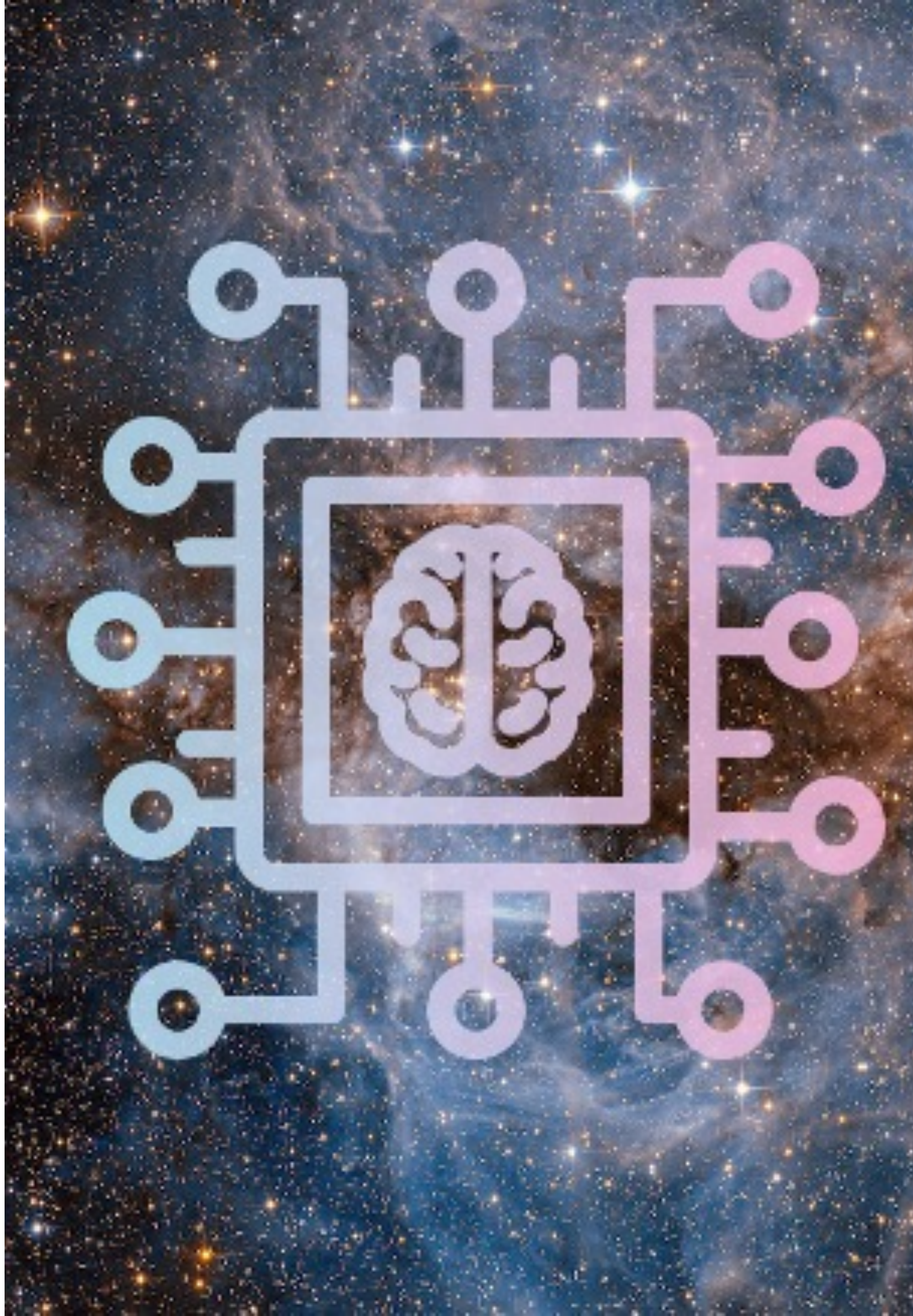
UNIVERSIDAD TECNICA FEDERICO SANTA MARIA
DEPARTAMENTO DE INFORMÁTICA
CCTVal CENTRO CIENTÍFICO TECNOLÓGICO DE VALPARAÍSO

Deep learning models for analyzing massive stars data: towards accurate and unbiased results

Raquel Pezoa (UTFSM), Michel Curé (UV), Ignacio Araya (UMayor), Felipe Ortiz (UV), José Montecinos (UTFSM), Daniela Turis (UV)

International Conference – Physics of Massive Stars
24 – 28 June 2024, Rio de Janeiro, Brazil
https://stel.asu.cas.cz/MassiveStars2024/



Outline

01

Machine Learning

Learning from data

02

Deep Learning

Artificial Neural Networks

03

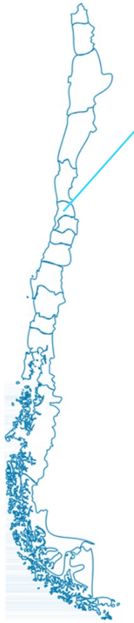
Probabilistic NN

Model's Confidence using
Bayesian Neural Networks

04

Applications

Classification of massive stars
and stellar parameters estimation



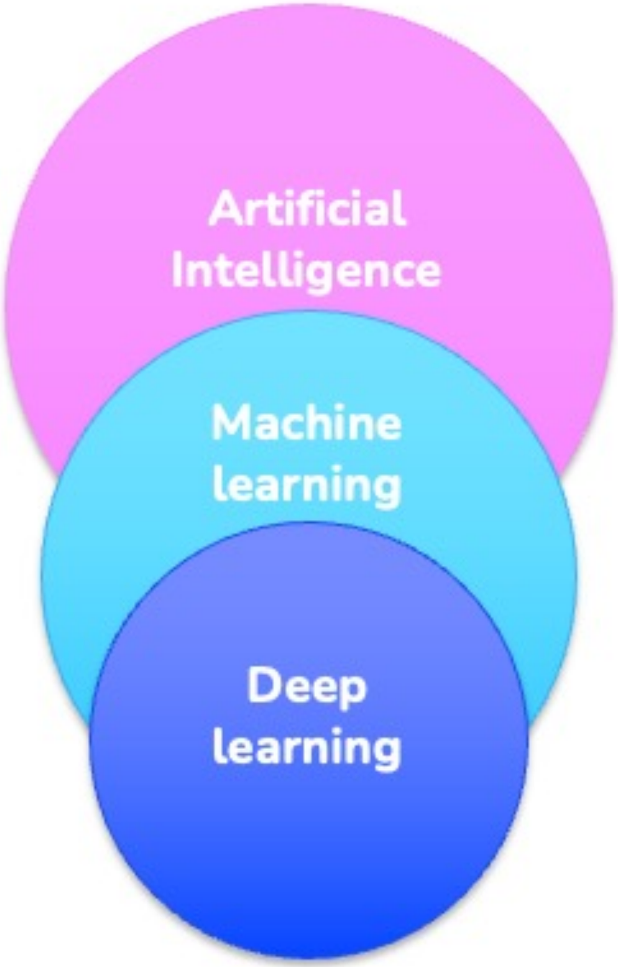
Mini-Bio

- My background is in computer science
- I am an academic at the Computer Science Department at Universidad Técnica Federico Santa Maria (UTFSM)
- My research is focused on machine learning and eXplainable Artificial intelligence, with applications on
  - particle physics (ATLAS/CERN), astrophysics, histopathological images

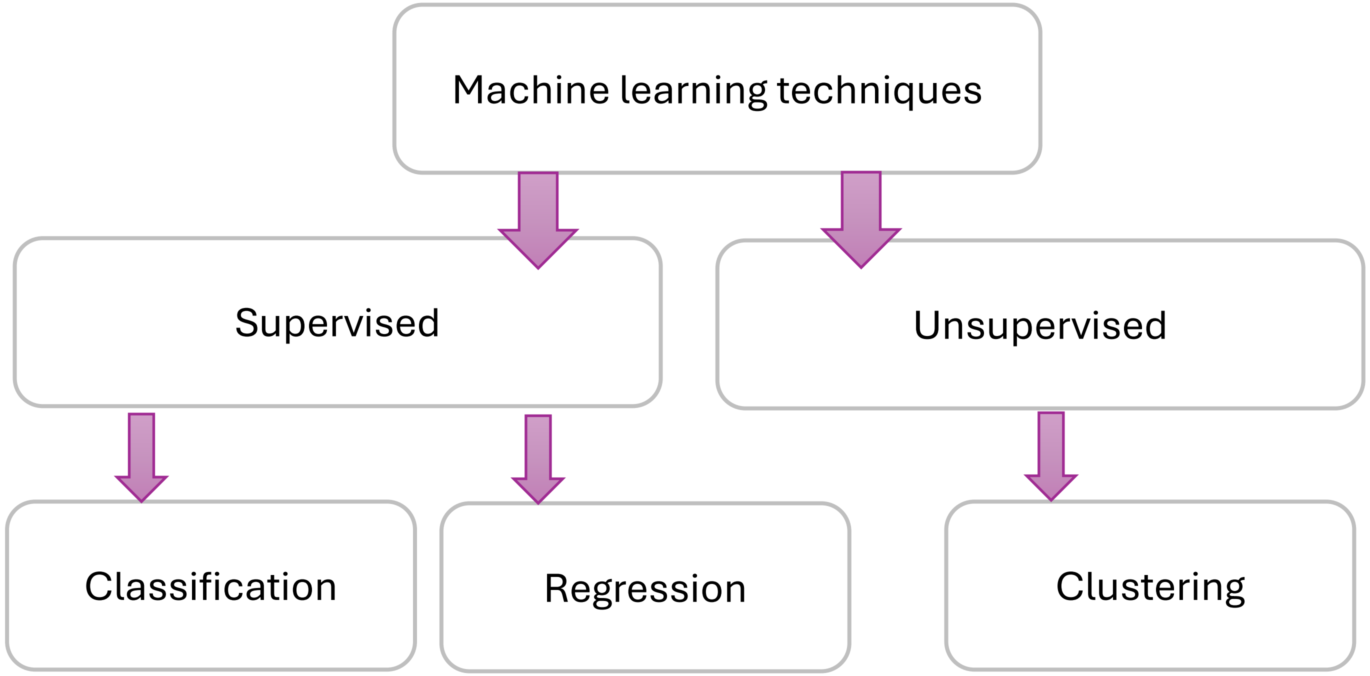
Machine Learning

- **Artificial Intelligence** is a computer science field that creates techniques that enables computers to perform tasks that typically require human intelligence.

- **Machine Learning** focuses on developing techniques that allow computers to "learn" from data, without explicitly being programmed.



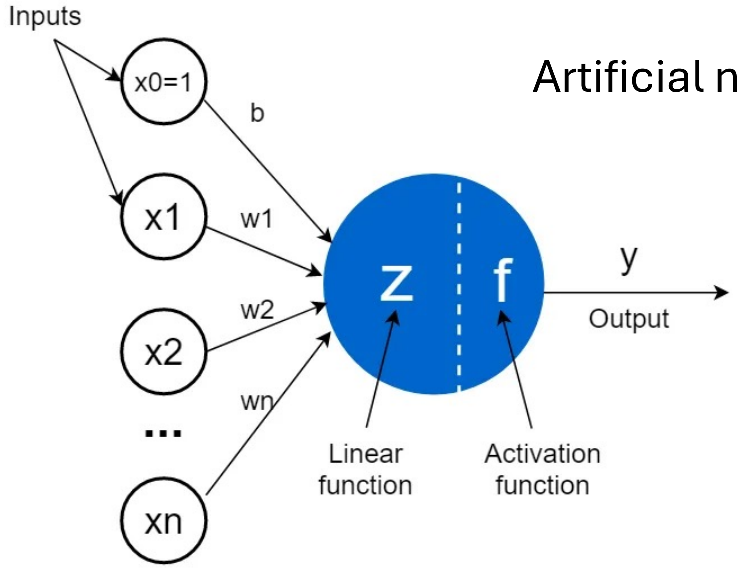
Machine Learning



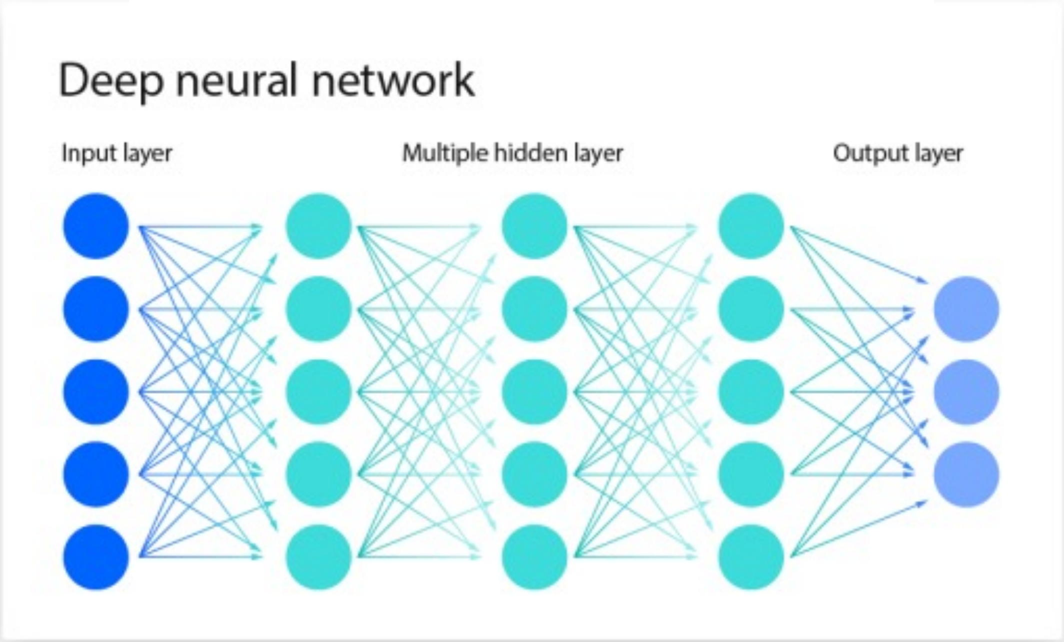
ML requires vast volumes of data
→ astrophysics is a field with an immense amount of data

Deep Learning

- A leading ML technique
- Artificial neural networks with a large number of hidden layers.
- Inspired by biological neural networks
- Outstanding performance in solving complex problems in science and industry.

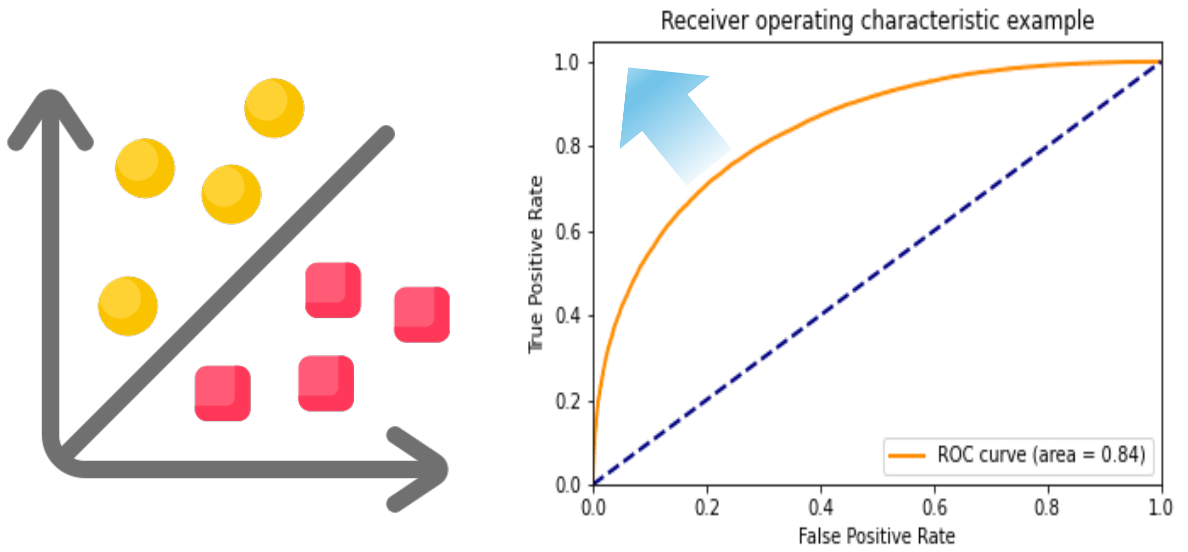


Artificial neuron

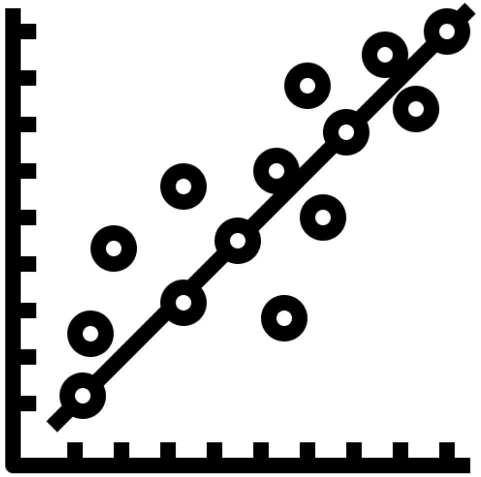


What do we expect of DL-based models?

- **Accurate** results → Models predict with high performance



Classification



$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \tilde{y}_i)^2$$

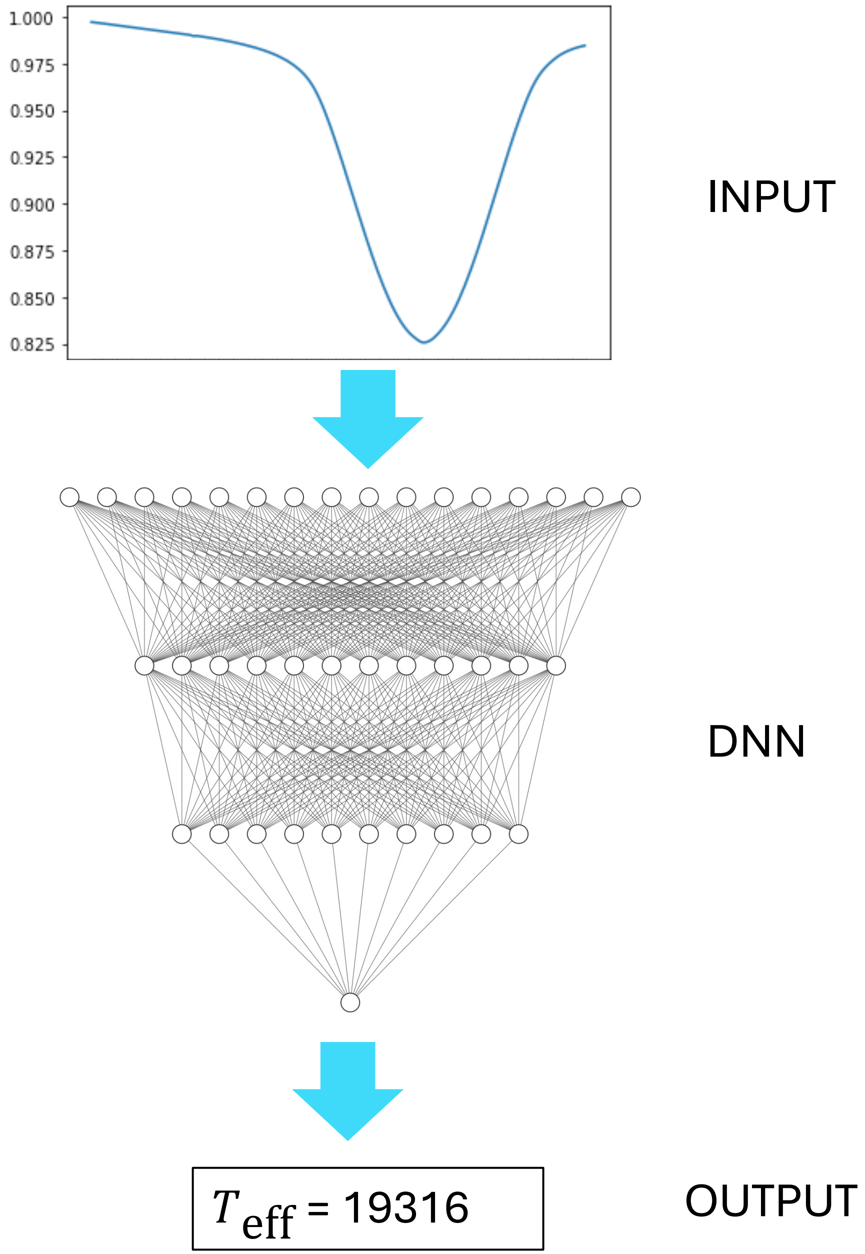
Regression

- But we also need to understand and trust in models' prediction

www.flaticon.com/free-icons

Model's confidence

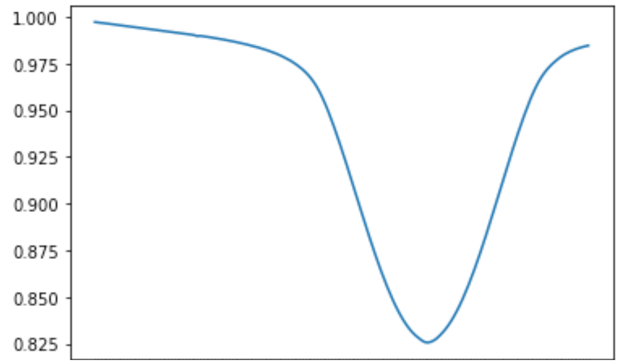
- In scientific fields, we'd like to fully trust DL models:
  - to make correct decisions,
  - to measure errors,
  - We want to know what the models know and what they don't know
- (Traditional) deep learning gives a point estimate in prediction



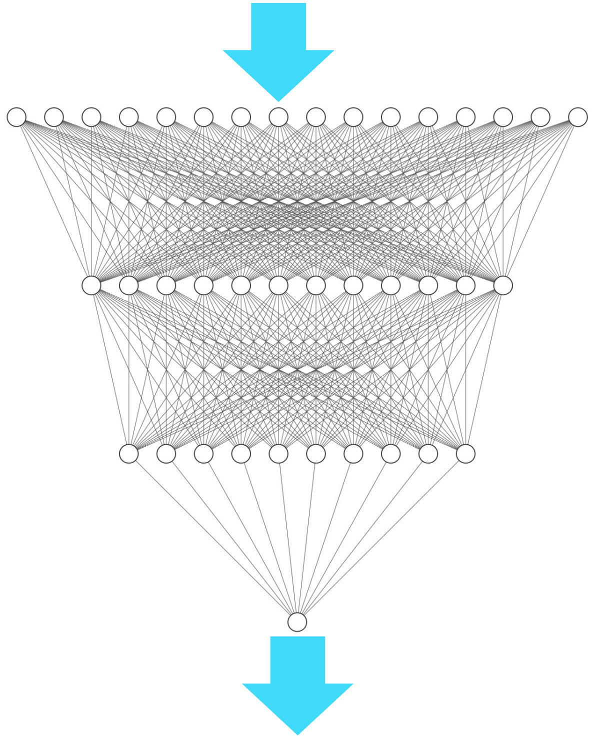
Model’s confidence

- We would like a DL model to give an answer to measure uncertainty and hence to
    - Make reliable decisions
    - Learn when you have limited, noisy and/or missing data
    - Get information about why a model failed

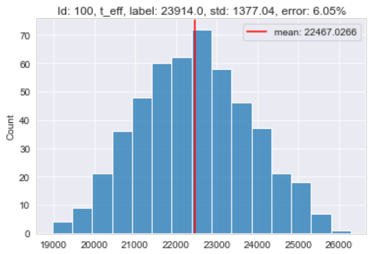
→ We need the DNN’s output to represent a predictive distribution



INPUT



DNN

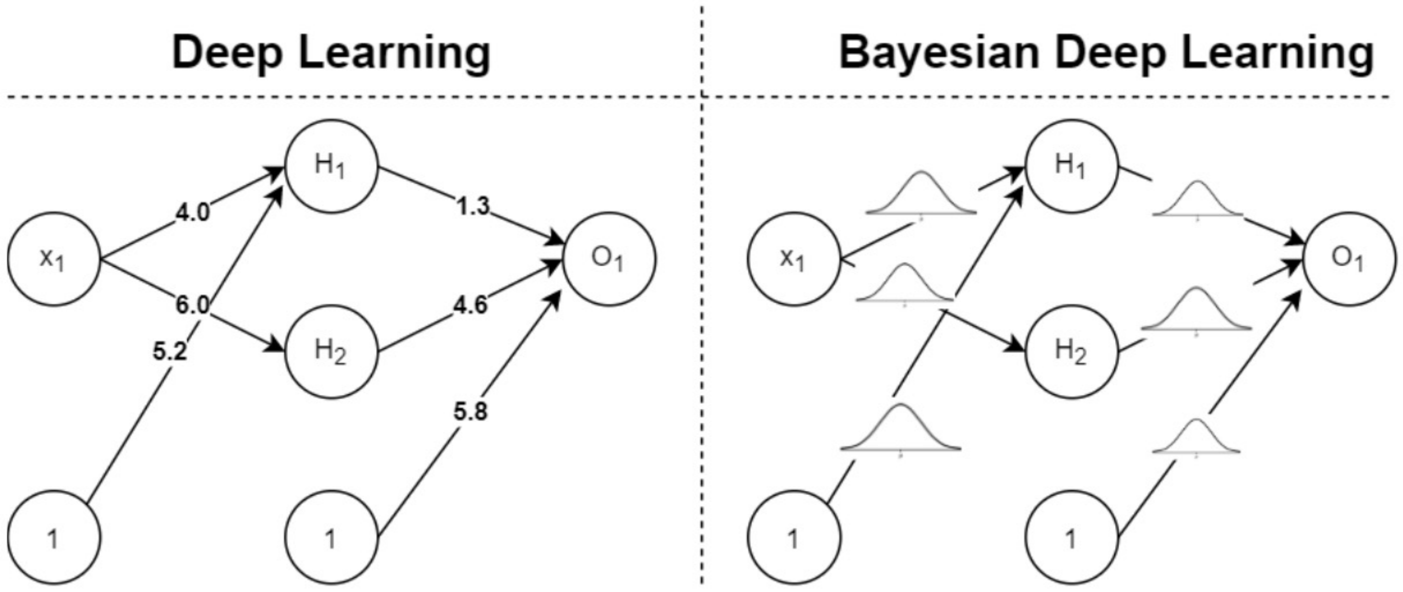


OUTPUT

DNN: Deep Neural Network

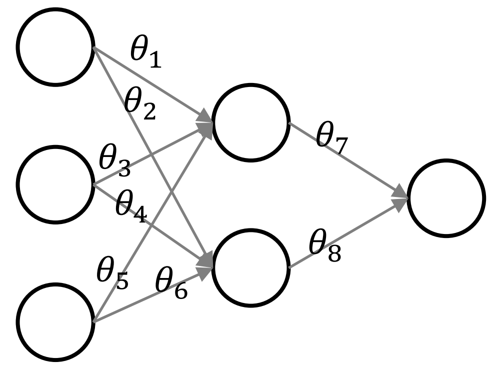
Probabilistic Neural Networks

- Probabilistic Neural Networks allows us to measure uncertainties of model's prediction.
- Point weights are "replaced" by a weight distribution



- How do we obtain weight distributions? → Bayesian Neural Networks

Bayesian Neural Networks



- DNN model defined by n weihgts → θ the vector of parameters.
- In the non-Bayesian case, is fixed, and in the Bayesian case, θ follows a distribution.
- The conditional posterior distribution $p(\theta|\,\mathcal{D})$ will allow us to estimate the uncertainty:

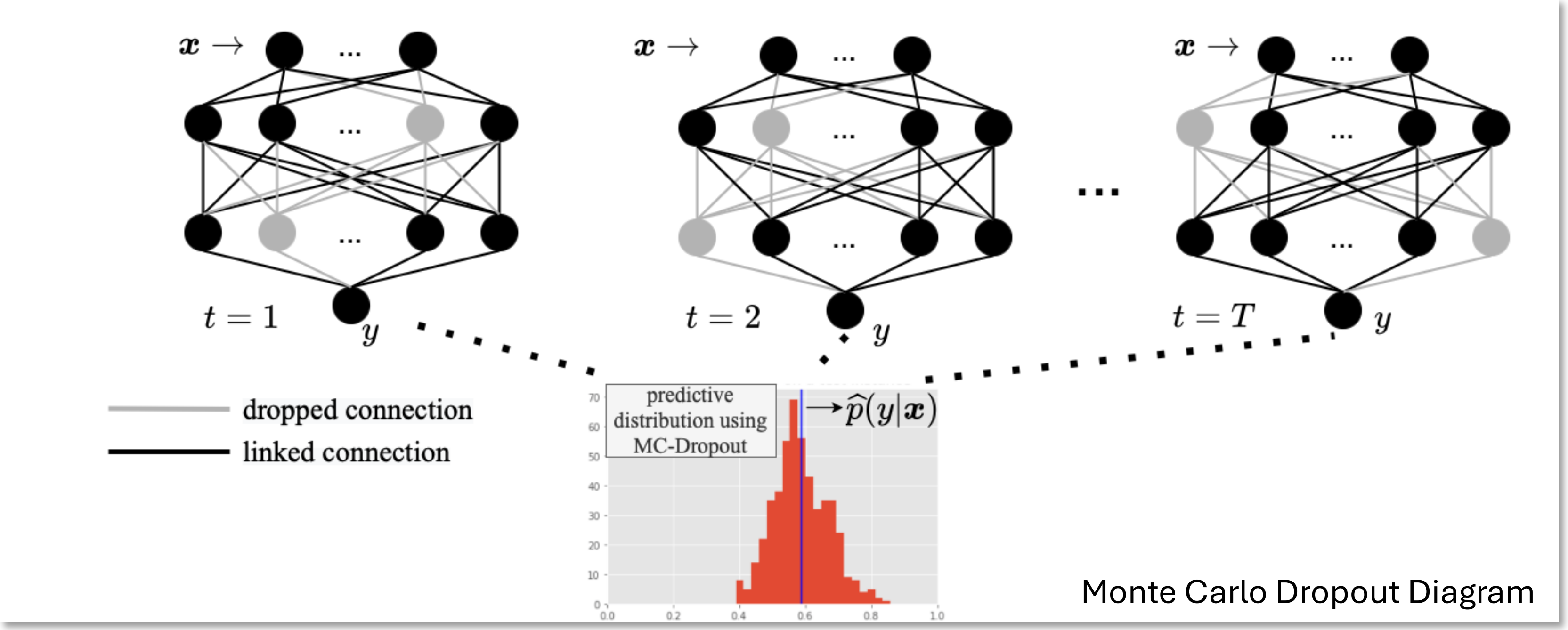
$$p(\theta|\mathcal{D}) = \frac{p(\mathcal{D}|\theta)\ p(\theta)}{p(\mathcal{D})}$$

Image source: Wikipedia

$\mathcal{D}$: the training data
$\mathcal{D} = \{\boldsymbol{x}_i, y_i\}_{i=1}^{N}, \boldsymbol{x} \in \mathbb{R}^d$
- Classification $y \in \{1, \dots, K\}$
- Regression $y \in \mathbb{R}$

How to obtain the weights distribution?

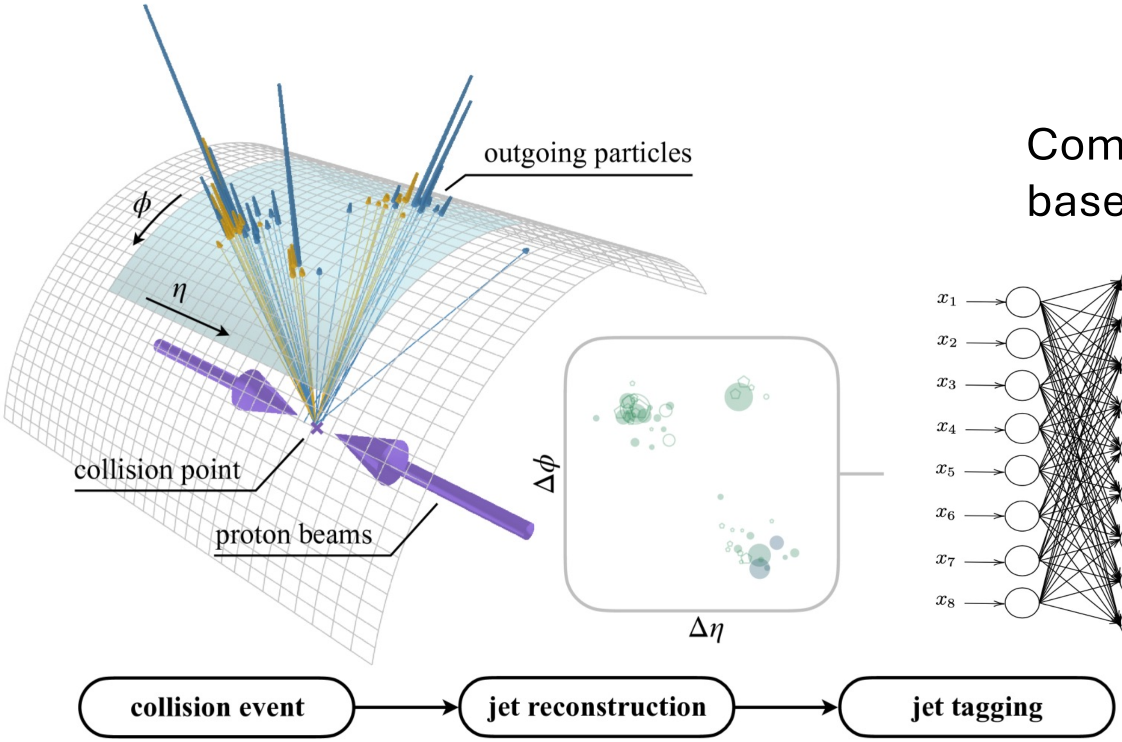
- Bayesian inference is computationally intractable
- Methods to approximate $p(\theta|\, \mathcal{D})$, like Variational Inference, Monte Carlo Dropout, Deep Ensemble, etc.



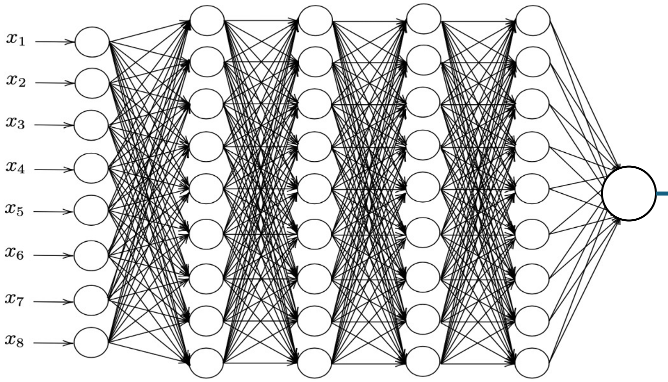
Monte Carlo Dropout Diagram

- Dropout during the the training and inference stages
- We obtain a predictive distribution – instead of a single output – and hence, we can estimate uncertainties

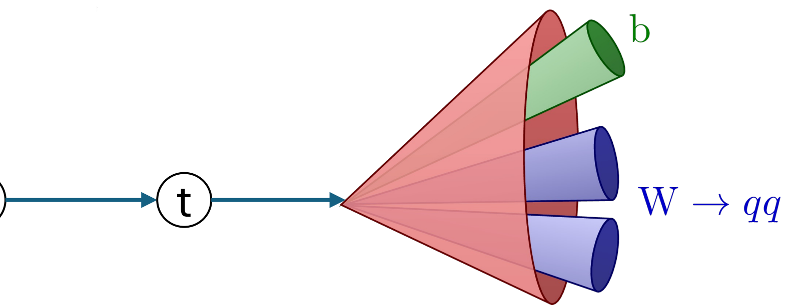
Explaining predictions in Particle Physics

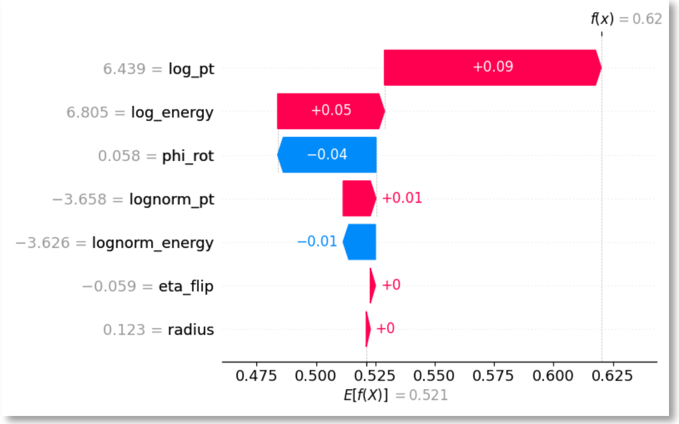


Complex DNN-based jet taggers



DNN output

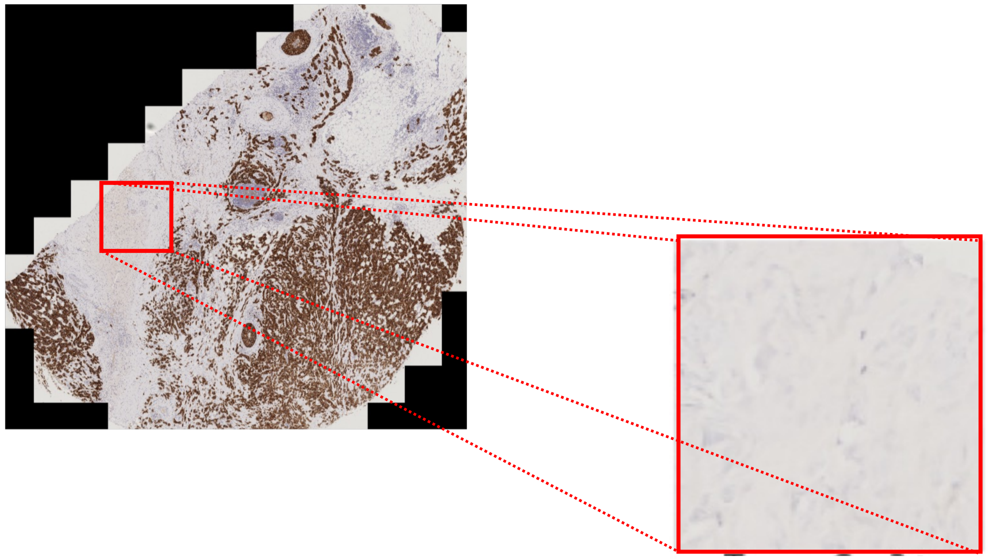




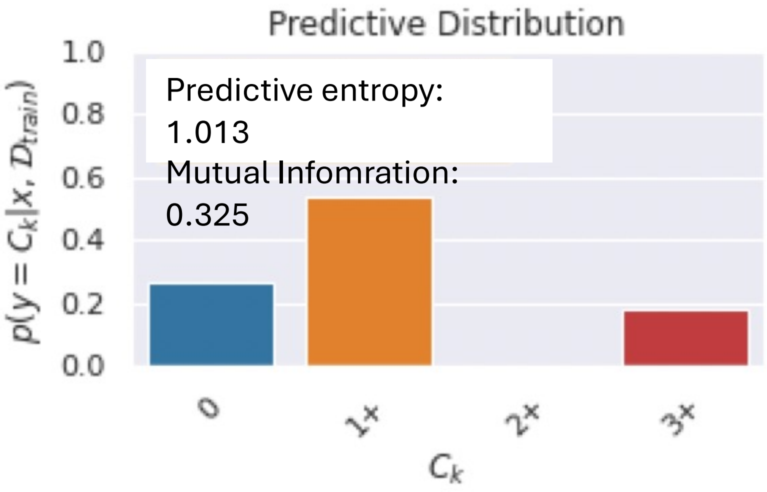
SHapley **A**dditive ex**P**lanation (SHAP)

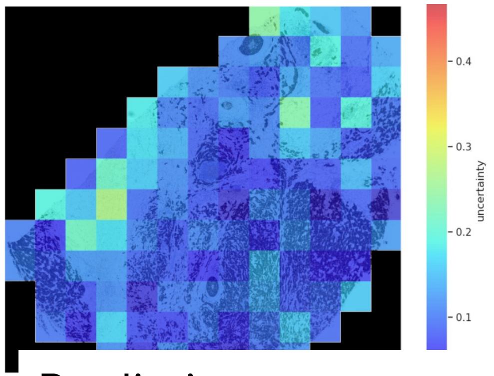
Pezoa, R., Salinas, L., & Torres, C. (2023, February). Explainability of High Energy Physics events classification using SHAP. In Journal of Physics: Conference Series (Vol. 2438, No. 1, p. 012082). IOP Publishing.

Measuring Uncertainties in Biomedical Images

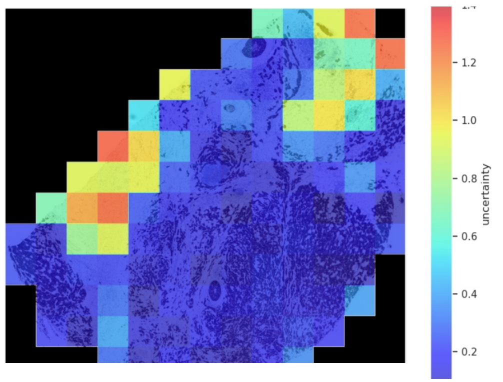


Predicción: 1+
Etiqueta: 3+





Predictive entropy



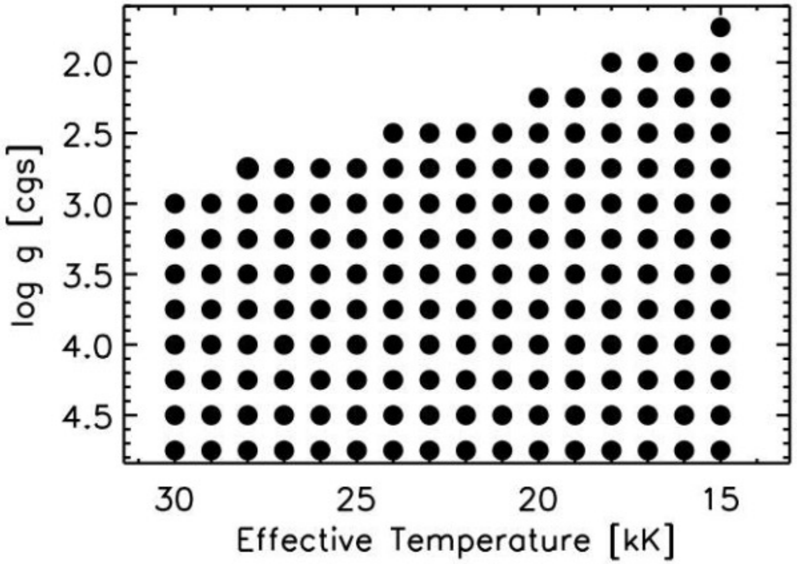
Mutual Information

Bórquez, S., Pezoa, R., Salinas, L., & Torres, C. E. (2023). Uncertainty estimation in the classification of histopathological images with HER2 overexpression using Monte Carlo Dropout. Biomedical Signal Processing and Control, 85, 104864.

Estimating stellar parameters using Deep Ensemble

- We train an ensemble of M DNNs with different random initializations.
- Data →Tlusty, Synspec and Rotin, using synthetic models

B stars, metallicity 1



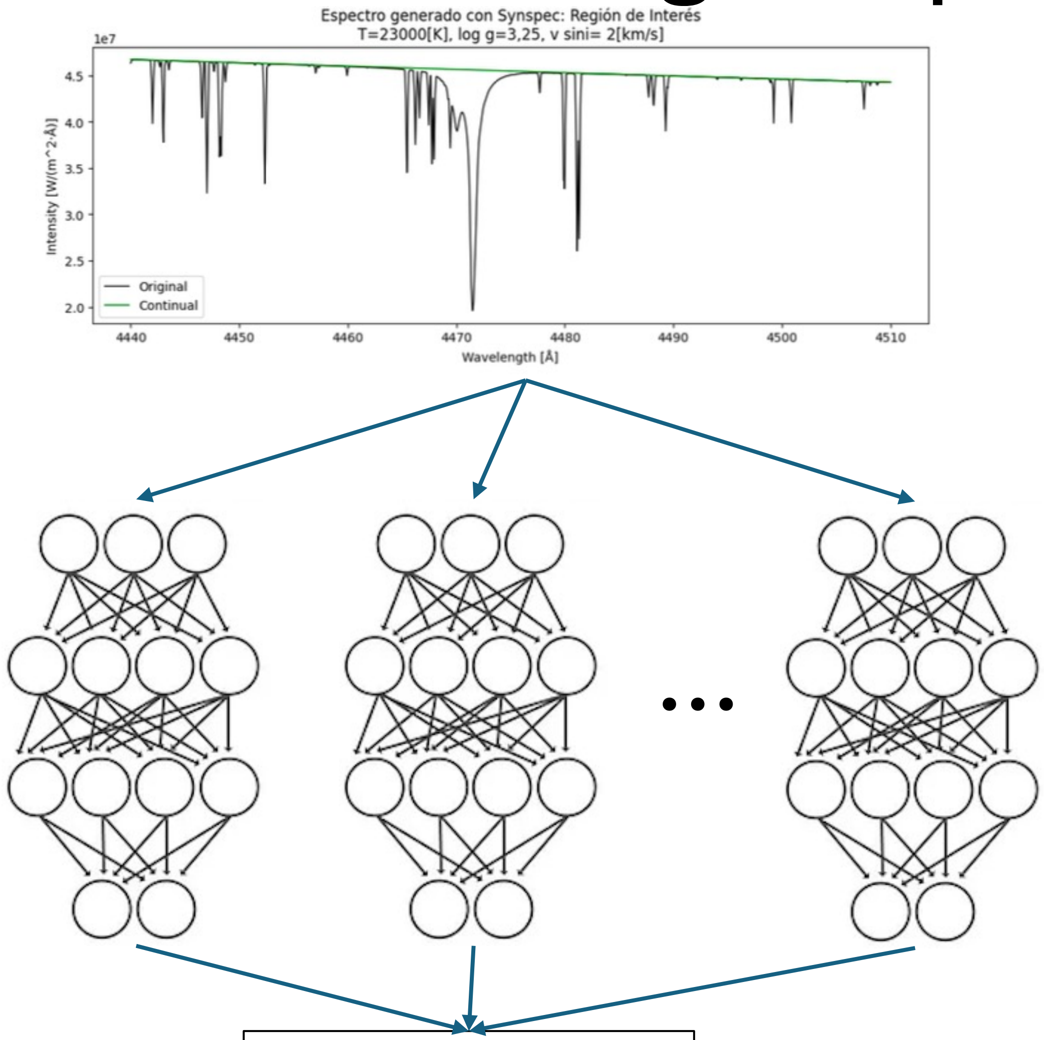
Model	Tlusty/step
T_{eff}[K]	15000-30000/1000
log g[1]	1.75-4.75/0.25
Solar metallicity	$Z/Z_0 = 1{,}0$
v_{sini}[km/s]	3-450/3

Train (60 %)	**Validation** (20 %)	**Test** (20 %)	**Total** (100 %)
14670	4890	4890	24450

Undergraduate thesis, José Montecinos (UTFSM)
Supervisors: Raquel Pezoa (UTFSM) and Michel Curé (UV)

Estimating stellar parameters using Deep Ensemble

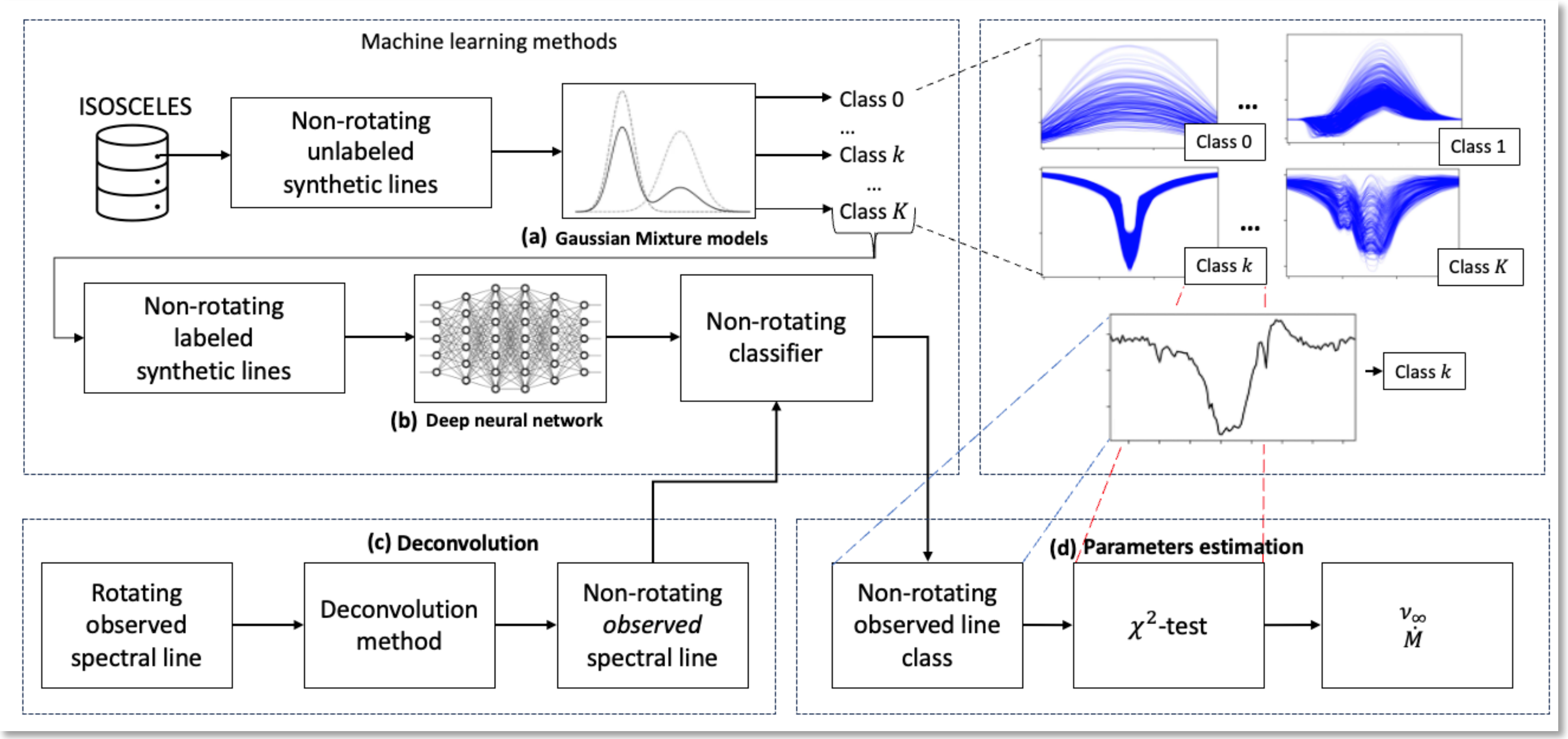
- We train an ensemble of M DNNs with different random initializations.



T_{eff} = 19316 ± 398
log g = 3.1 $\pm$ 0.1
$v \sin i$ = 302 $\pm$ 5

Estimation of wind parameters using GMM and DNN

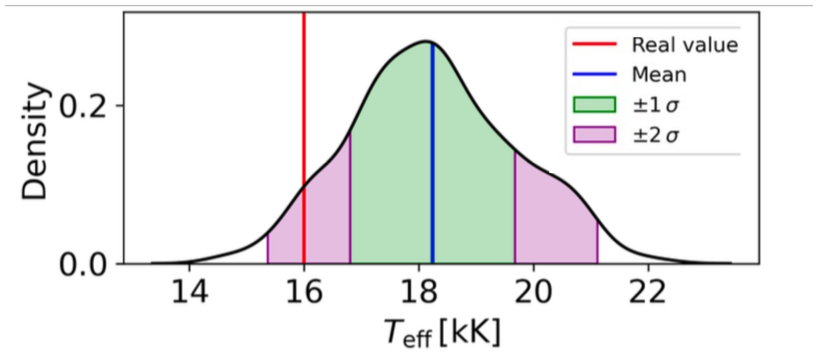
- Hα lines from δ-slows models from ISOSCELES database.
- Unsupervised and supervised ML techniques



Tuesday, June 25, 10:10
Ignacio Araya
"Stellar Atmosphere and Hydrodynamic Modeling with ISOSCELES v2.0<"

BDL – Monte Carlo Dropout

ISOSCELES

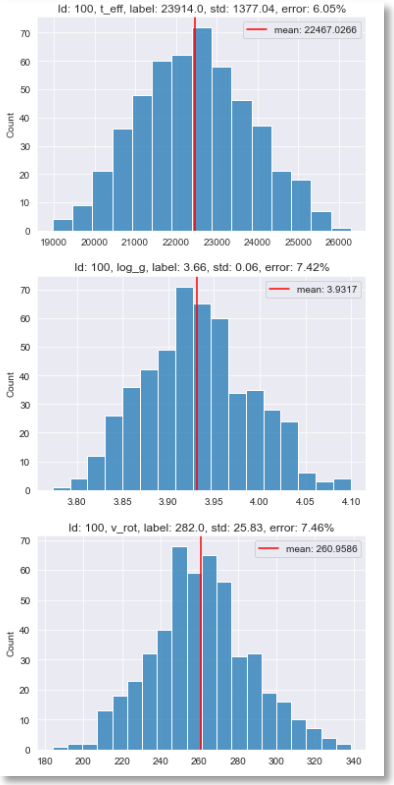


ZPEKTR data

```python
df = pd.read_csv("df_ZPEKTR_limb_lineal.csv")
```

```python
df
```

	4460.1	4460.2	4460.3	4460.4	4460.5	4460.6	4460.7	446(
0	0.996583	0.996529	0.996452	0.996362	0.996264	0.996160	0.996052	0.9959
1	0.996234	0.996208	0.996180	0.996141	0.996108	0.996074	0.996027	0.9959
2	0.997128	0.996958	0.996770	0.996573	0.996368	0.996163	0.995957	0.9957
3	0.993683	0.993541	0.993413	0.993277	0.993126	0.992986	0.992836	0.9926
4	0.995026	0.994887	0.994749	0.994607	0.994464	0.994324	0.994183	0.9940
...	...	...	...	...	...	...	...	
10686	0.997252	0.997177	0.997096	0.997015	0.996925	0.996834	0.996746	0.9966
10687	0.996881	0.996865	0.996835	0.996777	0.996702	0.996621	0.996535	0.9964



Poster session: Felipe Ortiz, "Classification of spectral Hα lines from massive stars using machine learning methods"

Poster session: Daniela Turis, "Determining Inclination Angles in Rapidly Rotating Massive Stars via Spectroscopic Analysis"

Closing Remarks

- DL-based models are an important tool for analyzing large samples of stellar spectra → increasing efficiency and accuracy.
- High accuracy is not enough → Bayesian Neural Networks is a powerful to estimate uncertainties and generate unbiased results
- In addition to developing DL and BDL-based methods, we are also working on the systematically analyzing of the observed data noise and DNN/BLD models.
- What's next?
  - Improving BDL-based algorithm to estimate wind and stellar parameters, and to include XAI techniques.

UNIVERSIDAD TECNICA FEDERICO SANTA MARIA

EX UMBRA IN SOLEM

DEPARTAMENTO DE INFORMÁTICA

CCTVal CENTRO CIENTÍFICO TECNOLÓGICO DE VALPARAÍSO

Deep learning models for analyzing massive stars data: towards accurate and unbiased results

GRACIAS!

OBRIGADA!

THANKS!

This work is supported by FONDECYT Postdoc 3190740, ANID PIA/APOYO AFB230003, FONDECYT Projects 1230131 and 11190147. Powered@NLHPC: This research was partially supported by the NLHPC's supercomputing infrastructure (ECM-02).