Constructing Impactful Machine Learning Research for Astronomy

DANIELA HUPPENKOTHEN
MICHELLE NTAMPAKA

Matthew Ho, Morgan Fouesneau, Brian Nord, J.E.G. Peek, Mike Walmsley, John F. Wu, C. Avestruz, Tobias Buck, Massimo Brescia, Douglas P. Finkbeiner, Andy D. Goulding, T. Kacprzak, Peter Melchior, Mario Pasquato, Nesar Ramachandra, Yuan-Sen Ting, Glenn van de Ven, Soledad Villar, V.A. Villar, Elad Zinger

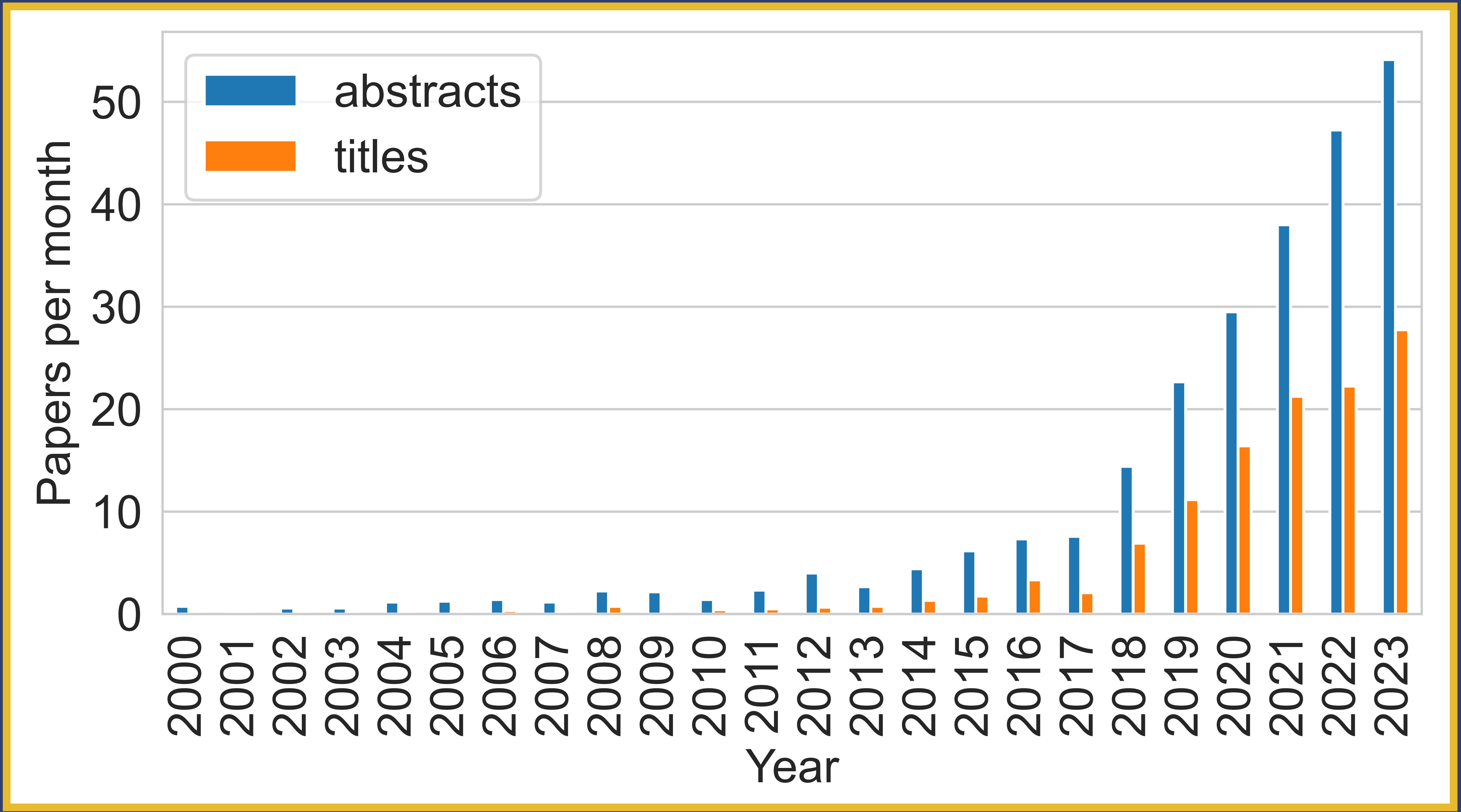
arXiv:2310.12528

SRON Netherlands Institute for Space Research

d.huppenkothen@sron.nl

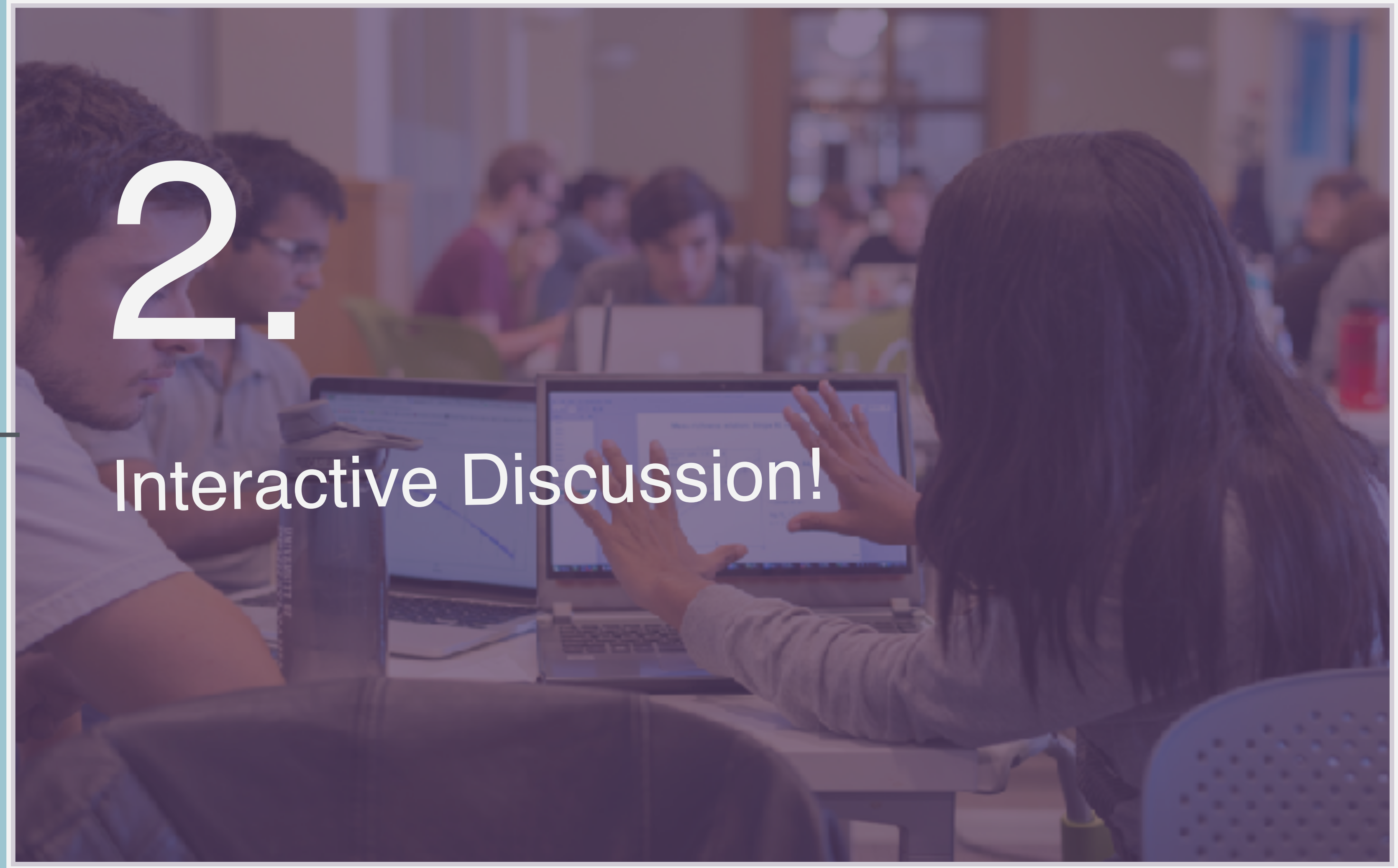

Why?

Machine Learning Papers in Astronomy



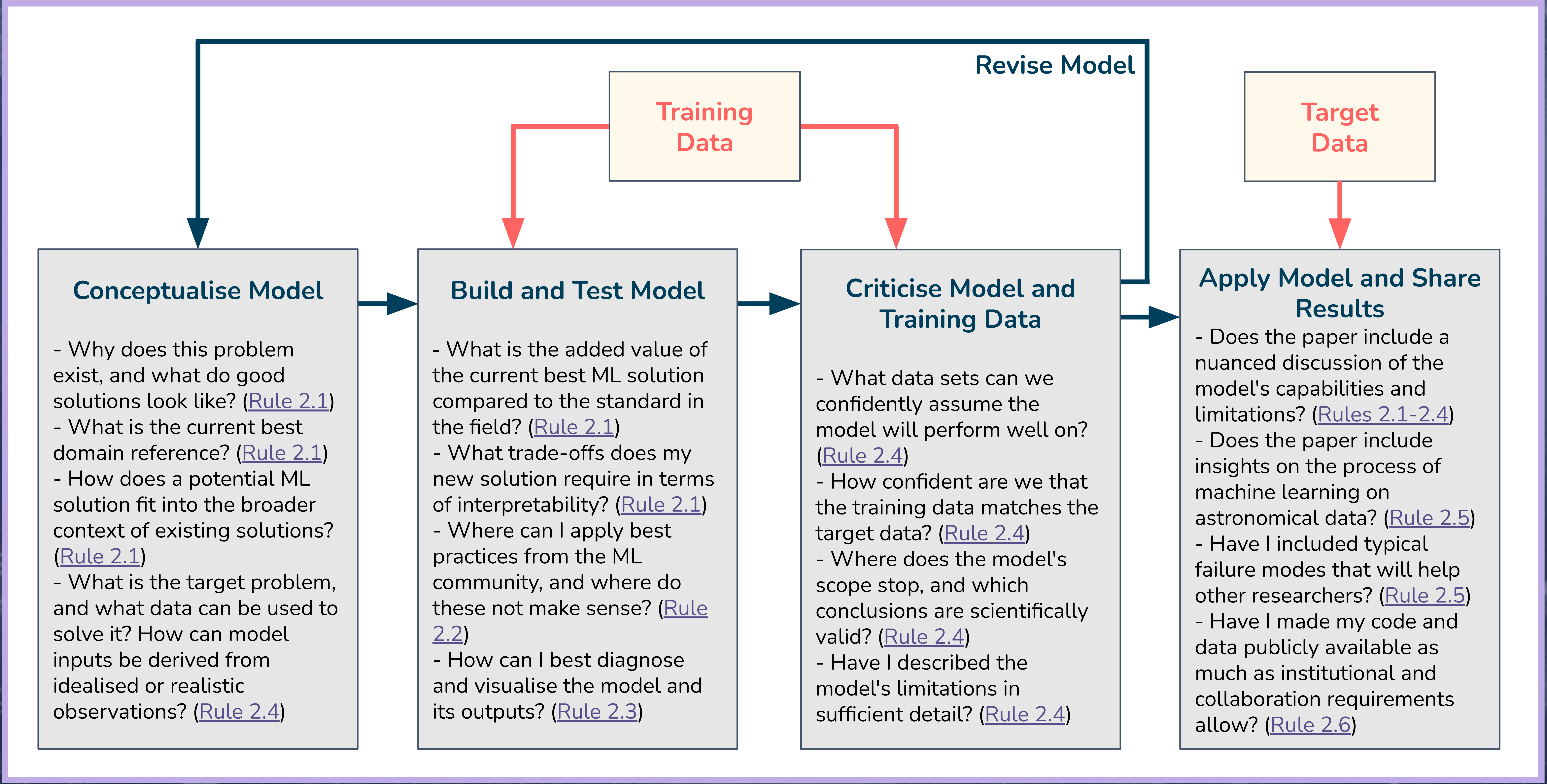
Key challenges: trust, robustness, interpretability


1.
6 Rules for Constructing Impactful Machine Learning in Astronomy


2.
Interactive Discussion!

Today

Box Loop for Machine Learning in Astronomy



Adapted from: Blei (2014)

Did you write a paper that showcases a rule? Do you know someone who did? Please send it to me!

d.huppenkothen@sron.nl

Rule 1: Compare against a domain reference and put result into larger context

Key reasons for implementing machine learning models in astronomy

Key reasons for implementing machine learning models in astronomy

1. Decrease the computational time and costs associated with scalability to larger data sets and samples

Key reasons for implementing machine learning models in astronomy

1. Decrease the computational time and costs associated with scalability to larger data sets and samples
2. Improve the robustness or precision of solutions to astrophysical questions

Key reasons for implementing machine learning models in astronomy

1. Decrease the computational time and costs associated with scalability to larger data sets and samples
2. Improve the robustness or precision of solutions to astrophysical questions
3. Facilitate model sharing and automation to make collaboration easier

Key reasons for implementing machine learning models in astronomy

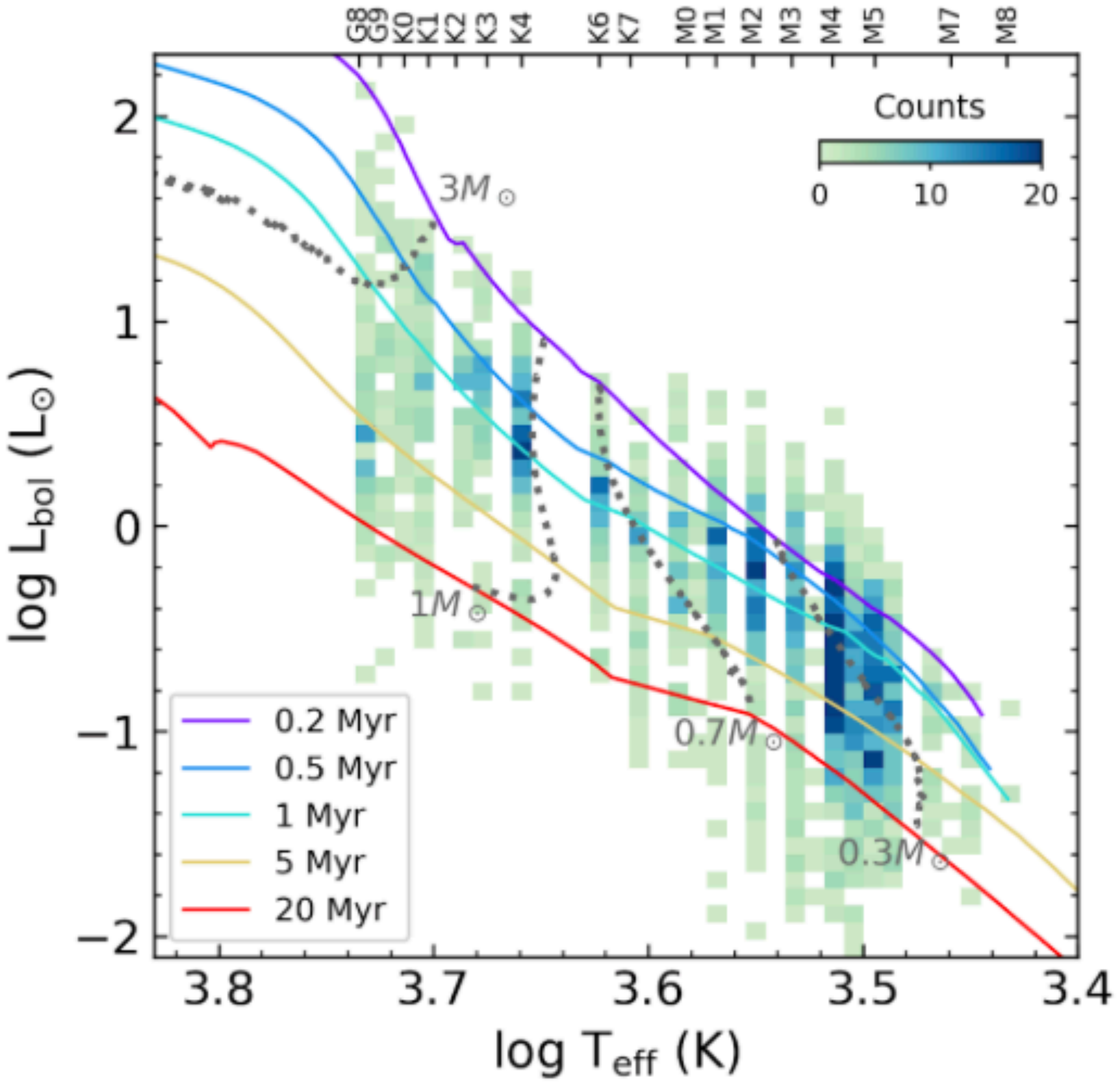
1. Decrease the computational time and costs associated with scalability to larger data sets and samples
2. Improve the robustness or precision of solutions to astrophysical questions
3. Facilitate model sharing and automation to make collaboration easier
4. Implement machine learning models for problems where no other solution exists

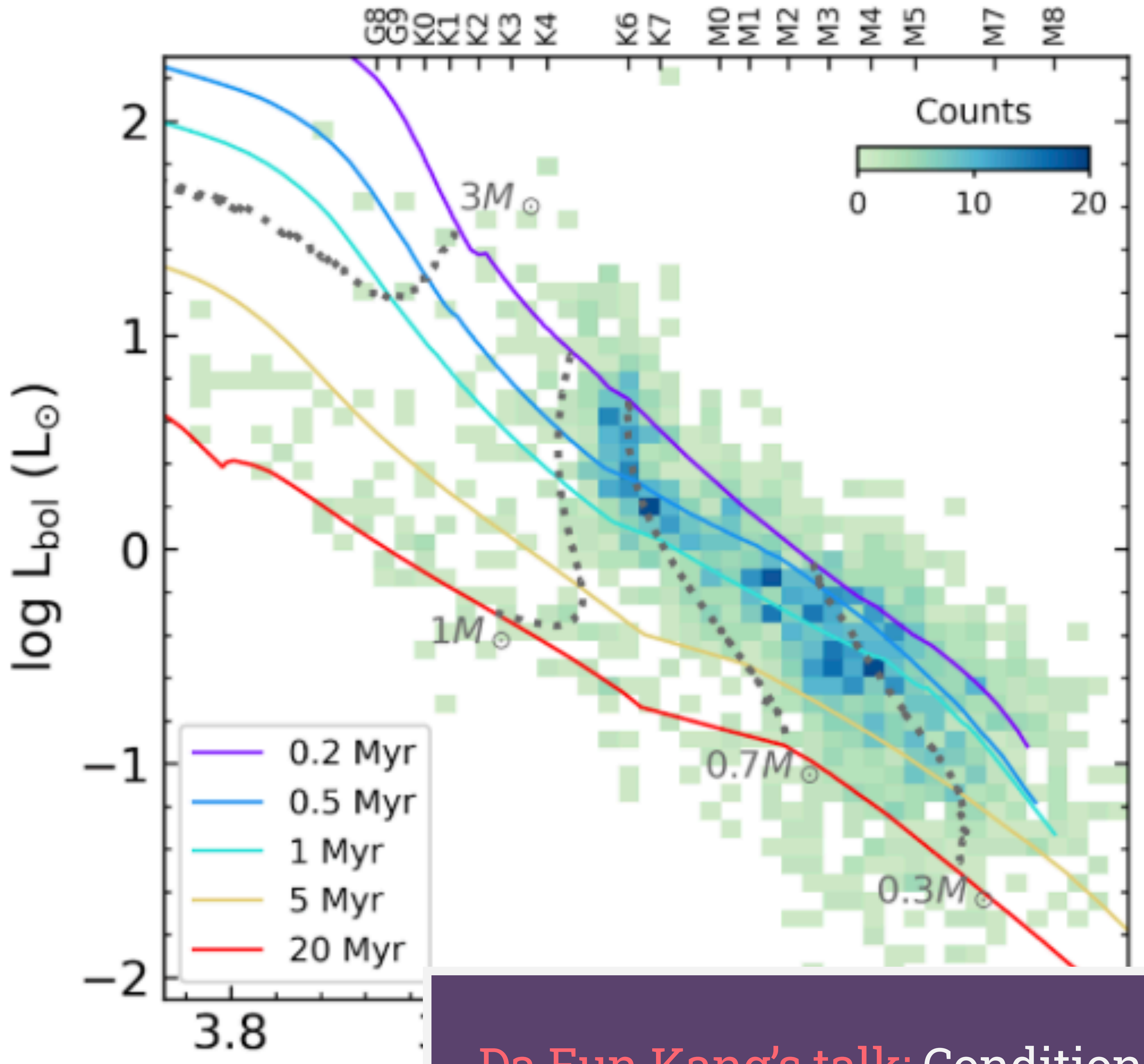
Key reasons for implementing machine learning models in astronomy

1. Decrease the computational time and costs associated with scalability to larger data sets and samples
2. Improve the robustness or precision of solutions to astrophysical questions
3. Facilitate model sharing and automation to make collaboration easier
4. Implement machine learning models for problems where no other solution exists

What value is added by implementing a machine learning model, compared to the trade-offs (e.g. interpretability, biases in population inference)

cINN (MAP) vs. Template fitting (lit)





Bolometric luminosity

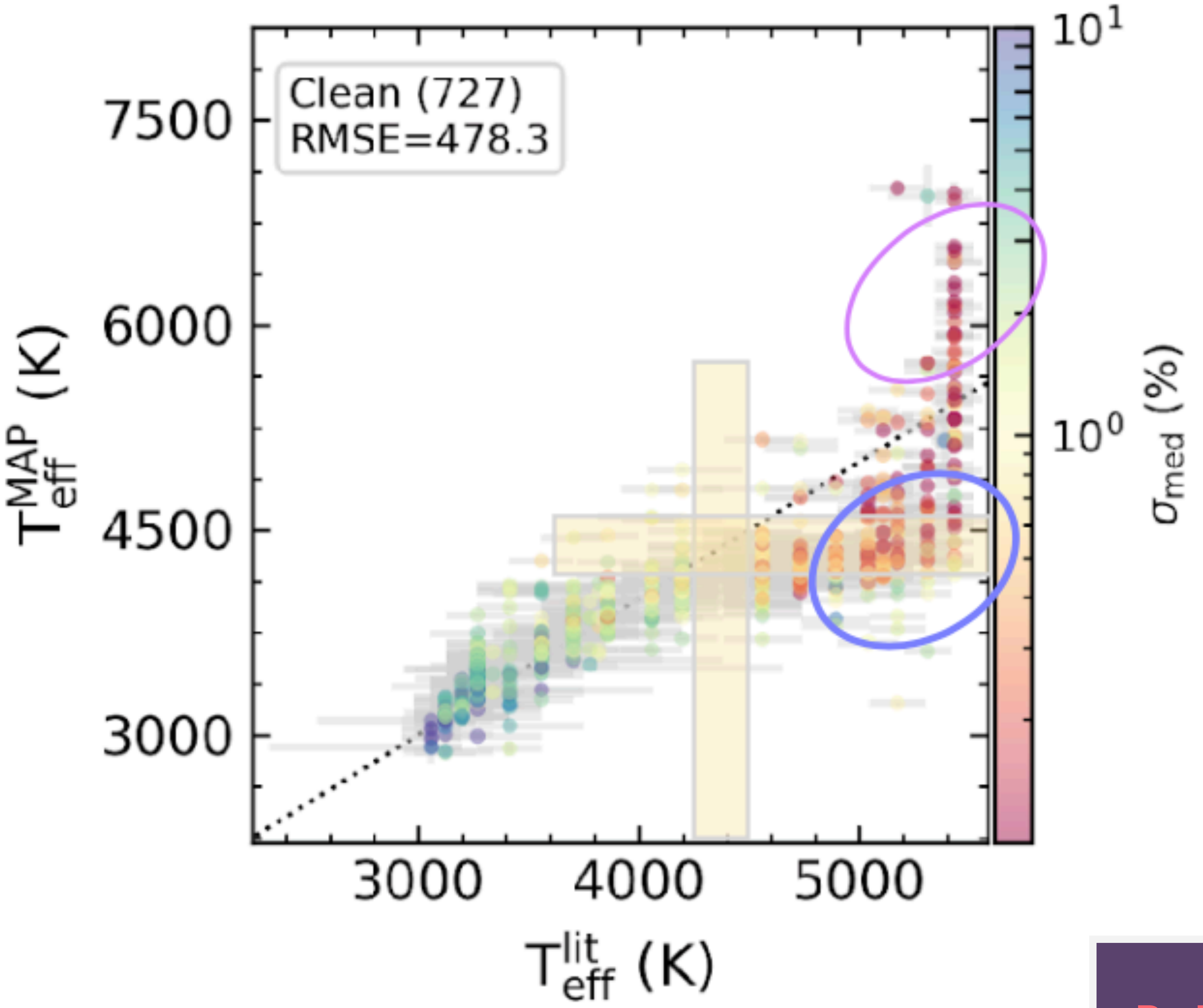
- J-band mag (HAWK-I, VISTA)
- extinction (A_V)
- Colors and bolometric correction = f(T_{eff})

- Average stellar age: 0.7Myr
- Two stellar populations
  - Young: ~ 0.7Myr

Da Eun Kang's talk: Conditional Invertible Neural Networks for enhanced analysis of young, low-mass stars in Trumpler 14

cINN (MAP) vs. Template fitting (lit)

cINN



Literature

Templates used in literature: M9.5 – G8.0

- No templates higher than G8 (5430K)
- No templates between K6 and K4 (ΔT ~360K)

cINN training range: 2600K – 7000K

Da Eun Kang's talk: Conditional Invertible Neural Networks for enhanced analysis of young, low-mass stars in Trumpler 14

Rule 2: Adopt best practices from the ML community

Use best practices in machine learning to establish trust in results

Use best practices in machine learning to establish trust in results

1. Designing and preparing data sets

Use best practices in machine learning to establish trust in results

1. Designing and preparing data sets
2. Choosing algoriothms

Use best practices in machine learning to establish trust in results

1. Designing and preparing data sets
2. Choosing algoriothms
3. Choosing evaluation metrics

Use best practices in machine learning to establish trust in results

1. Designing and preparing data sets
2. Choosing algoriothms
3. Choosing evaluation metrics
4. Exploring and reporting outputs

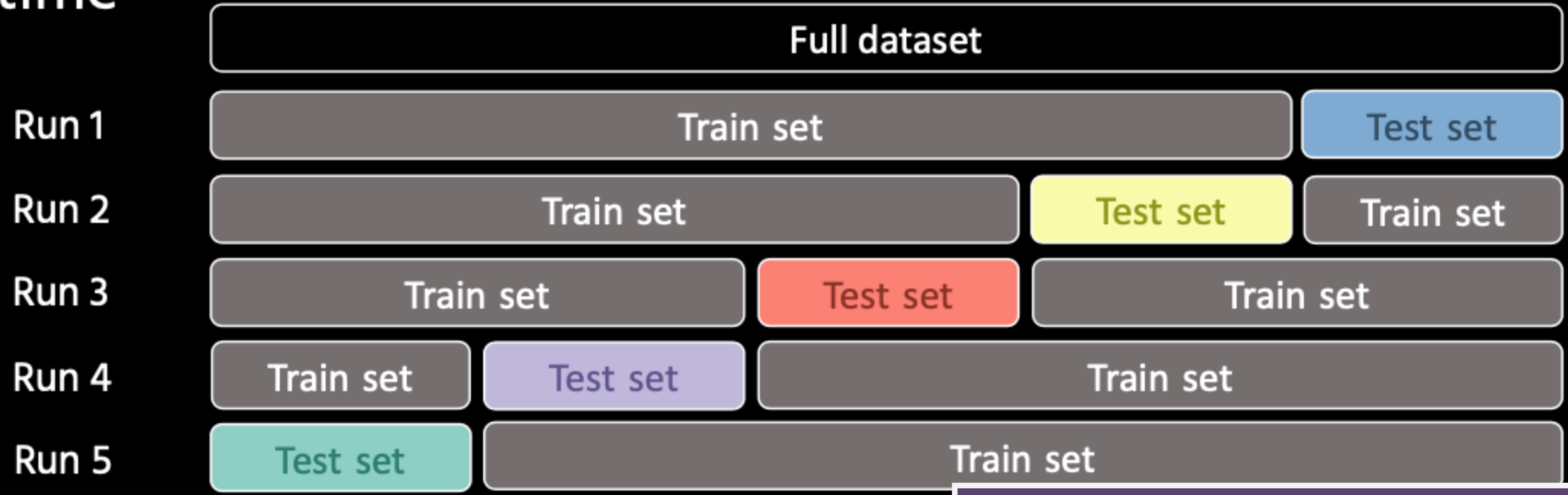
Use best practices in machine learning to establish trust in results

1. Designing and preparing data sets
2. Choosing algoriothms
3. Choosing evaluation metrics
4. Exploring and reporting outputs

Machine learning is not just a set of algorithms, but a community of practice with rules, conventions and best practices to guard against challenges (e.g. overfitting, lack of generalisation)

K-fold cross-validation

- We only have 101 finetuning samples, or ~20 samples for testing
- Train the model k times, reserving different data for testing each time



Louisa Canepa's talk yesterday: Measuring intracluster light with machine learning

Rule 3: Interpret, Diagnose and/or Visualise Models

Develop and/or apply innovative diagnosis and evaluation techniques

Develop and/or apply innovative diagnosis and evaluation techniques

- Monitor in-distribution validation of ML models

Develop and/or apply innovative diagnosis and evaluation techniques

- Monitor in-distribution validation of ML models
- Check for outliers

Develop and/or apply innovative diagnosis and evaluation techniques

- Monitor in-distribution validation of ML models
- Check for outliers
- Visualize model output beyond summary statistics

Develop and/or apply innovative diagnosis and evaluation techniques

- Monitor in-distribution validation of ML models
- Check for outliers
- Visualize model output beyond summary statistics
- Select performance metrics according to the problem being solved

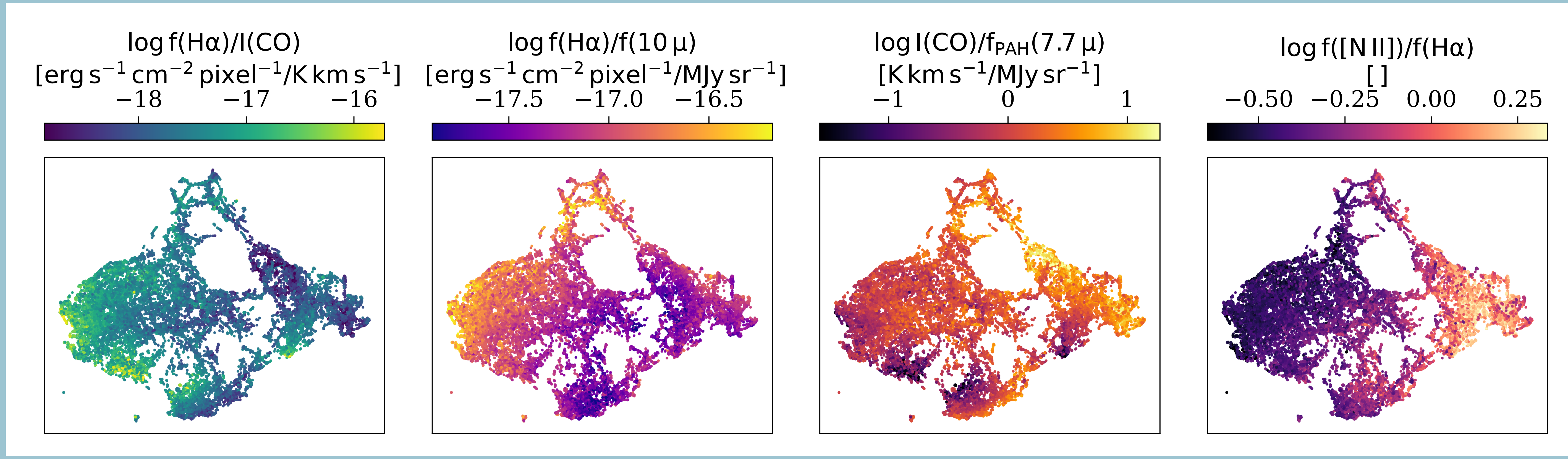
Develop and/or apply innovative diagnosis and evaluation techniques

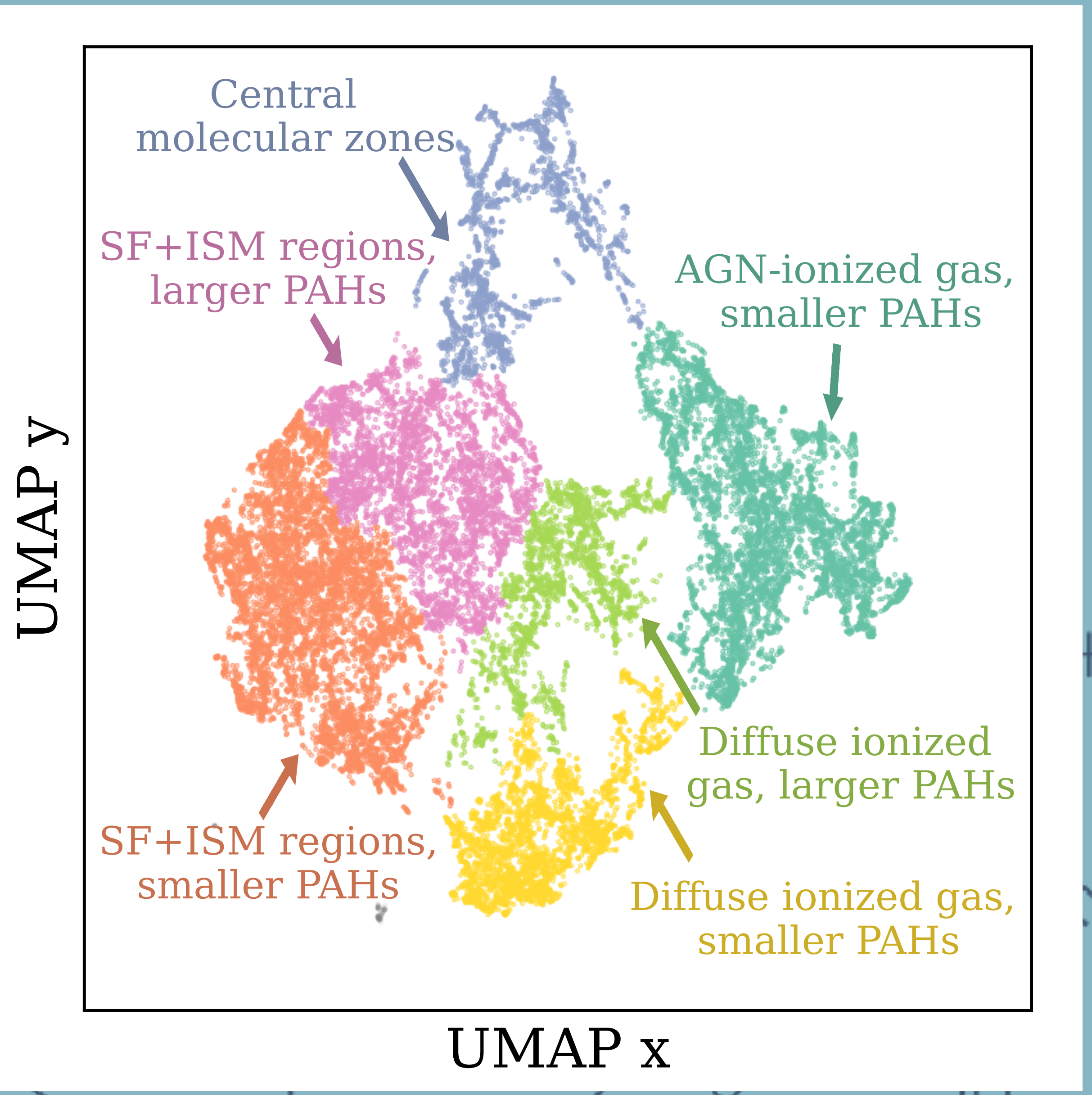
- Monitor in-distribution validation of ML models
- Check for outliers
- Visualize model output beyond summary statistics
- Select performance metrics according to the problem being solved

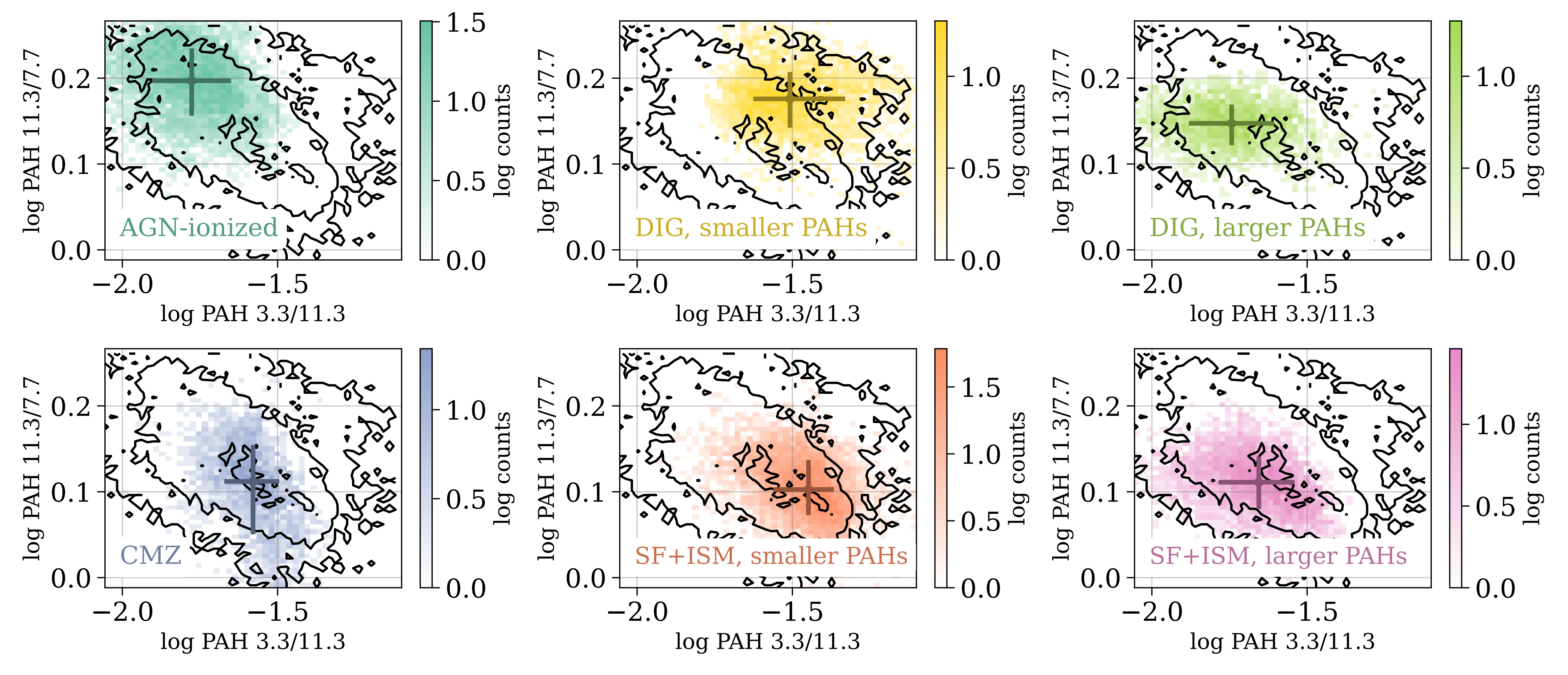
ML models have numerous algorithmic failure modes, and blindly trusting their output may lead to biased inferences about astrophysical problems

PHANGS-ML: dissecting multiphase gas and dust in nearby galaxies using machine learning

Baron et al (2024)







Rule 4: Explore limits and scope of the model

What problems can we confidently assume the model will perform well on?

What problems can we confidently assume the model will perform well on?

- Determine target problem and data with which it can be solved

What problems can we confidently assume the model will perform well on?

- Determine target problem and data with which it can be solved
- Generate training data from idealised or realistic observations

- Determine target problem and data with which it can be solved
- Generate training data from idealised or realistic observations
- Select appropriate modeling and testing methodologies in alignment with scientific use case (e.g. changes in noise)

What problems can we confidently assume the model will perform well on?

What problems can we confidently assume the model will perform well on?

- Determine target problem and data with which it can be solved
- Generate training data from idealised or realistic observations
- Select appropriate modeling and testing methodologies in alignment with scientific use case (e.g. changes in noise)
- Identify limits of the model scope: e.g. causal versus correlational

What problems can we confidently assume the model will perform well on?

- Determine target problem and data with which it can be solved
- Generate training data from idealised or realistic observations
- Select appropriate modeling and testing methodologies in alignment with scientific use case (e.g. changes in noise)
- Identify limits of the model scope: e.g. causal versus correlational
- Don't forget about rare, previously unseen events!

What problems can we confidently assume the model will perform well on?

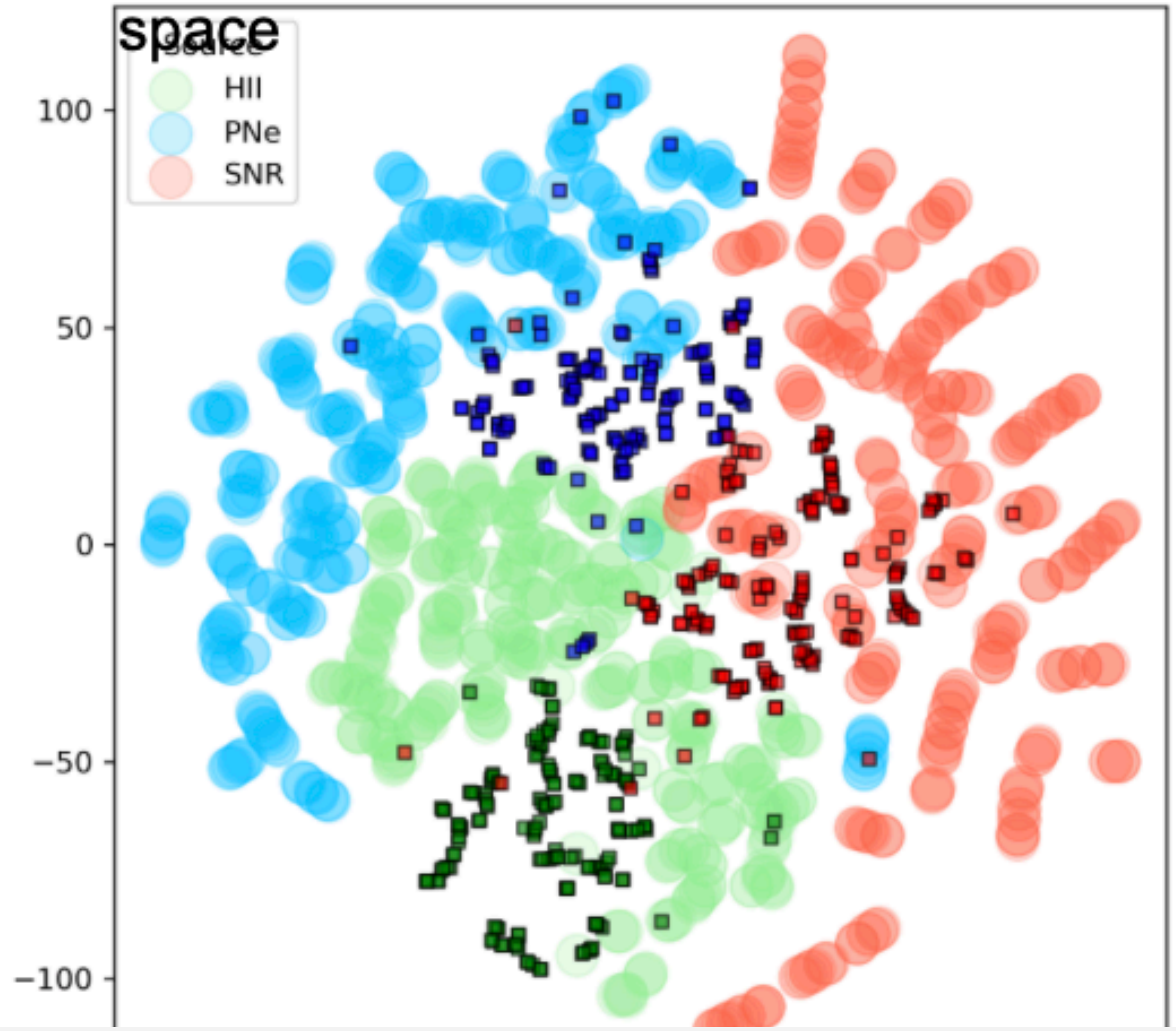
- Determine target problem and data with which it can be solved
- Generate training data from idealised or realistic observations
- Select appropriate modeling and testing methodologies in alignment with scientific use case (e.g. changes in noise)
- Identify limits of the model scope: e.g. causal versus correlational
- Don't forget about rare, previously unseen events!

In many situations where training data is imperfect or incompletely understood, it may be impossible to build unbiased models: transparency if crucial to be able to apply results

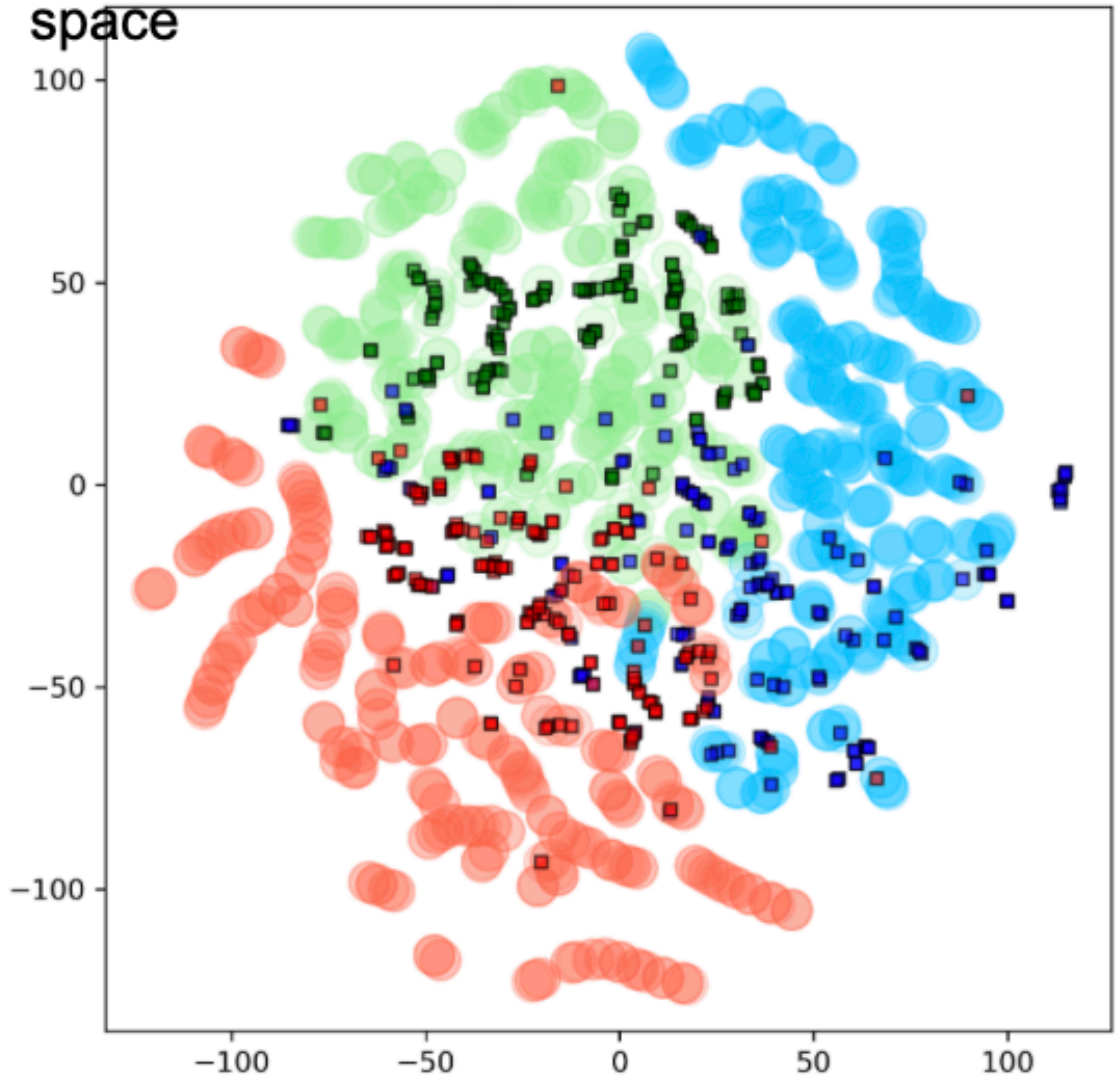
Did we successfully align domains?

3

Tsne representation of the input space



Tsne representation of the embeddings space



Francesco Belfiore's talk yesterday: A domain-adaptation approach to classify ionised nebulae in nearby galaxies

ancesco Belfiore

Rule 5: Share and Discuss Lessons Learned

Share Lessons Learned to Avoid Duplication

Share Lessons Learned to Avoid Duplication

1. Adapting ML to astronomical problems involves both successes and informative failures

1. Adapting ML to astronomical problems involves both successes and informative failures
2. Publications are biased towards successes

Share Lessons Learned to Avoid Duplication

Share Lessons Learned to Avoid Duplication

1. Adapting ML to astronomical problems involves both successes and informative failures
2. Publications are biased towards successes
3. Some models may never get published for not reaching thresholds over previous approaches

Share Lessons Learned to Avoid Duplication

1. Adapting ML to astronomical problems involves both successes and informative failures
2. Publications are biased towards successes
3. Some models may never get published for not reaching thresholds over previous approaches
4. Write manuscripts that are pedagogical as well as scientific

Share Lessons Learned to Avoid Duplication

1. Adapting ML to astronomical problems involves both successes and informative failures
2. Publications are biased towards successes
3. Some models may never get published for not reaching thresholds over previous approaches
4. Write manuscripts that are pedagogical as well as scientific

Biasing publication exclusively towards successful approaches could discourage model exploration to avoid risk of failure, an deprive the literature of these examples

ICBINB @ NeurIPS2021

Call for Papers Guidelines Schedule Accepted Talks Panel Breakouts PC CoC Awards Feedback

I (Still) Can't Believe It's Not Better! Workshop

ICBINB@NeurIPS 2021 - A Workshop for "beautiful" ideas that *should* have worked

Rule 6: Make Software and Data Publicly Available

Consider Making Code, Models and Data Public!

Consider Making Code, Models and Data Public!

1. Enables others to reproduce research results, verify correctness

Consider Making Code, Models and Data Public!

1. Enables others to reproduce research results, verify correctness
2. Enables reuse of model architectures, loss functions or whole pipelines

Consider Making Code, Models and Data Public!

1. Enables others to reproduce research results, verify correctness
2. Enables reuse of model architectures, loss functions or whole pipelines
3. Provides a foundation for new projects

Consider Making Code, Models and Data Public!

1. Enables others to reproduce research results, verify correctness
2. Enables reuse of model architectures, loss functions or whole pipelines
3. Provides a foundation for new projects
4. But: data and model sharing rights might be complicated!

Consider Making Code, Models and Data Public!

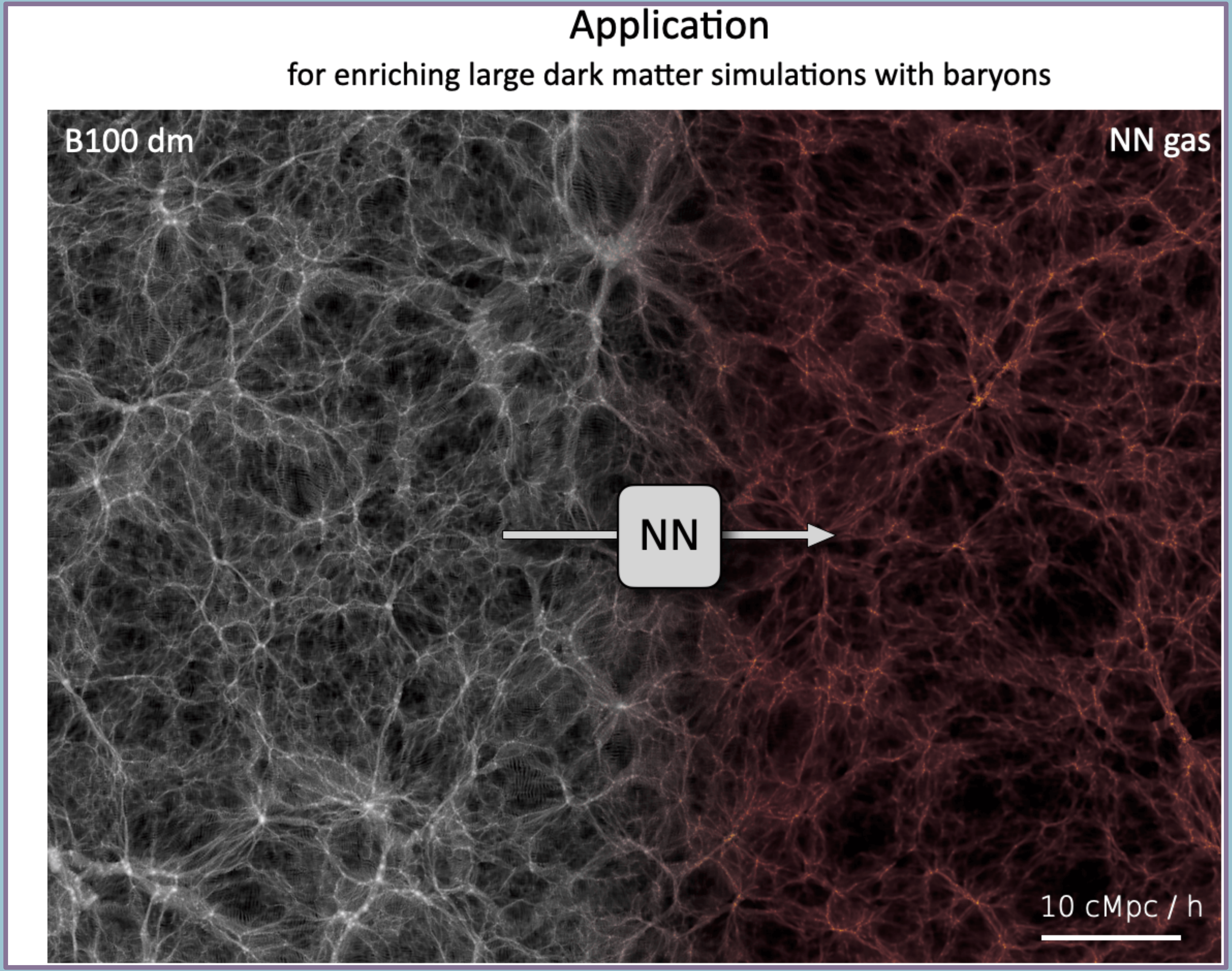
1. Enables others to reproduce research results, verify correctness
2. Enables reuse of model architectures, loss functions or whole pipelines
3. Provides a foundation for new projects
4. But: data and model sharing rights might be complicated!

Benchmarking new models against existing ones, and publishing the results, are best done when existing models are public!

EMBER: emulating baryons from dark matter-only simulations over cosmic time

Bernadini et al (2022)

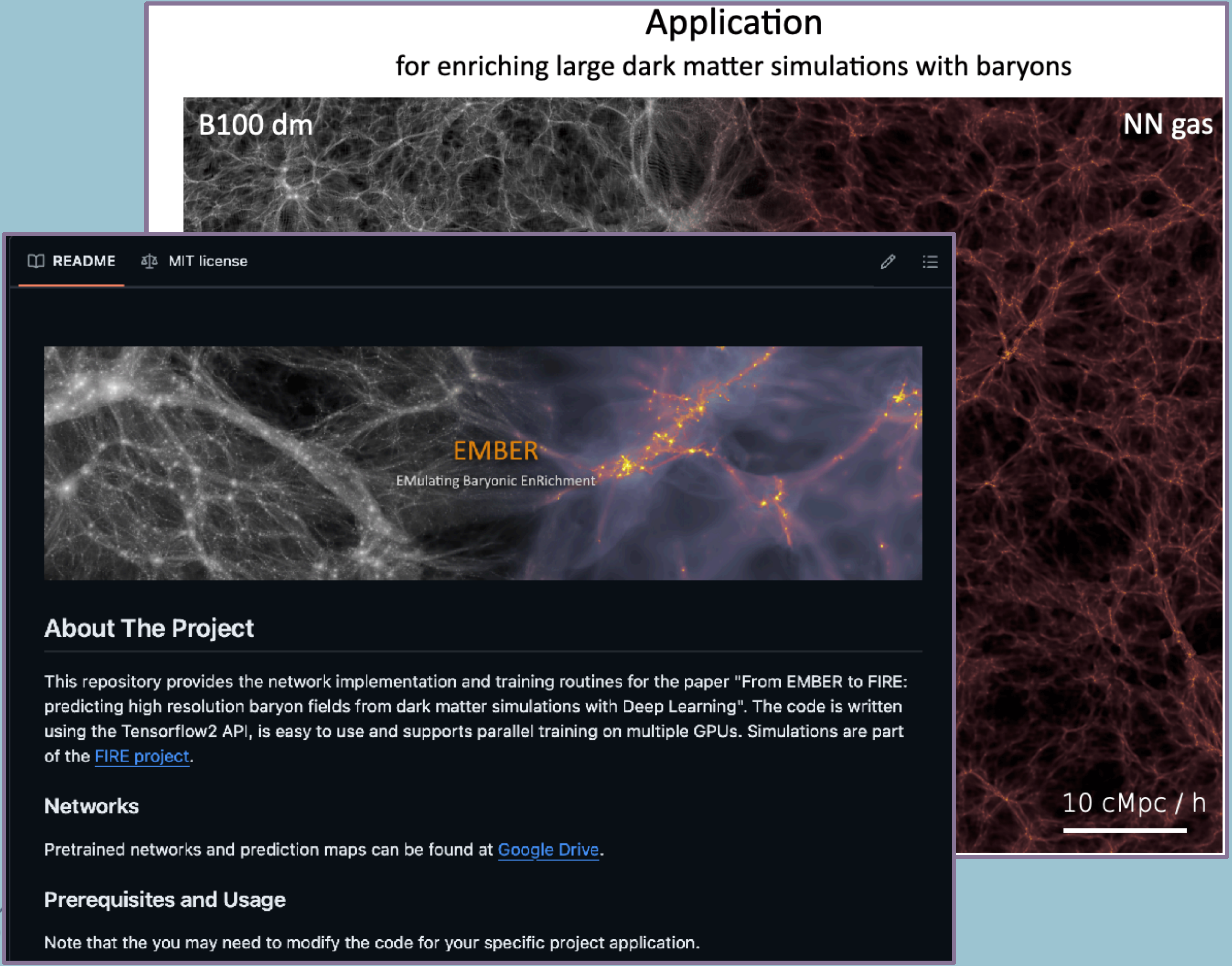
See talk on Wednesday morning!



EMBER: emulating baryons from dark matter-only simulations over cosmic time

Bernadini et al (2022)

See talk on Wednesday morning!



README MIT license

About The Project

This repository provides the network implementation and training routines for the paper "From EMBER to FIRE: predicting high resolution baryon fields from dark matter simulations with Deep Learning". The code is written using the Tensorflow2 API, is easy to use and supports parallel training on multiple GPUs. Simulations are part of the FIRE project.

Networks

Pretrained networks and prediction maps can be found at Google Drive.

Prerequisites and Usage

Note that the you may need to modify the code for your specific project application.

Rules are meant to be broken!

It’s your turn!

Interactive Activity

Interactive Activity

a. Which of the rules resonate with you?
b. Which do you disagree with? Why?
c. What did we miss?

Interactive Activity

- Think on your own (2 minutes)
- Discuss with a partner (4 minutes)
- Discuss in group of 4 (8 minutes)
- Share with full group (6 minutes)

a. Which of the rules resonate with you?
b. Which do you disagree with? Why?
c. What did we miss?

Interactive Activity

- Think on your own (2 minutes)
- Discuss with a partner (4 minutes)
- Discuss in group of 4 (8 minutes)
- Share with full group (6 minutes)

a. Which of the rules resonate with you?
b. Which do you disagree with? Why?
c. What did we miss?

1. Compare against a domain reference and put result into larger context
2. Adopt best practices from the ML community
3. Interpret, Diagnose and/or Visualise Models
4. Explore limits and scope of the model
5. Share and Discuss Lessons Learned
6. Make Software and Data Publicly Available

Interactive Activity: Share-Out

a. Which of the rules resonate with you?
b. Which do you disagree with? Why?
c. What did we miss?

Notes

- Start here

Questions?

SRON
Netherlands Institute for Space Research

d.huppenkothen@sron.nl