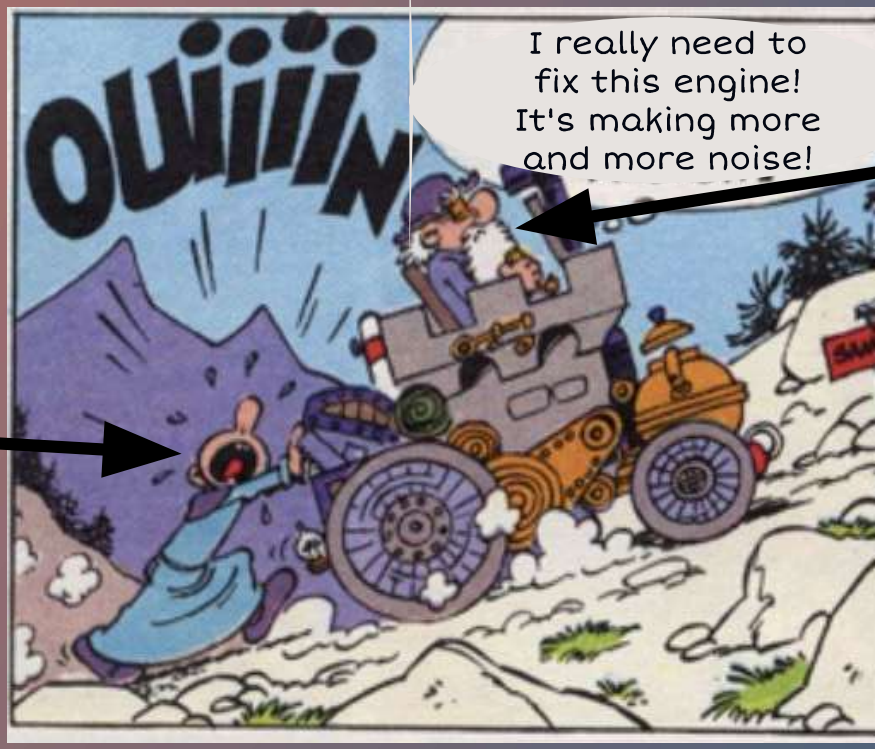


Lung & Bovy (2023)



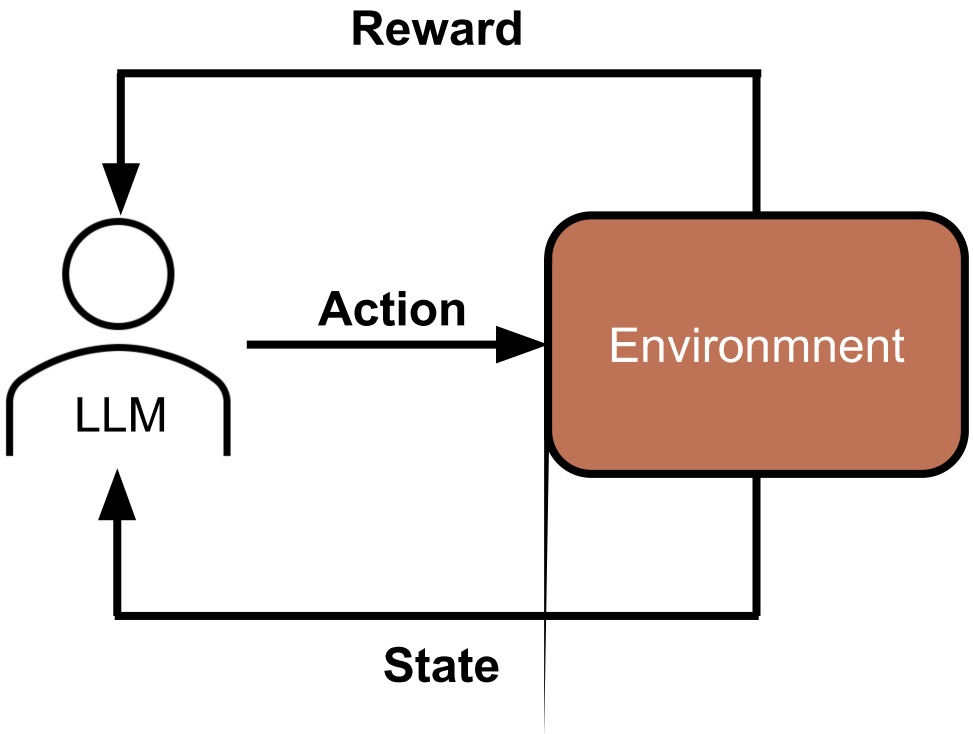


AI agent



Leonard, un genie en balade

What is an AI Agent



An artificial intelligence (AI) agent is a program that can interact with its environment, collect data, and use the data to perform self-determined tasks to meet predetermined goals.

Humans set goals, but an AI agent independently chooses the best actions it needs to perform to achieve those goals.

Can I have a cup of coffee please?

Sure! let me prepare it

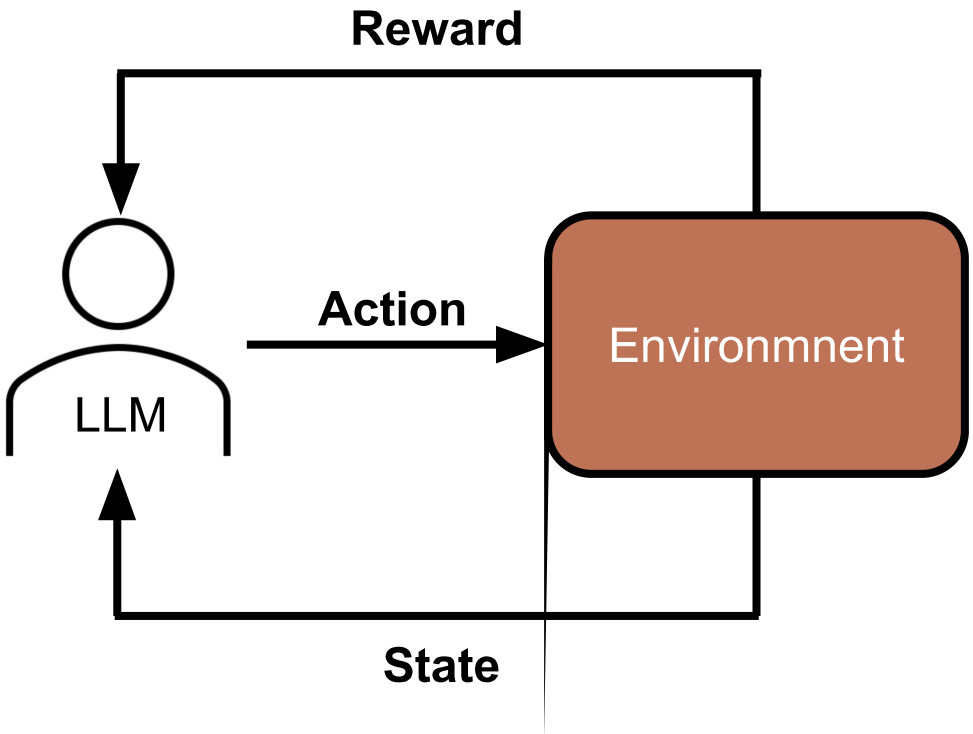
Planning:
The user wants me to prepar coffee. I need to check if the co machine is connected and de which coffee the user want.
Action:
{Env: coffee mach}: Is the co machine connected? is the tan water empty ?
{Env: clock}: what times is it?

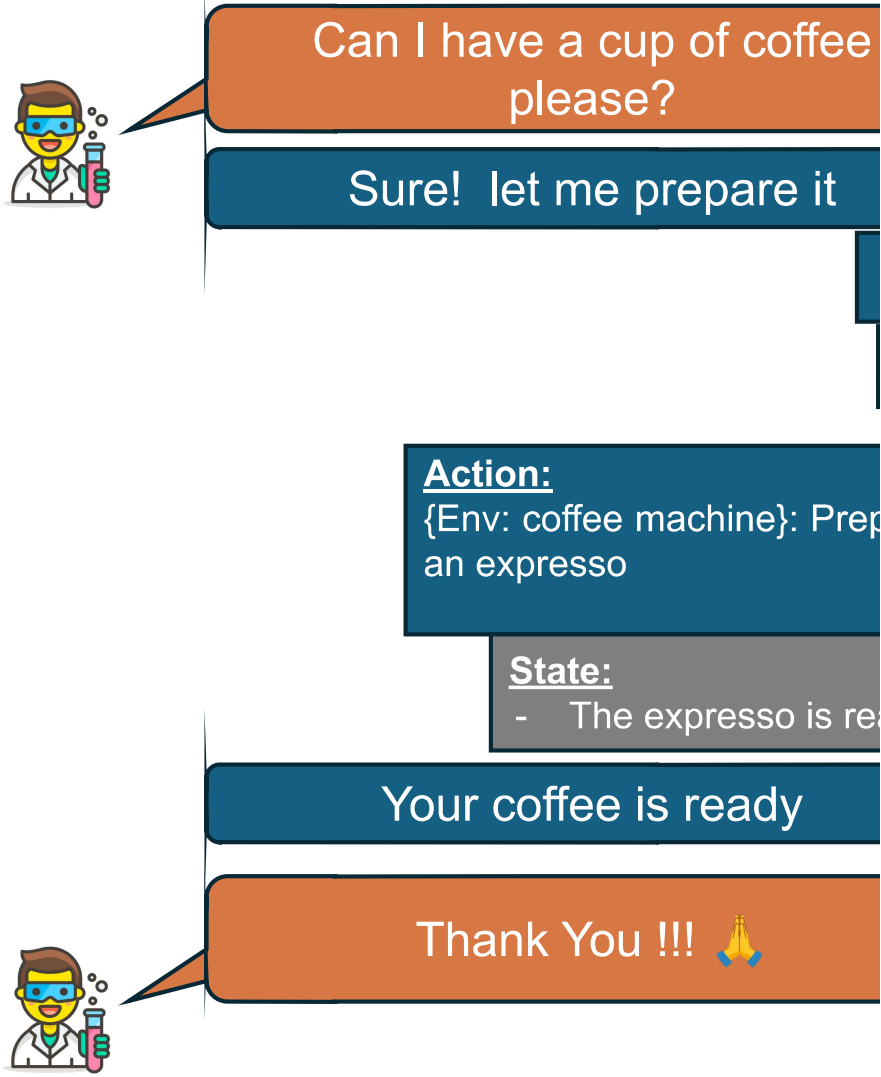
State:
- The coffee machine is connected
- The tank water is full

State:
- It is 15h30

Planning:
The coffee machine connected and ready prepare a coffee. It is 15 hence it is too late fo cappuccino. Let's go wit simple expresso

What is an AI Agent



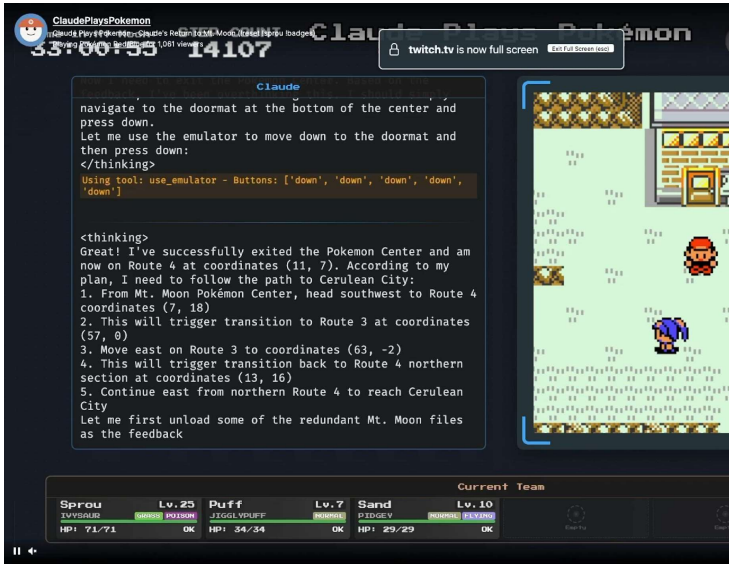


An artificial intelligence (AI) agent is a program that can interact with its environment, collect data, and use the data to perform self-determined tasks to meet predetermined goals.

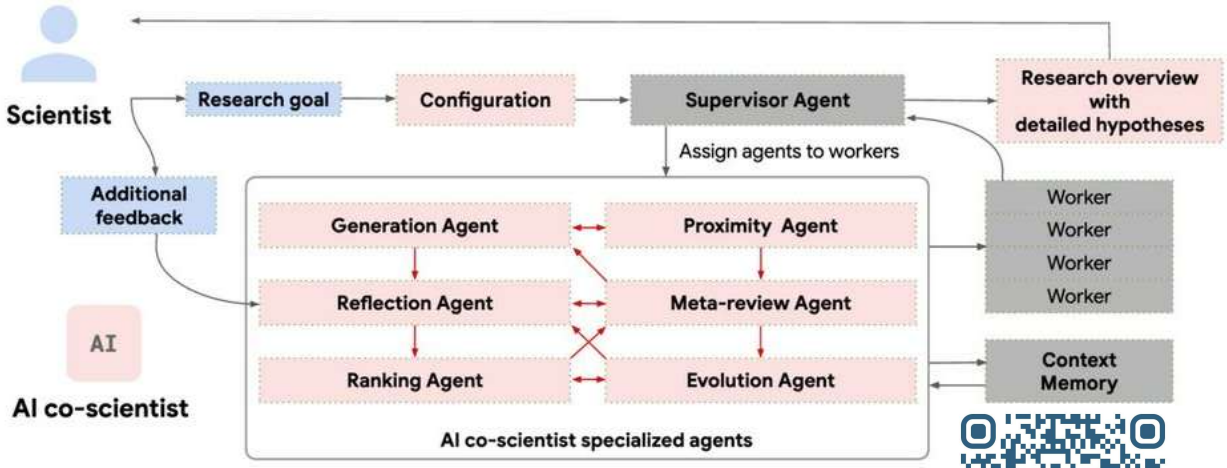
Humans set goals, but an AI agent independently chooses the best actions it needs to perform to achieve those goals.

Example of agents

Coding



Playing Pokemon

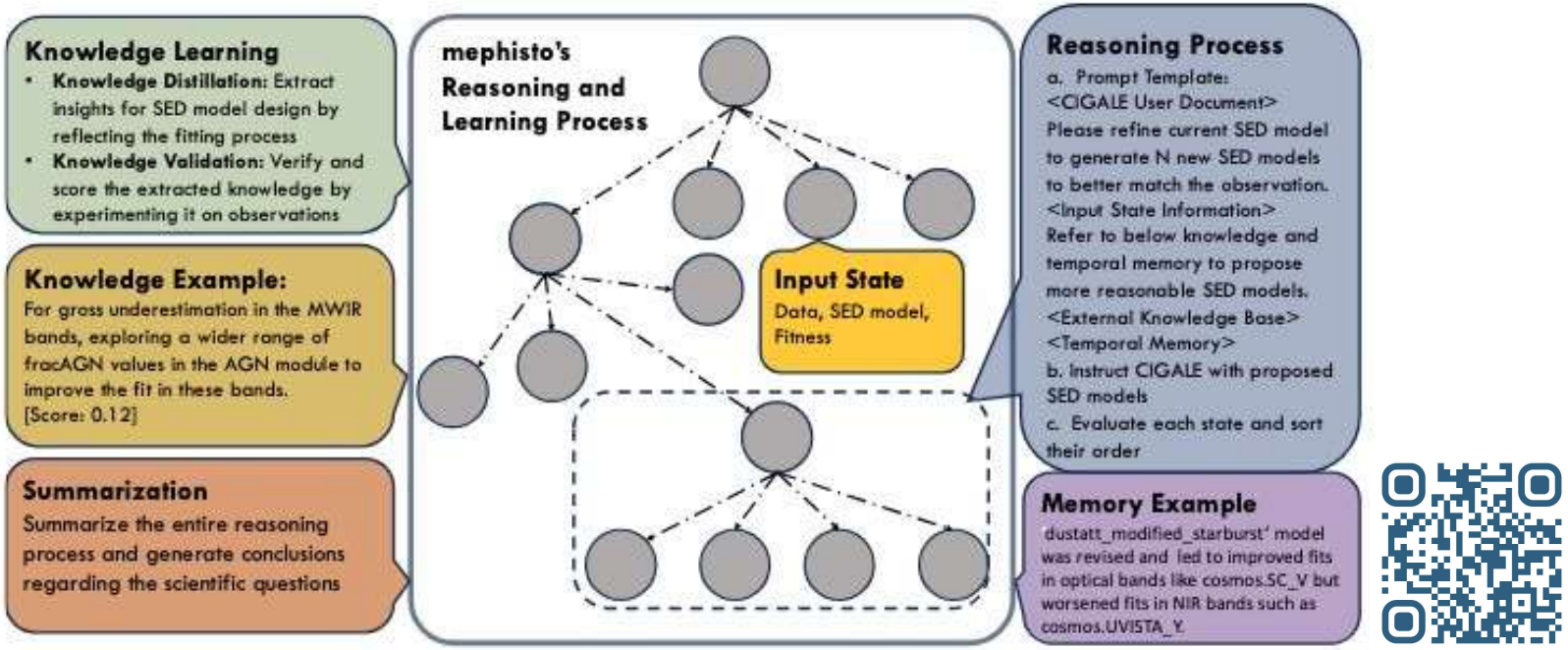


Co-scientist
(Google, Gottweis et al. 2025)

Data Science Gemin

Agent AI Astro

Mephisto, an AI agent ables to use CIGALE code to fit SED



Sun et al. 2024.

In 2040, AI agent will be use both for the processing of data and for the planning of observation (target decision, mode choice, opportunity decision…)

<CIGALE Documentation>
<Introduction to User Input>
Your task is to analyze user input and adhere to below rules a
provided CIGALE models to improve the fit quality of differe
the CIGALE module which should affect the fitting quality the
module, e.g., use another choice, adjust the parameter grid, ar

- Choice Selection: For each module in the model, se
  This decision should be based on the need to opt
  observational data.
- Parameter Grid Specification: For the selected cho
  parameter grid that the model will use to fit the data.
  - For discrete parameters, select grid values from
  - For continuous parameters, derive grid values wi
    a comprehensive exploration of parameter spac
- Mandatory Modules: The model configuration must i
  following modules:
  - sfh (Star Formation History): This module is es
    which a galaxy forms stars over time.
  - ssp (Simple Stellar Population): This module is
    collective properties of stars in a galaxy that for
    the same metallicity.
- For 'imf' parameter in 'bc03' and 'm2005', it should
  1 are accepteable, "imf": [1] are foridden
- For 'disk_type' parameter in 'fritz2006' and 'skirtor
  list, i.e., "disk_type": 0 are acceptable, "disk_type":

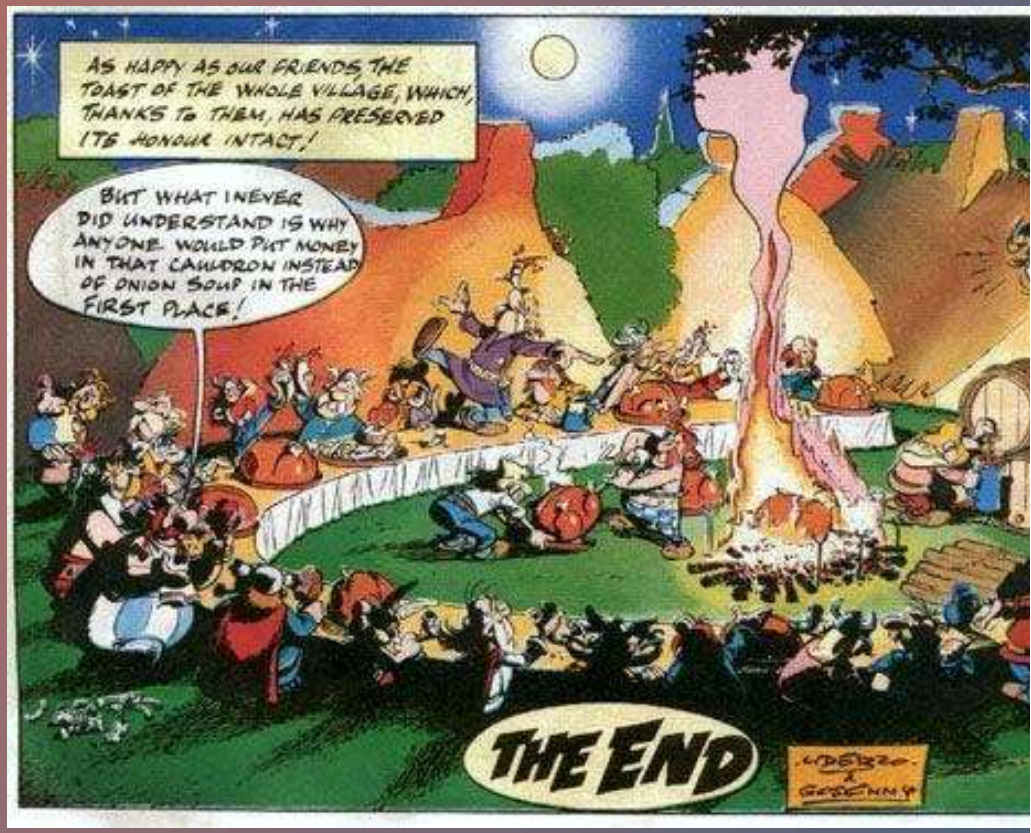
##CIGALE SED Knowledge##
To enhance the fit quality for different photometric bands in the
following addtional information when designing your CIGAL
%%KNOWLEDGE%%

##Temporary Memory##
Please take into account the following temporary memories t
superior fit quality in the updated CIGALE model:
%%MEMORY%%

##Expected Output Format##
Your output format should be structured as below list construc
able CIGALE model modifications:

```
[
    {
        "thinking": "thinking for CIGALE model
        "module": "module name for CIGALE model
        "name": "module choice name for CIGALE
        "parameters": [
            "parameter 1": ["parameter 1 grid h
            "parameter 2": ["parameter 2 grid h
            ...
            "parameter n": ["parameter n grid h
        ]
    },
    ...
]
```

Conclusion

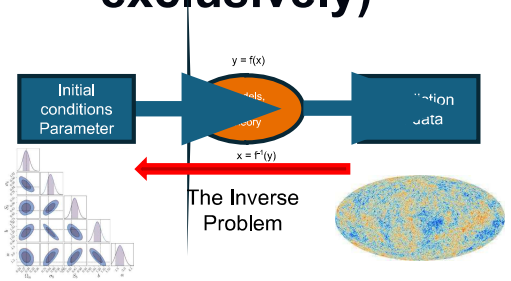


Asterix and Obelix and the cauldro

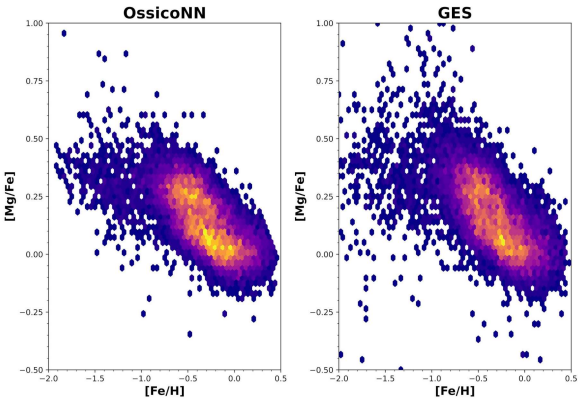
Conclusion

MALSP

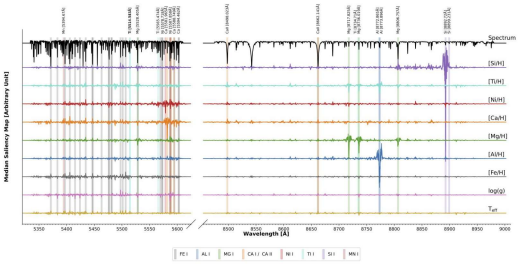
AI is important to tackle inverse problems (but not exclusively)

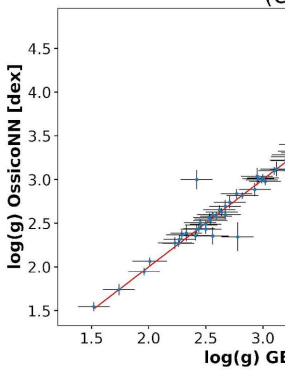


AI well reproduce Astrophysical relation

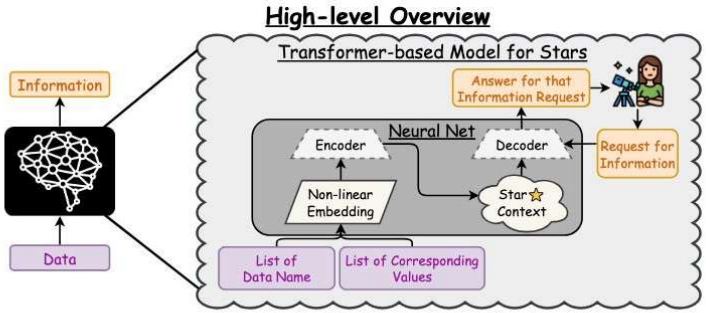


NOT A BLACK BOX: we get feature importance ar comprehensive uncertainty





In 2040, LLM will be used both for the processing of data and communication of results



By 2040, astrophysicists will primar planners, relying on AI agents to ex observational tasks and data ar

Agent

AI & cINN | Applied to stars | LLM | AI Agent | Co

Thanks !!!!