Machine Learning for Cosmology

Michelle Ntampaka

STScI | SPACE TELESCOPE SCIENCE INSTITUTE

Astronomy is the ideal sandbox for machine learning.

Astronomy is the ideal sandbox for machine learning.

- Minimal privacy concerns.
- Culture of sharing data.
- Cosmological surveys are rich with information.
- Well-posed questions.
- Public interest and support.
- Data are non-monetizable.
- *This does not exempt us from ethical concerns!*

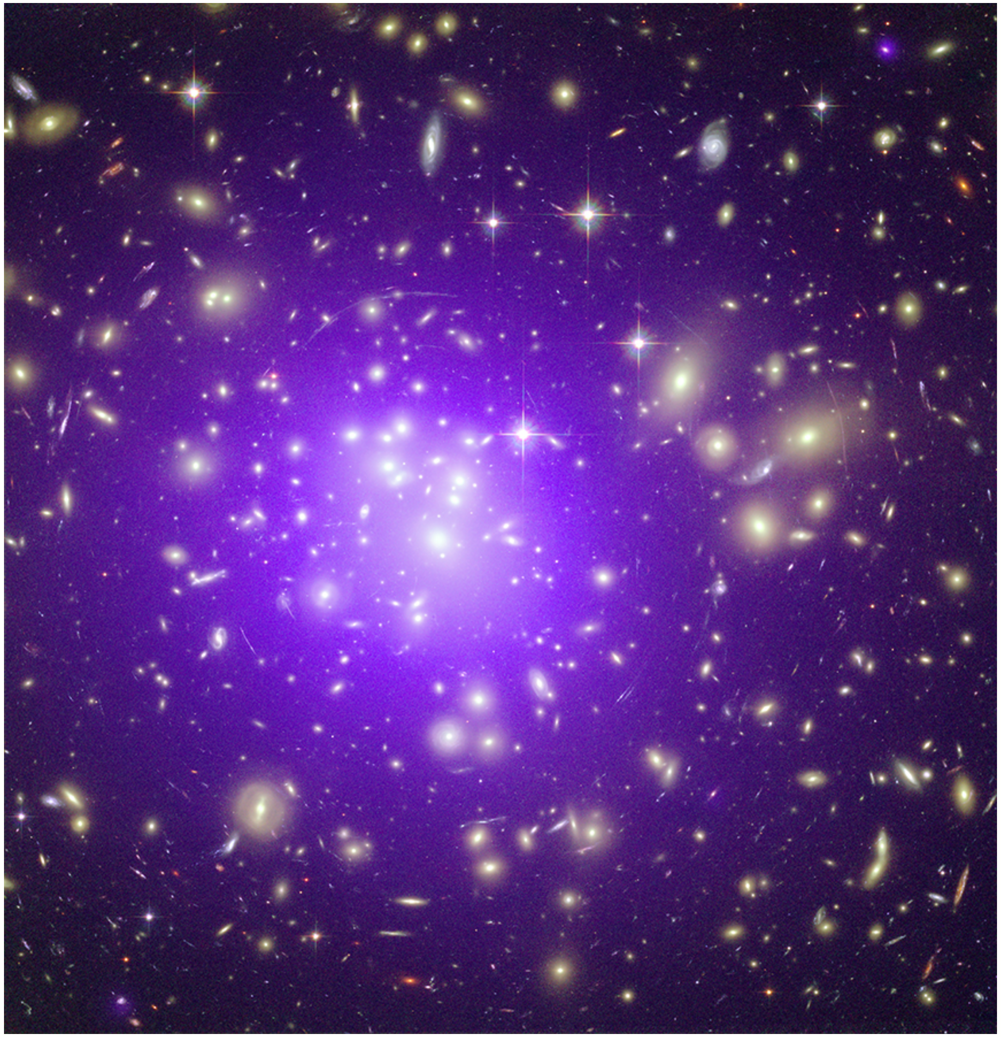
Is machine learning the right tool for astronomy?

Is machine learning the right tool for astronomy?

1. Can ML be trusted?
2. Can ML be used to make physical discoveries?
3. What role will machine learning play in the future of astronomy?

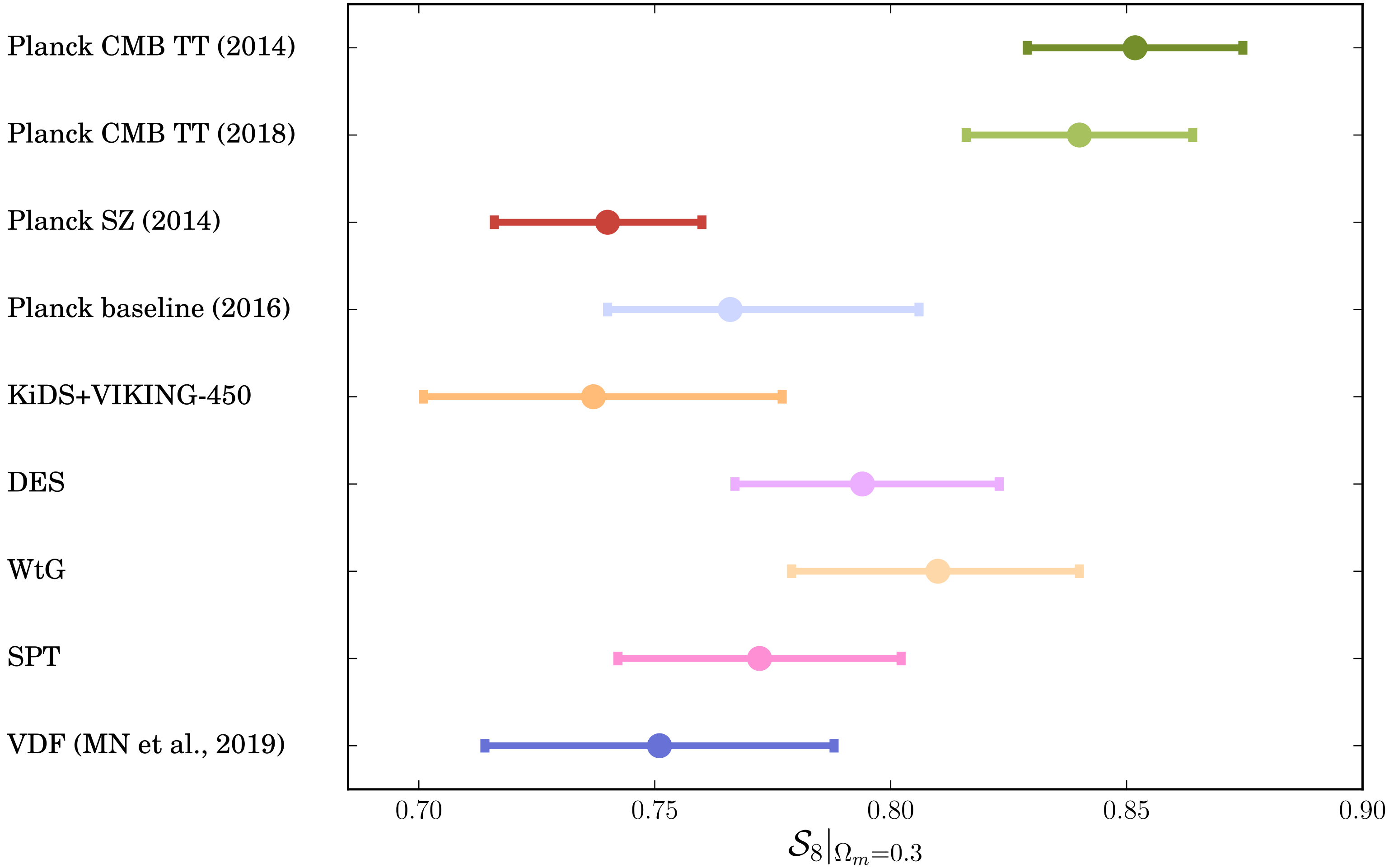
A Case Study in Interpretability: Galaxy Clusters as a Cosmological Probe

Galaxy Clusters



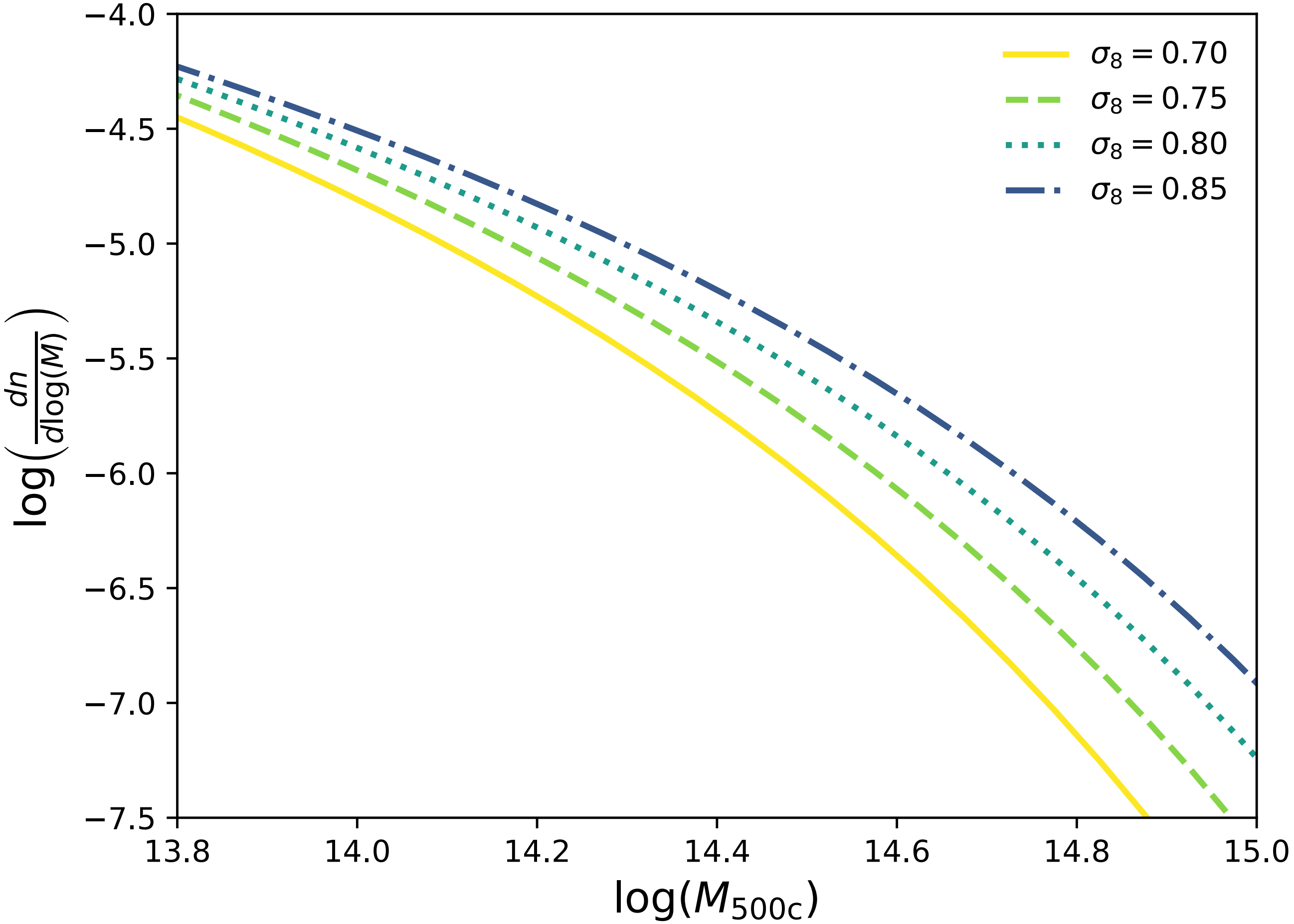
X-ray image credit: NASA/CXC/MIT/E.-H Peng et al; Optical image credit: NASA/STScI

Tensions in the current cosmological model: S_8 (CMB vs. LSS)



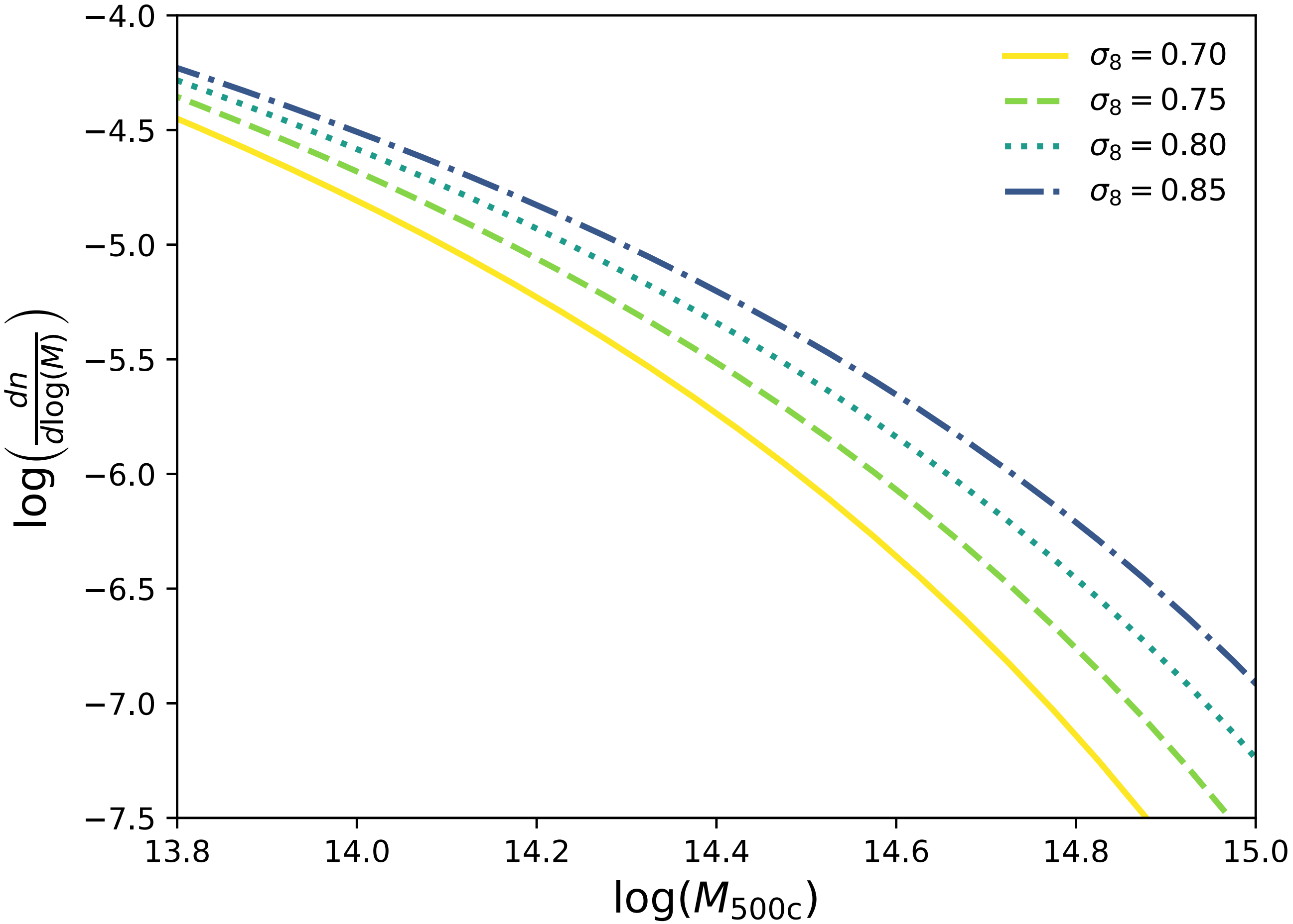
Ntampaka+ 2019b

Clusters as a Cosmological Probe



Cluster abundance is sensitive to the underlying cosmology (especially to σ_8!)

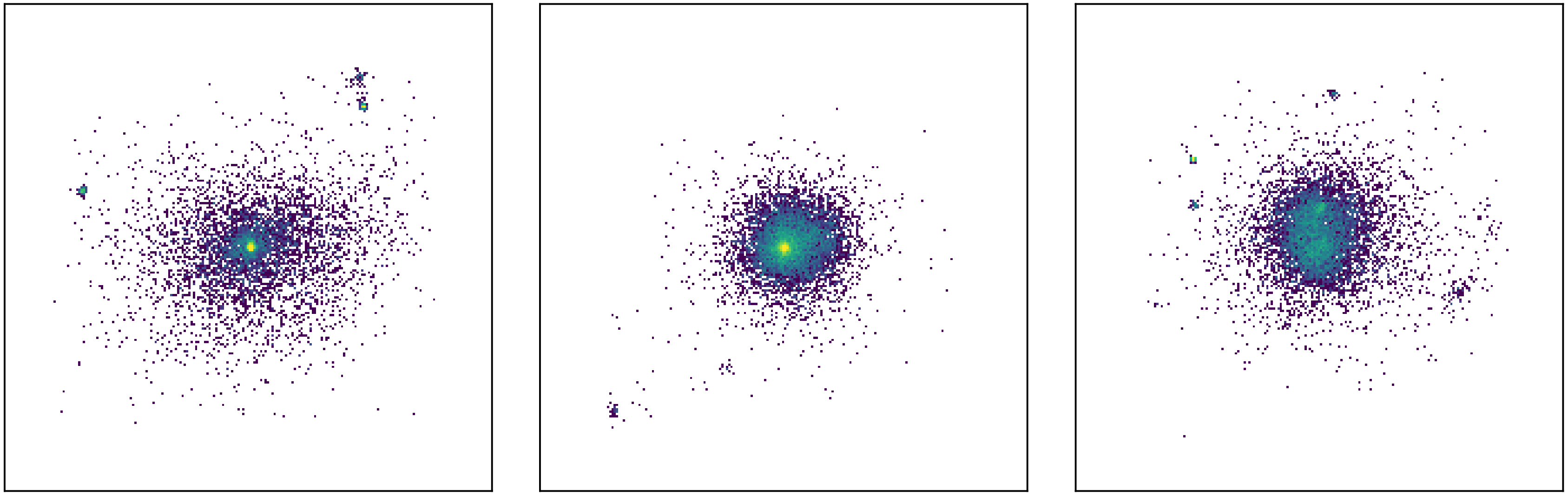
Mass Bias vs. Cosmology



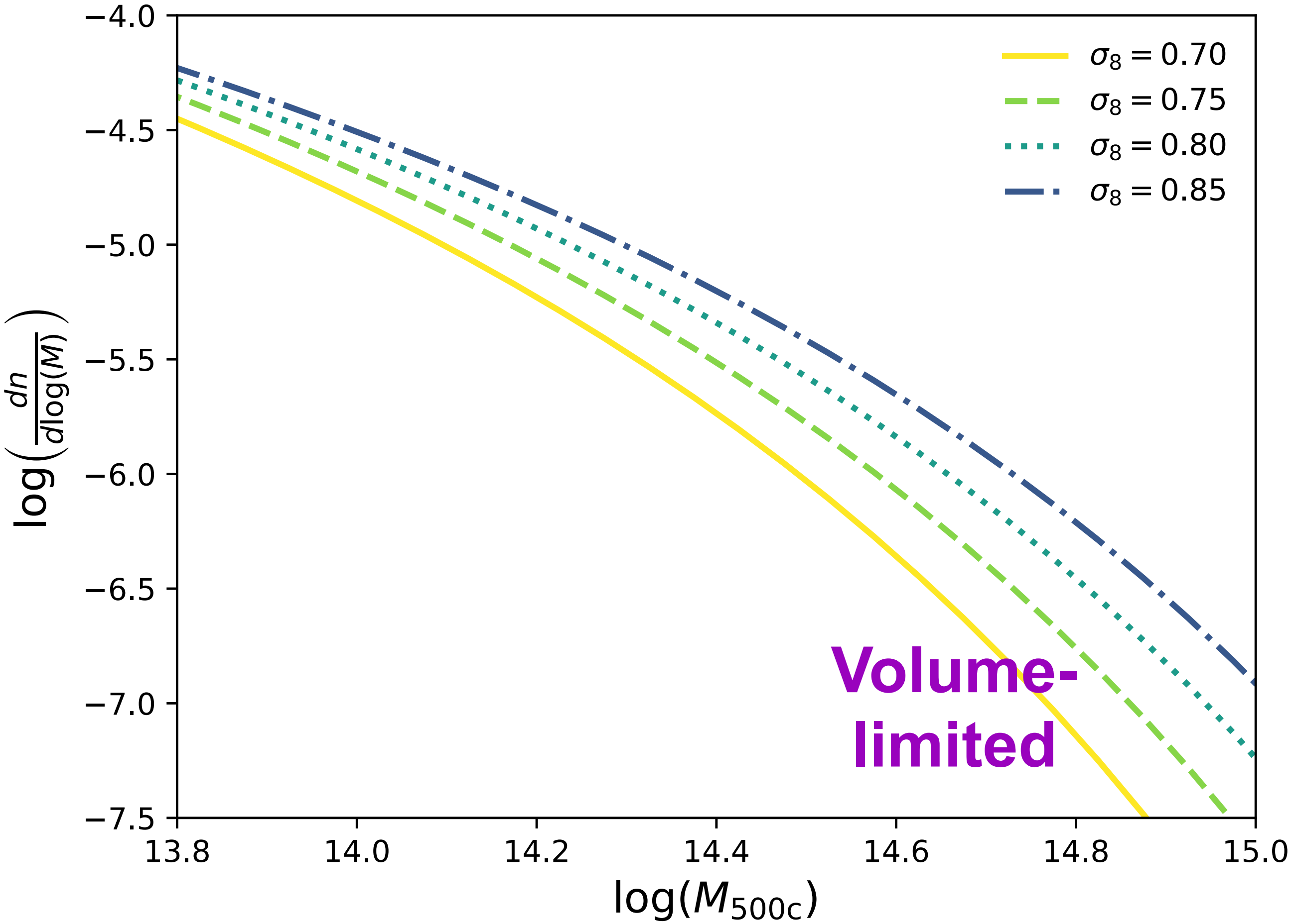
It is difficult to discern between biased cluster masses and a change in cosmology (WL masses can help us with mass calibration).

Simulated Galaxy Clusters

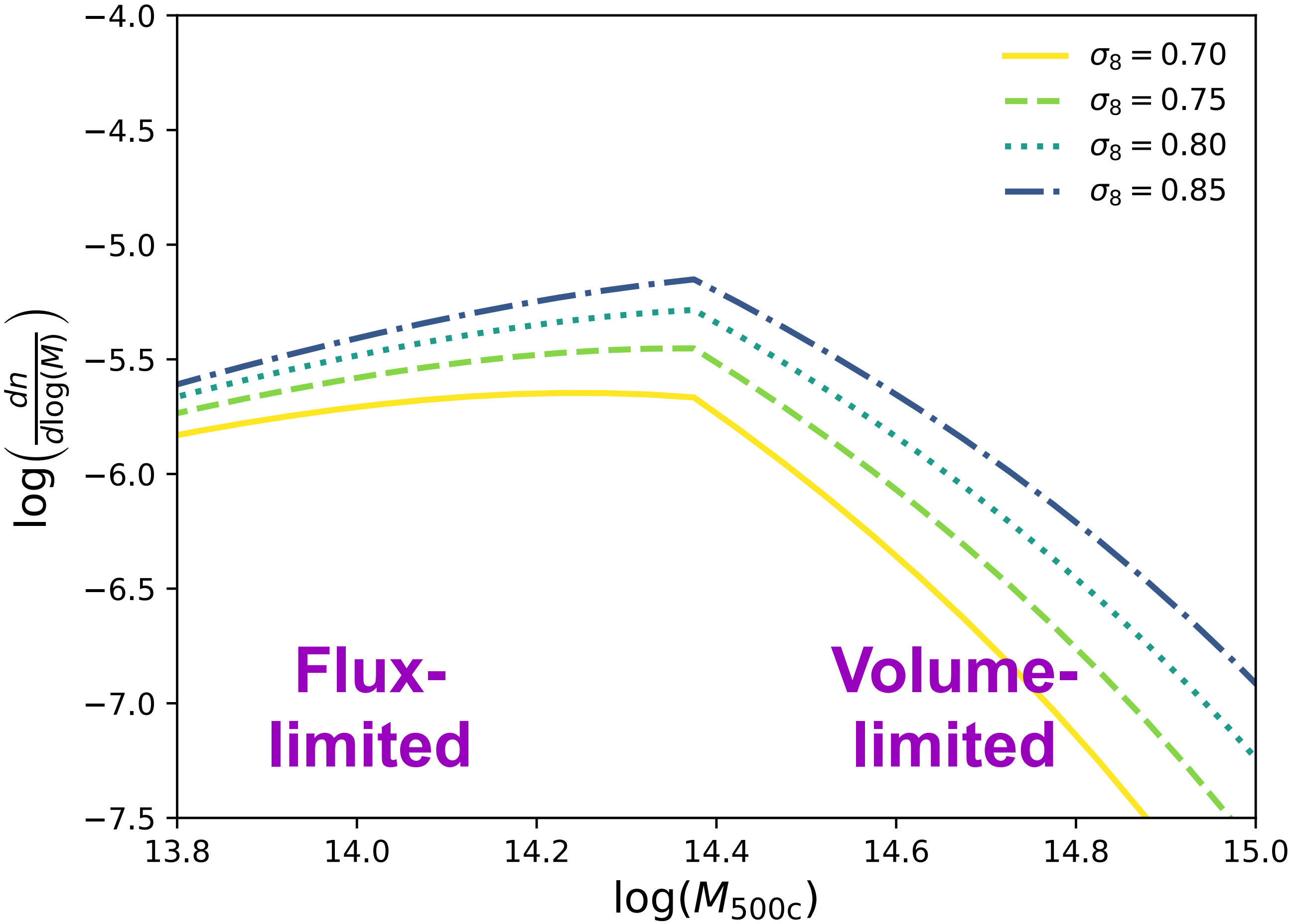
- Magneticum Cosmological Hydrodynamical Simulation
- WMAP7 Cosmology
- Box length = $352h^{-1}$ Mpc
- *Chandra*-like mock observation



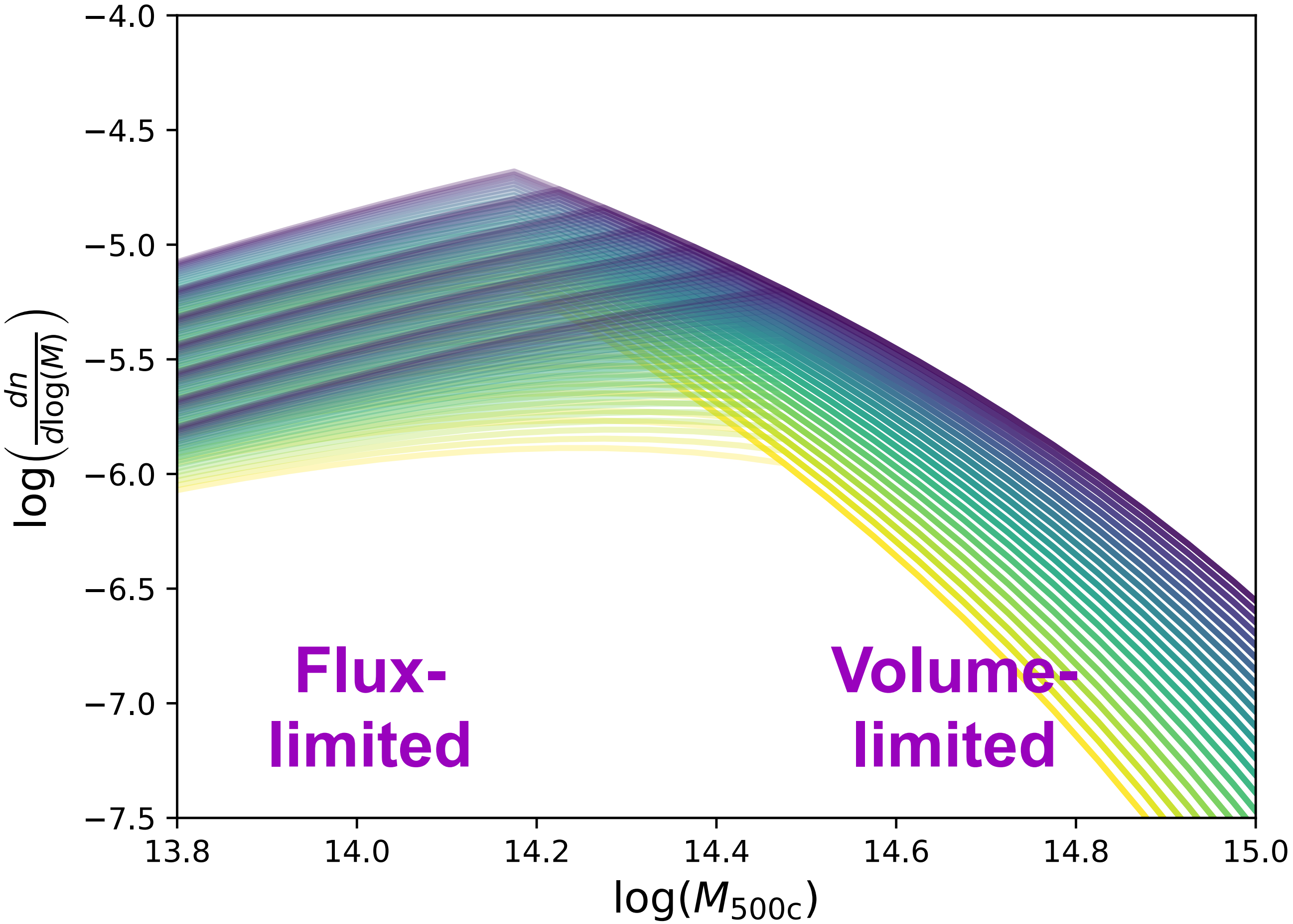
Many Cosmologies from One Simulation



Many Cosmologies from One Simulation

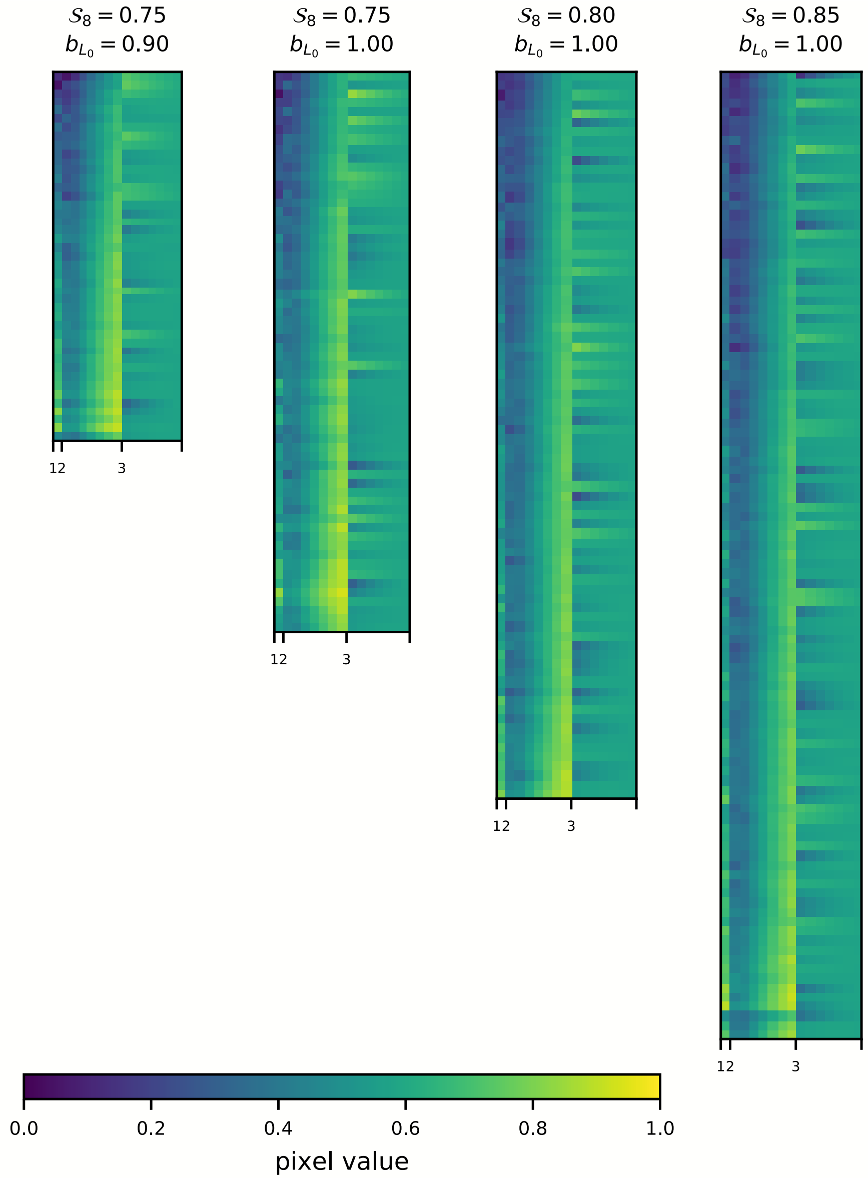


Many Cosmologies from One Simulation



Typical X-ray surveys include both flux- and volume-limited regimes, and also uncertainty in the mass at which these two join.

Simulated Cluster Samples



Pixel = cluster observable
Row = cluster
Image = simulated cluster sample

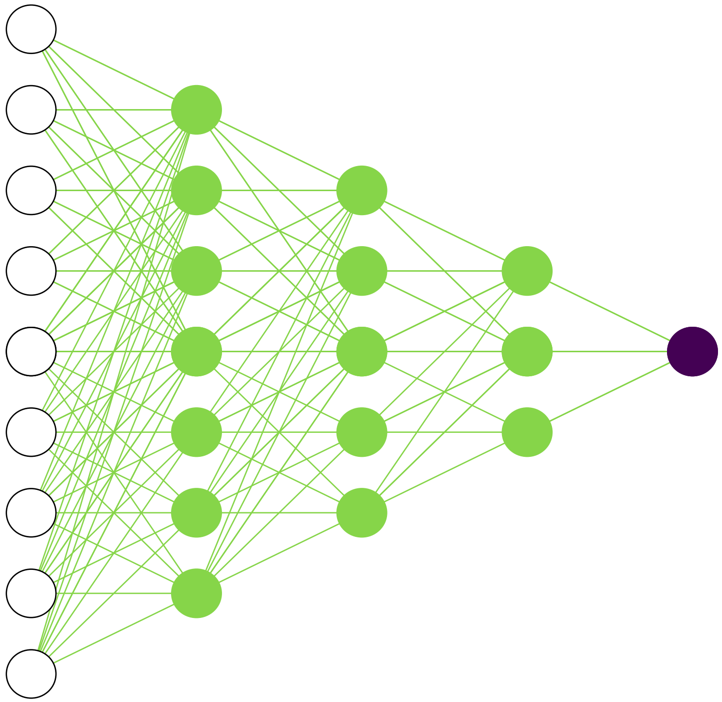
Pixels are cluster observables:

1. Temperature
2. Gas mass profile
3. Gas density slope profile

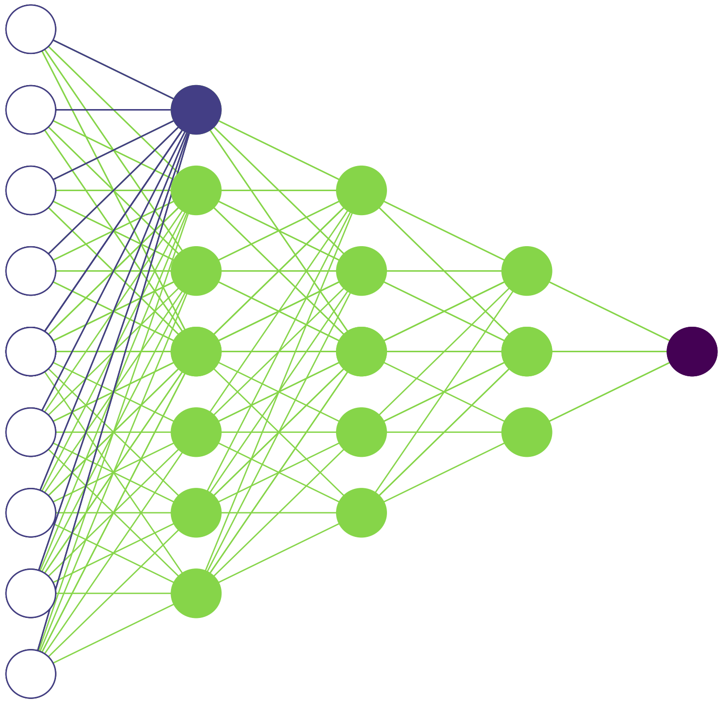
We have simulated realistic cluster samples
for many cosmologies & have cast each into a 2D array.

ML tools:
Neural Networks &
Encoders

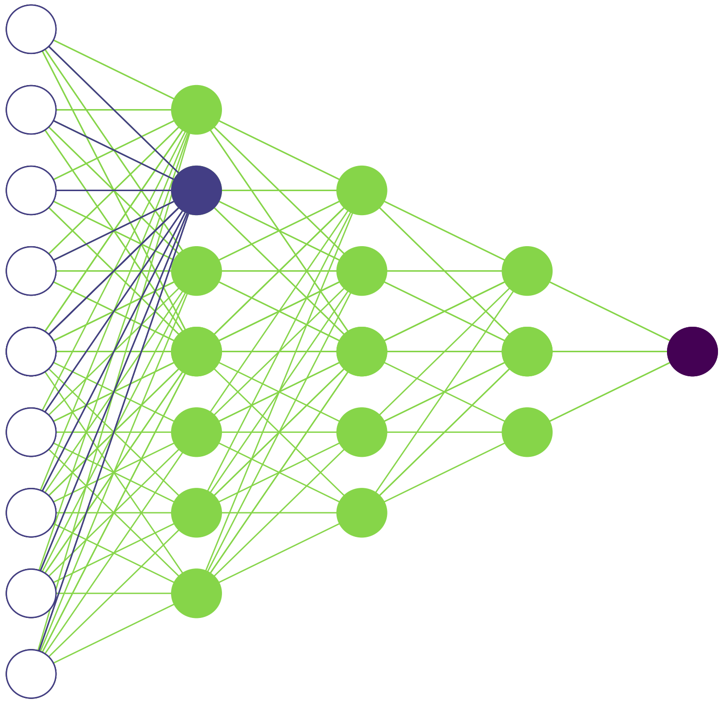
Neural Network



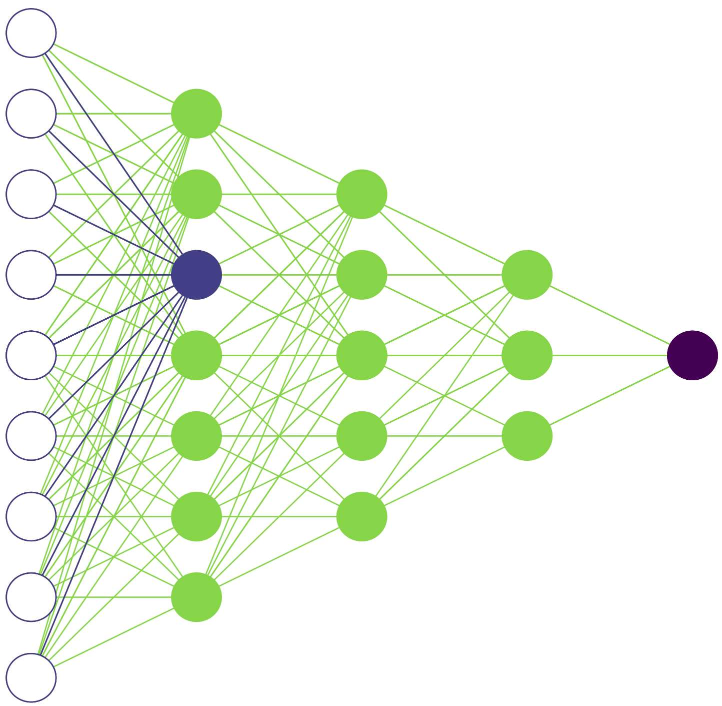
Neural Network



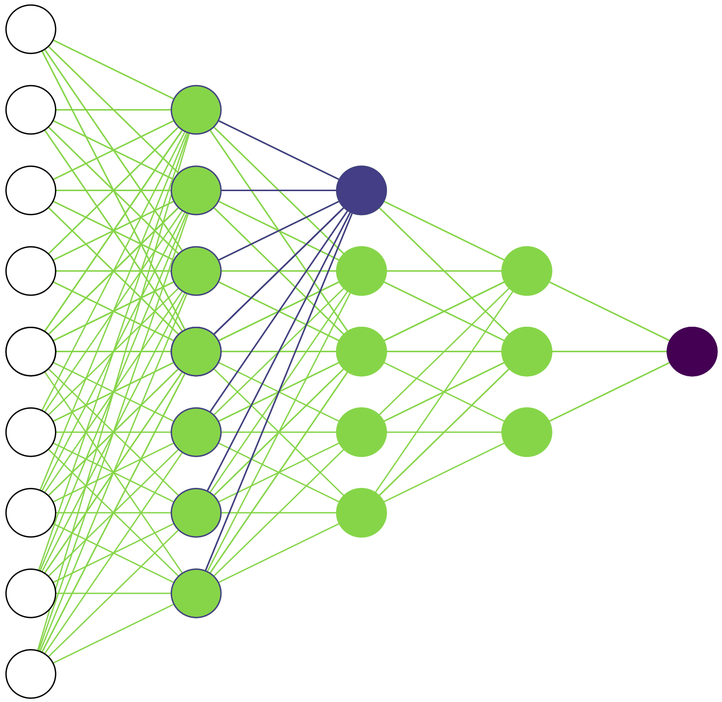
Neural Network



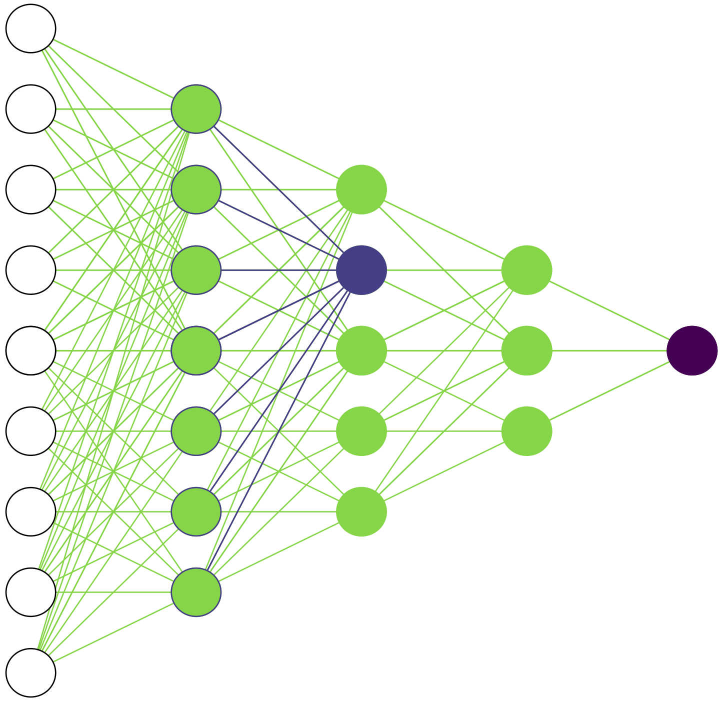
Neural Network



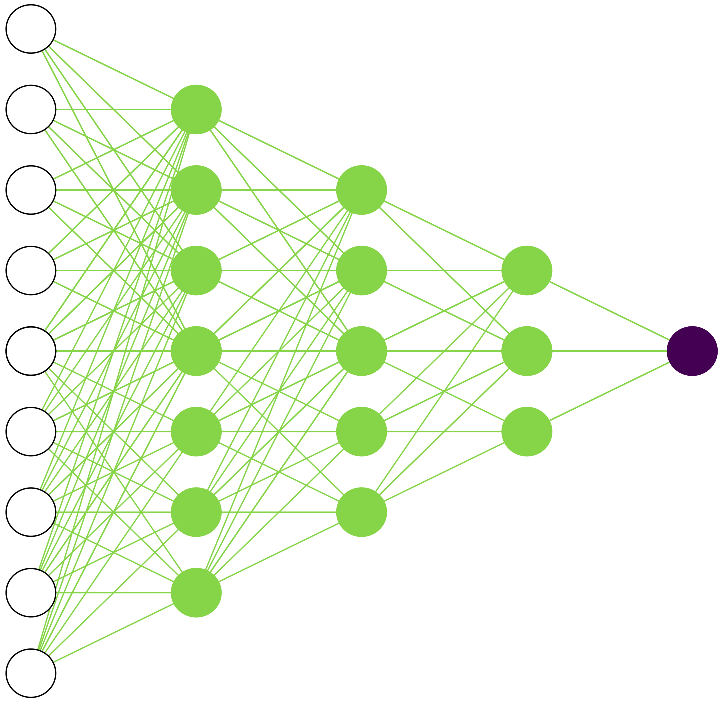
Neural Network



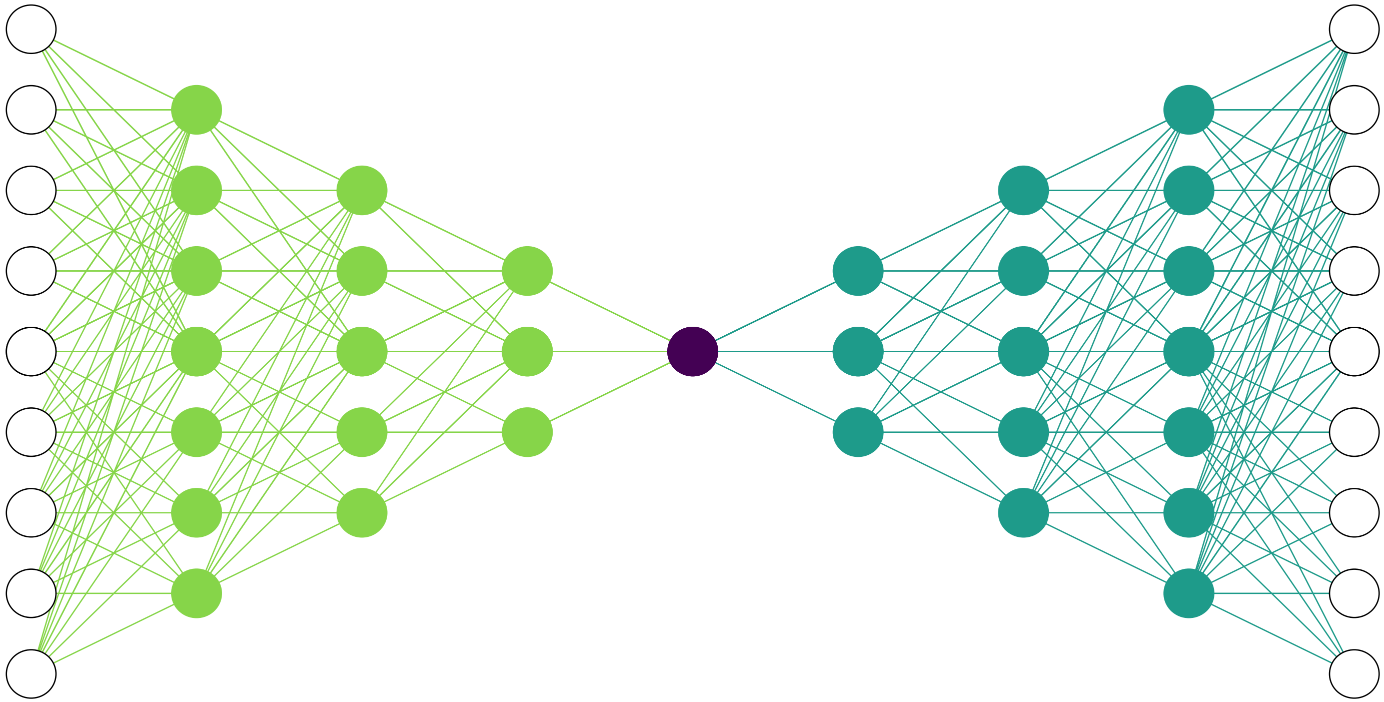
Neural Network



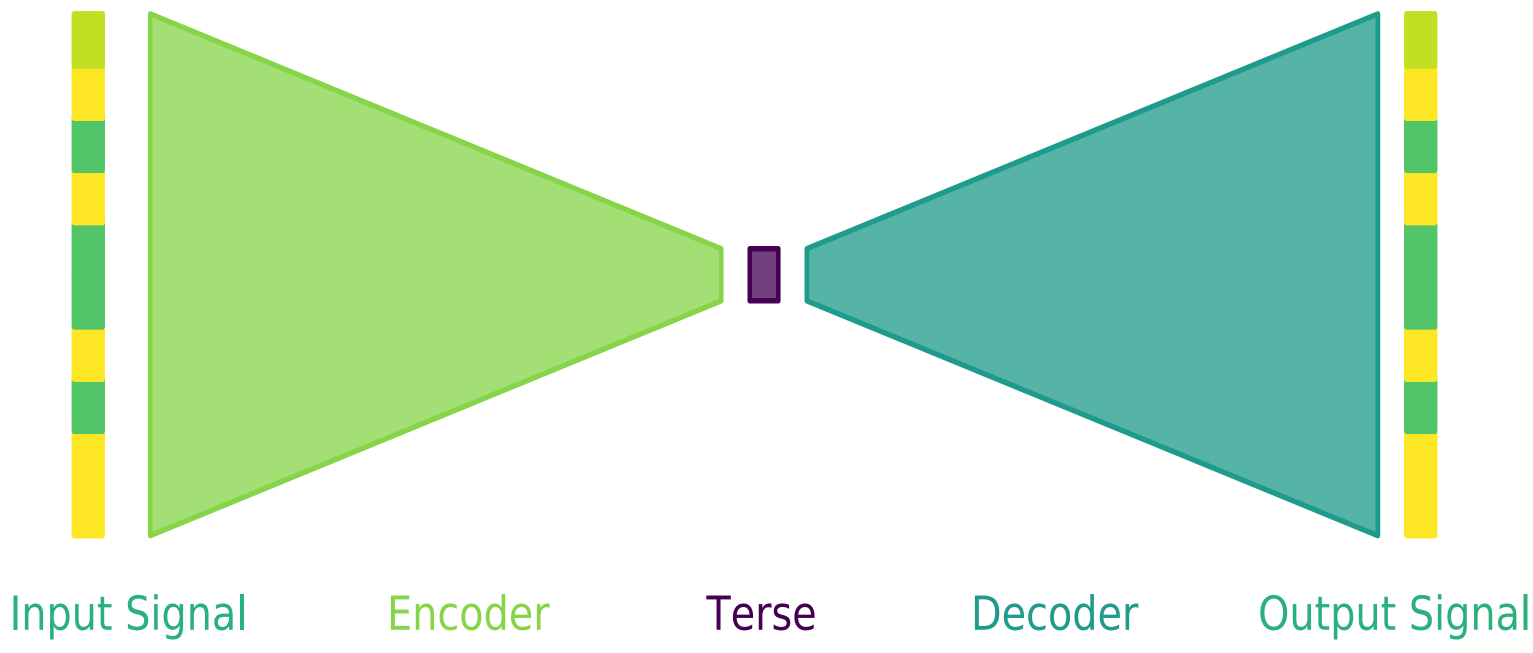
Neural Network



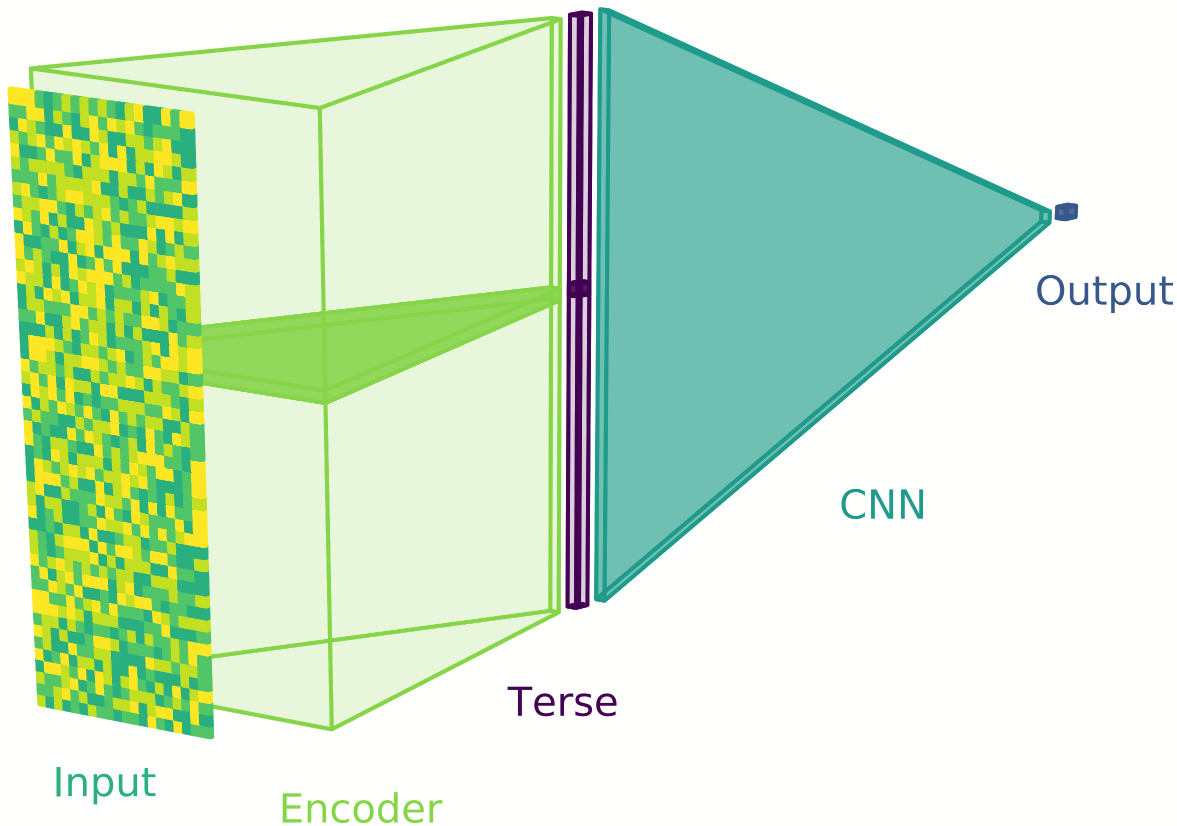
Autoencoder



Autoencoder

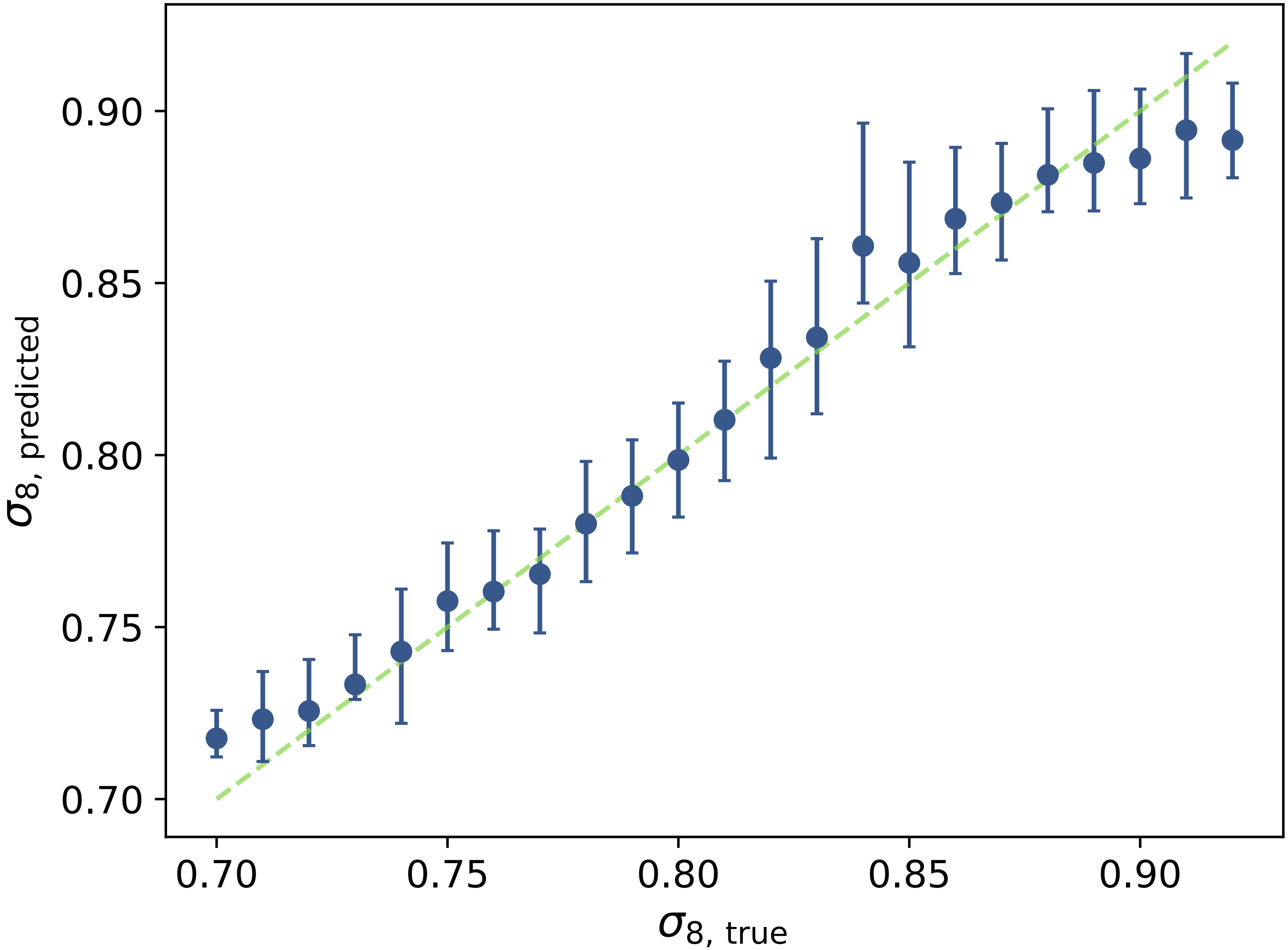


Supervised Encoder



This encoder is engineered to mimic a human approach to the task:
cluster observables ➔ cluster mass proxy
list of mass estimates ➔ cosmology

Results: σ_8 Predictions



Interpretation

Can ML be trusted?
Can it drive physical understanding?

Can ML be trusted?
Can it drive physical understanding?

1. Terse Value Correlations: to assess whether the model will generalize.
These can help us to trust in our ML models.

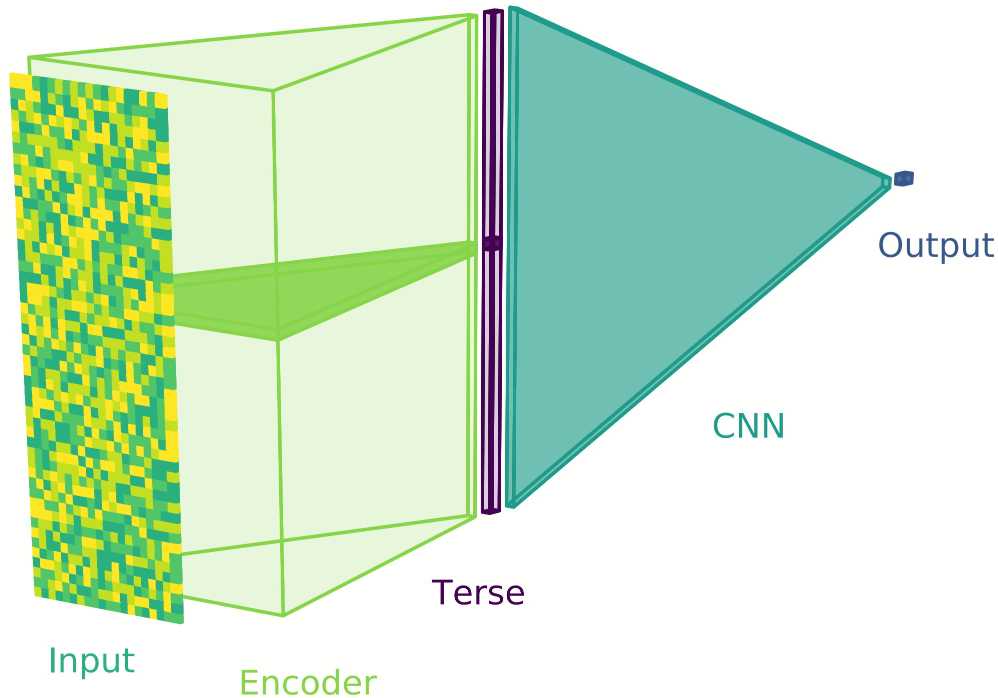
2. Saliency maps: identify what part of the cluster carries the most cosmological information.
These can lead to ML-driven discoveries.

1. Correlations ➔ Trust

Terse Value Correlations

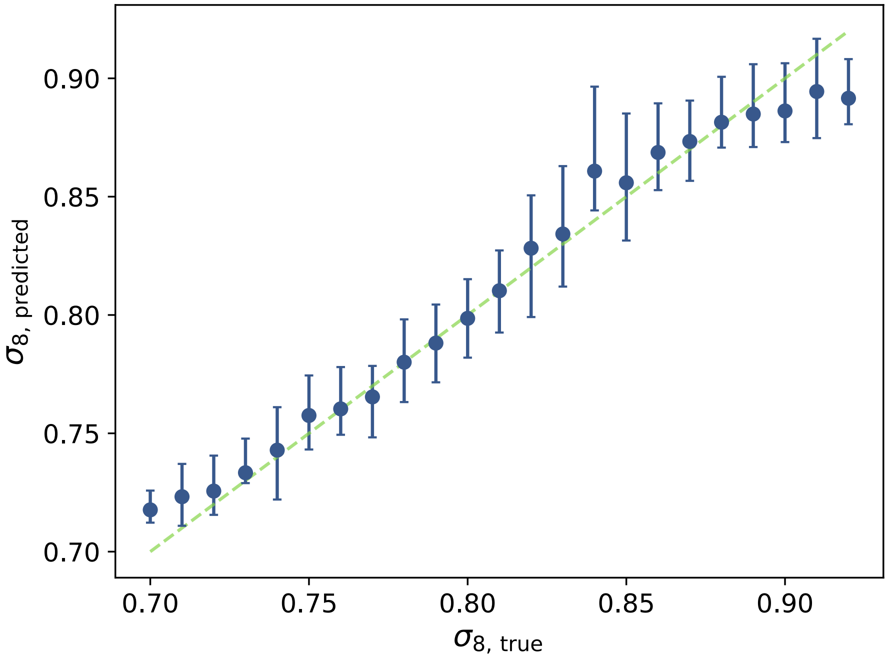
Big idea: visualize how the terse layer correlates with cluster parameters.

Build trust in the model by asking “Is the terse layer summarizing sensible features?”

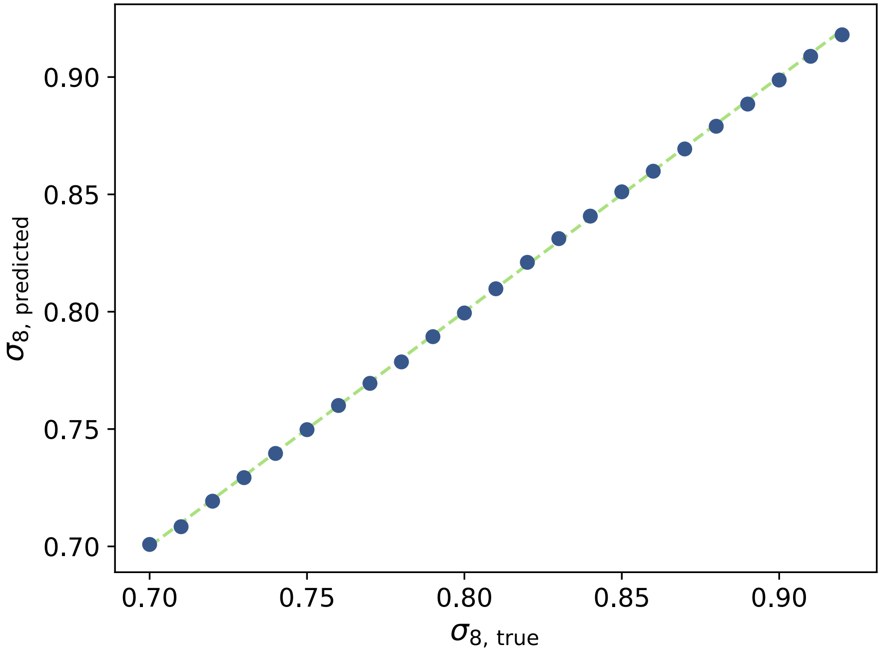


These results look *suspicious*! Did the ML cheat?

Trustworthy Results:



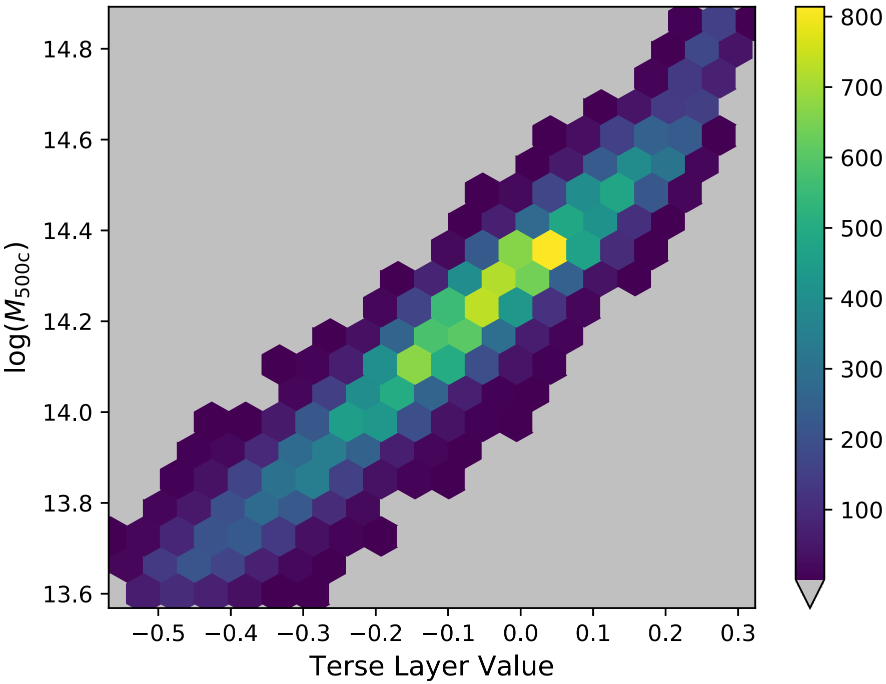
Suspicious Results:



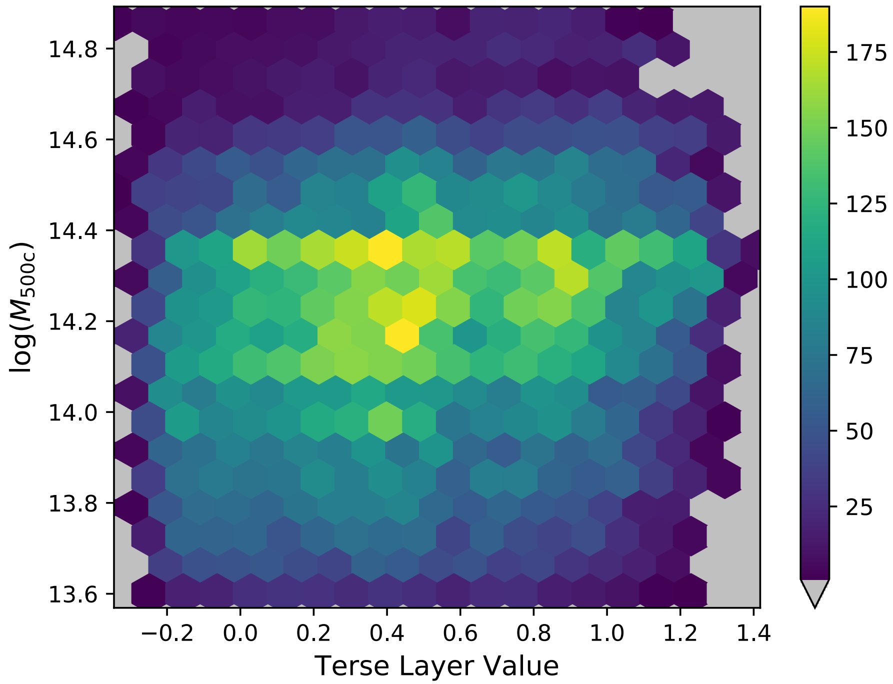
Here, we find a factor-of-10 better results. But is it trustworthy? Will it generalize to real data?

These results look *suspicious*! Did the ML cheat?

Trustworthy Results:



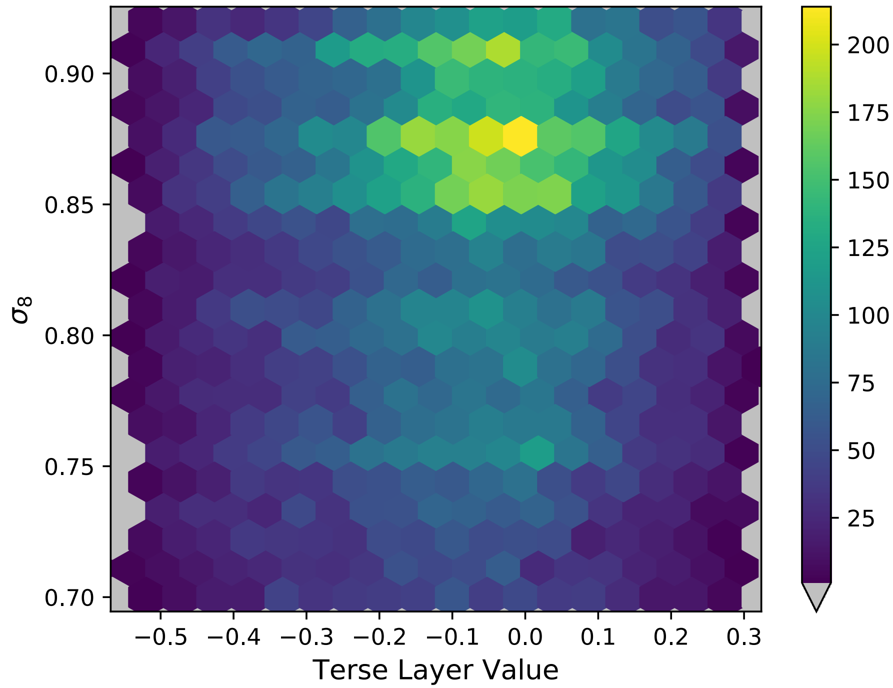
Suspicious Results:



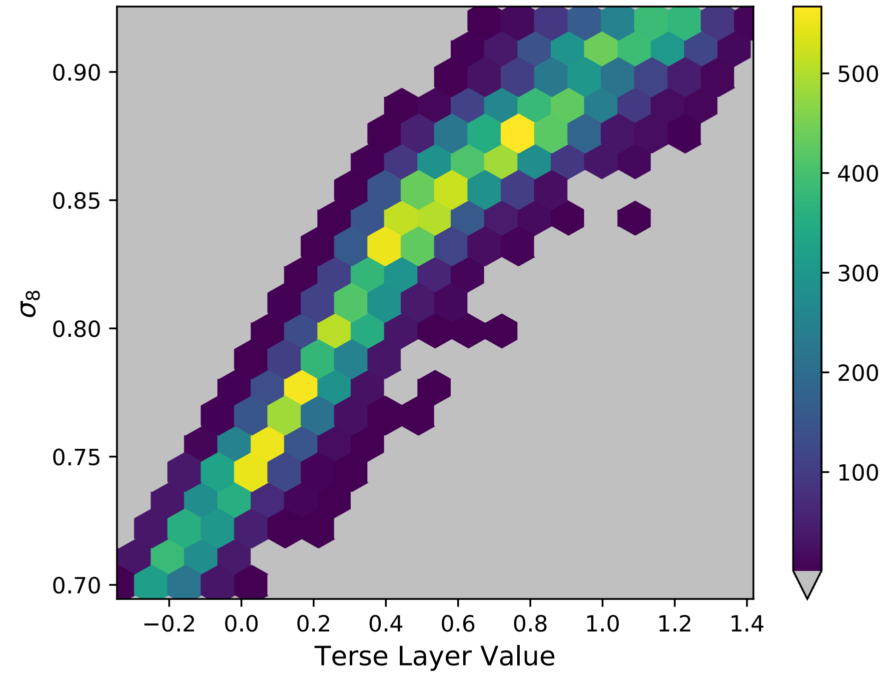
Why is the terse value uncorrelated with mass? This is a red flag that something isn't right.

These results look *suspicious*! Did the ML cheat?

Trustworthy Results:



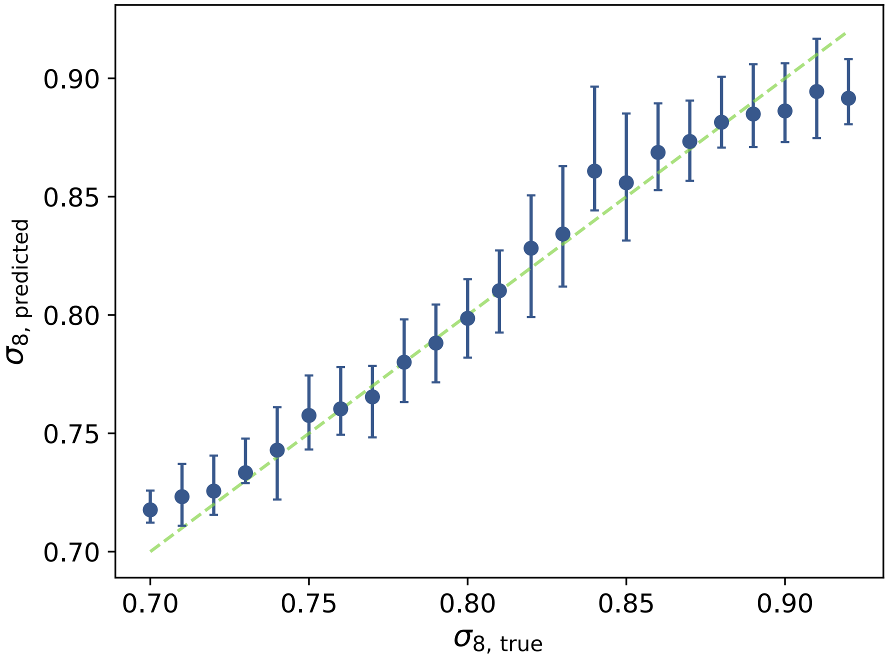
Suspicious Results:



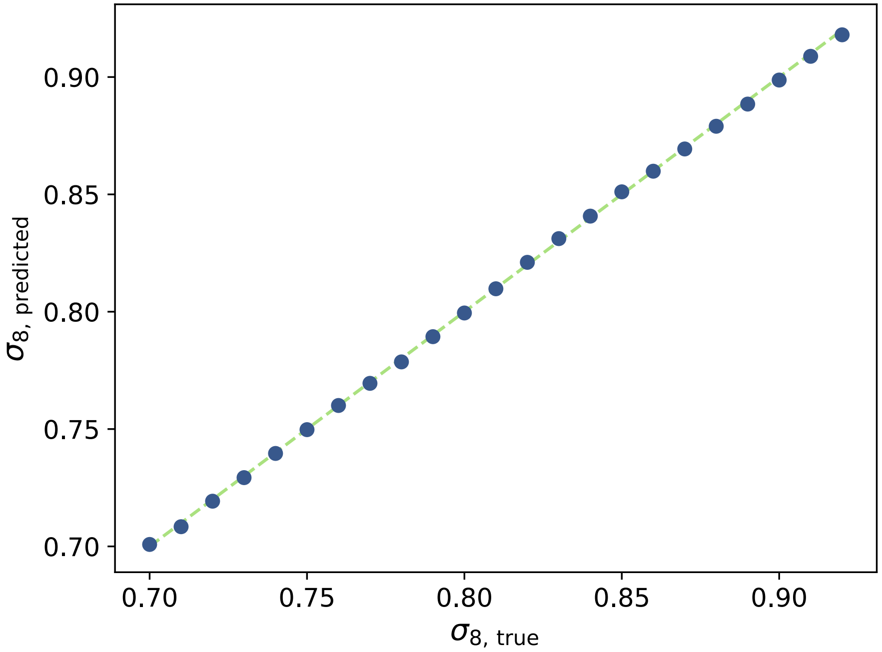
Here's the problem. The model has picked up on a cosmology-dependent simulation artifact and is able to infer σ_8 from just one cluster.

These results look *suspicious*! Did the ML cheat?

Trustworthy Results:



Suspicious Results:

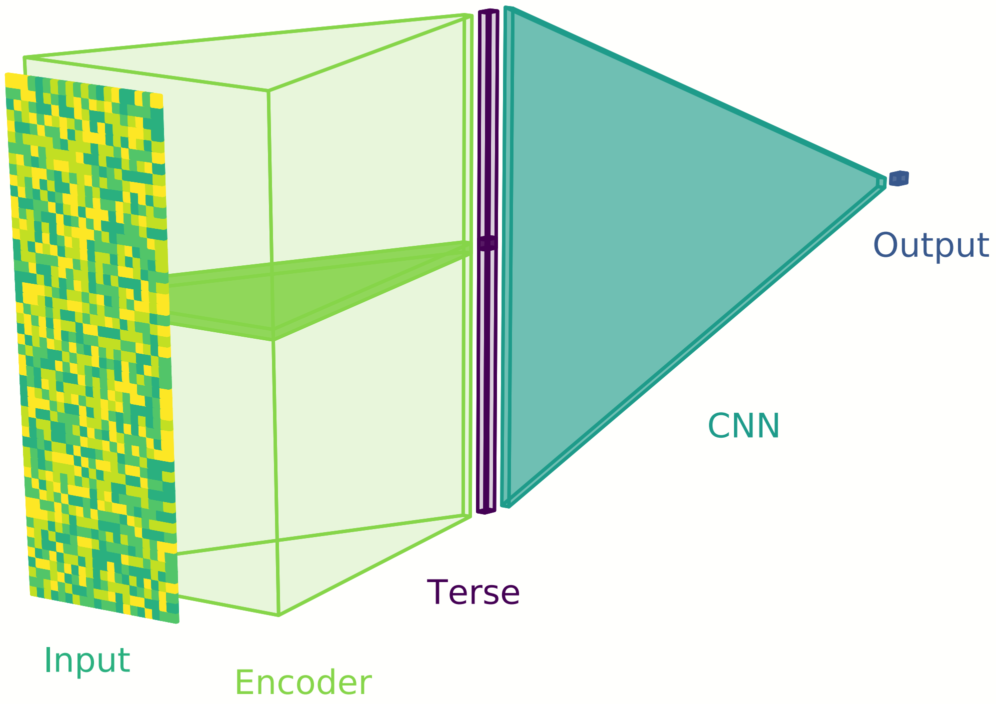


The verdict? For the "Suspicious Results," ML cheated.
The "Suspicious Results" are not robust. The model will not generalize to real observations because it depends on a simulation artifact.

Terse Value Correlations

Big idea: visualize how the terse layer correlates with cluster parameters.

Build trust in the model by asking “Is the terse layer summarizing sensible features?”



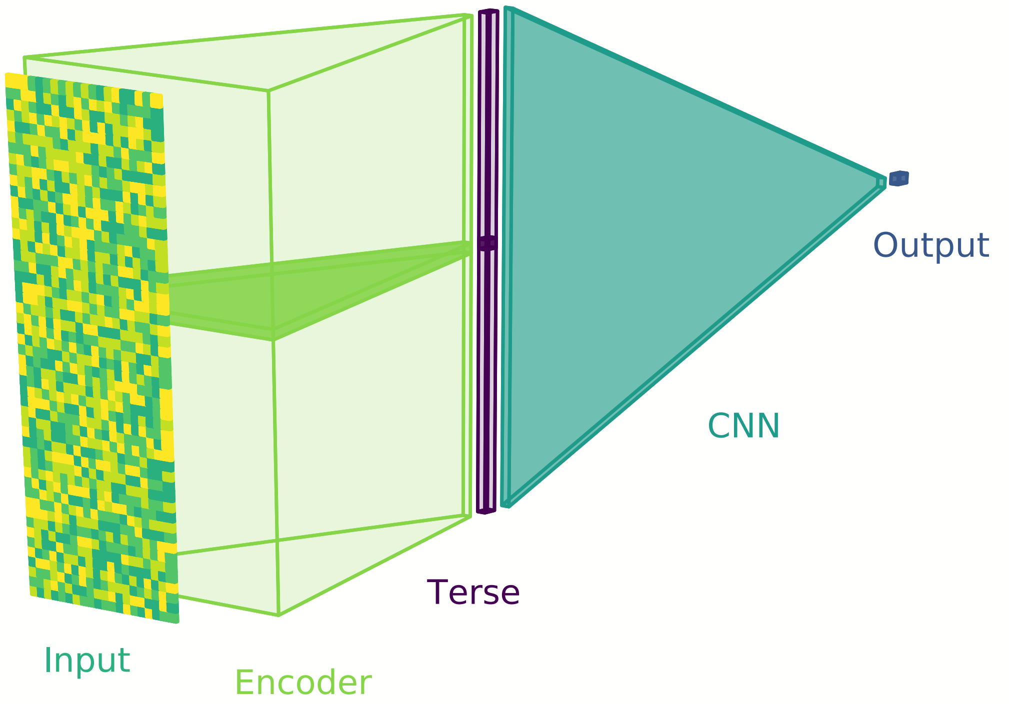
2. Saliency → Discovery

Saliency Maps

Big idea: assess the importance of each cluster feature by looking at gradients in the encoder network. “How does changing the input change the terse value?”

Develop a physical framework for understanding surprising results.

Simonyan+ 2014

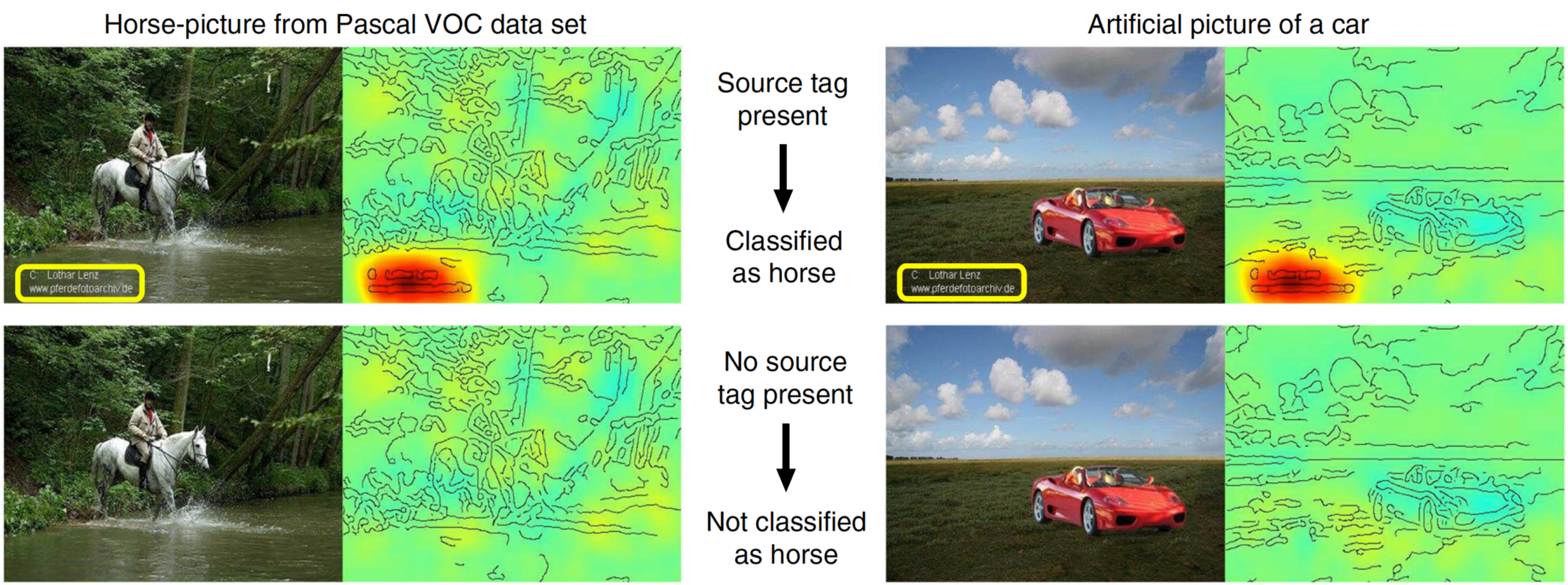


Saliency Maps

How can you tell that this is a horse? How can ML tell that this is a horse?

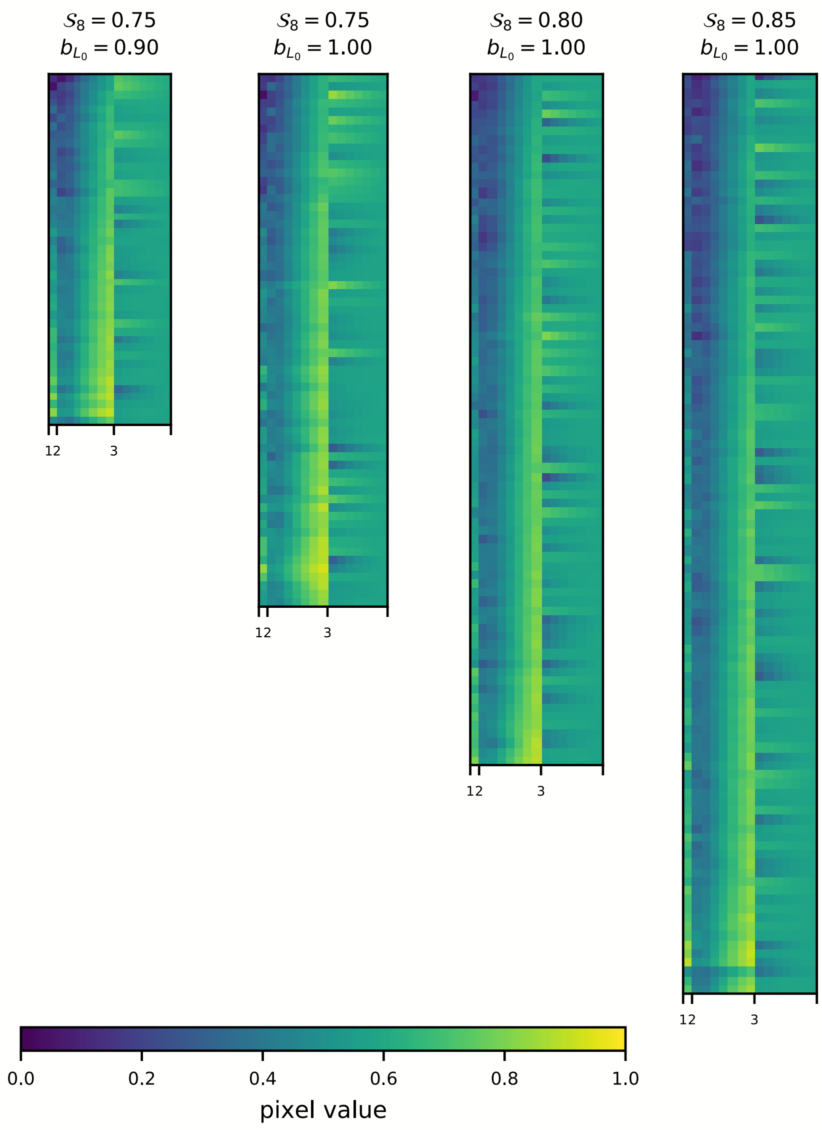
Saliency Maps

Saliency Maps

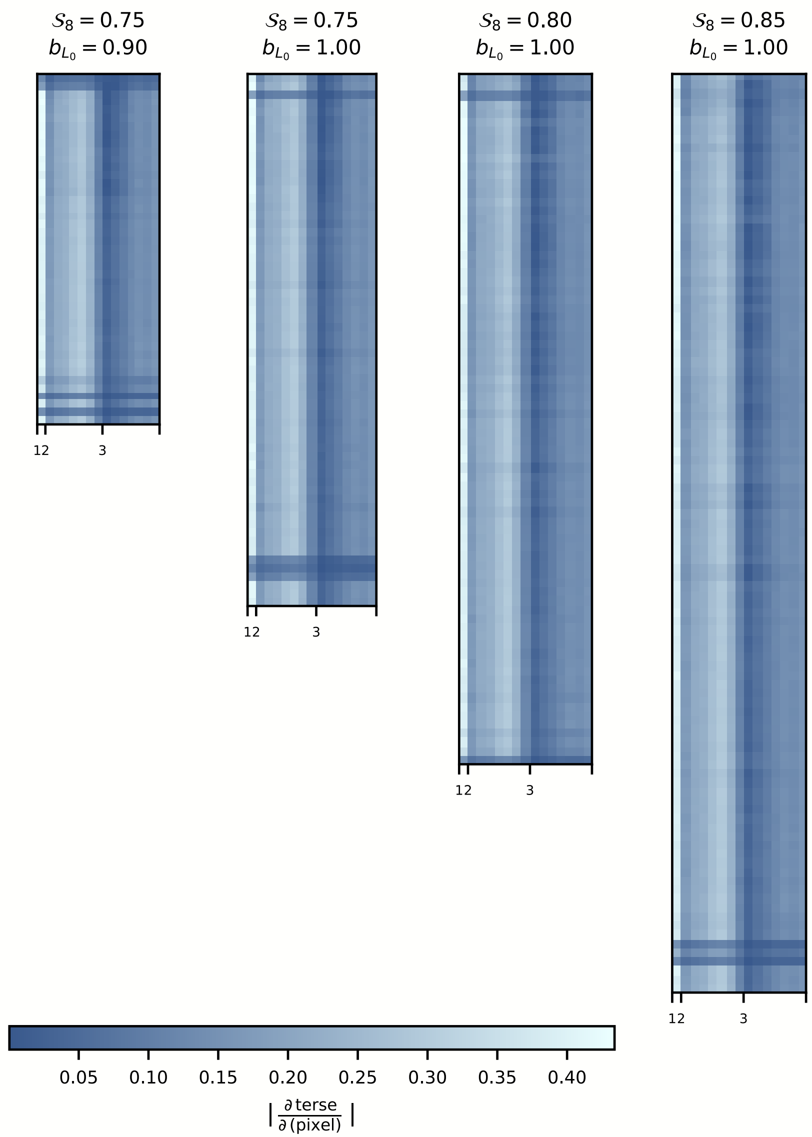


Lapuschkin et al., 2019

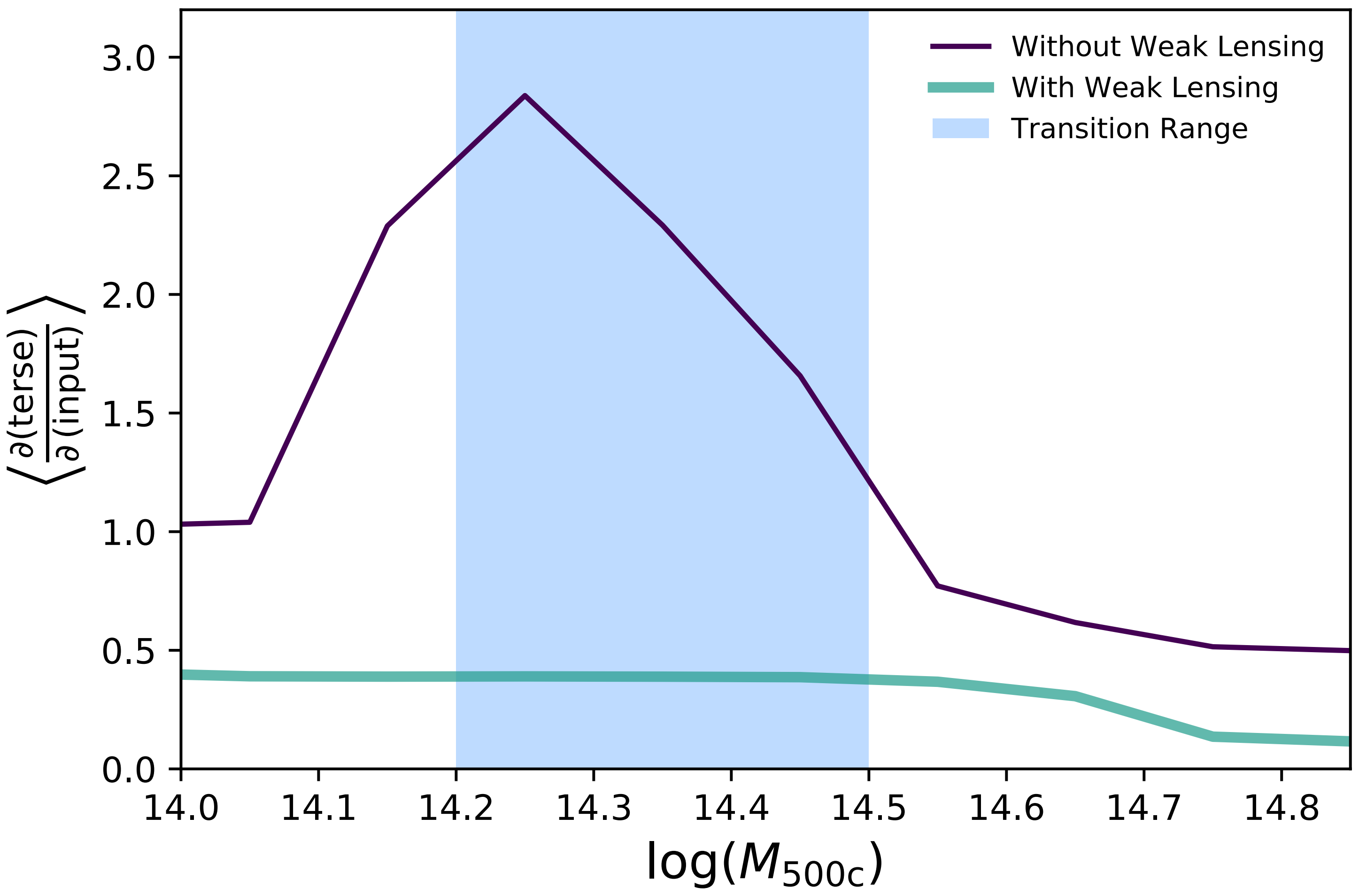
Input



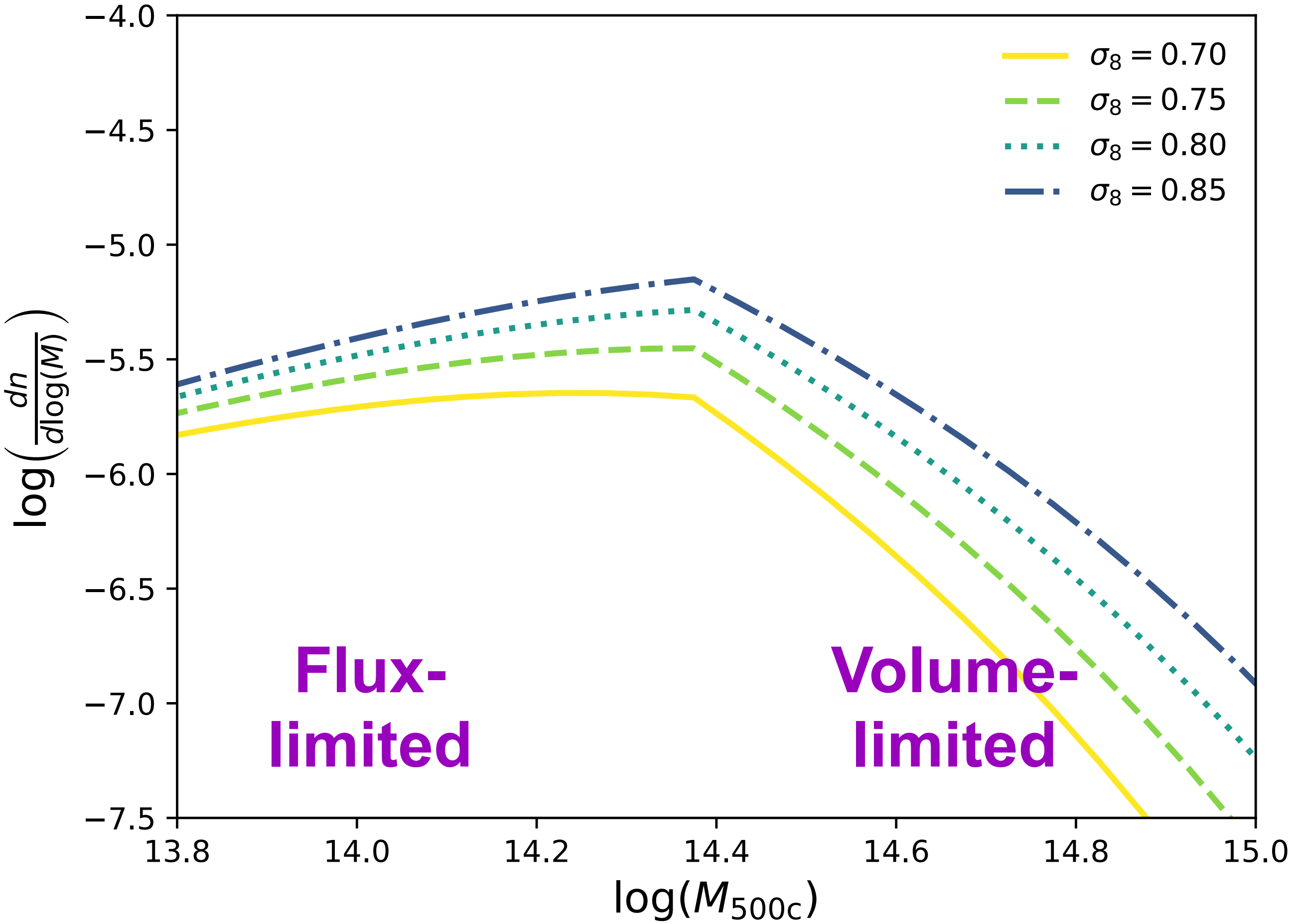
Saliency Maps



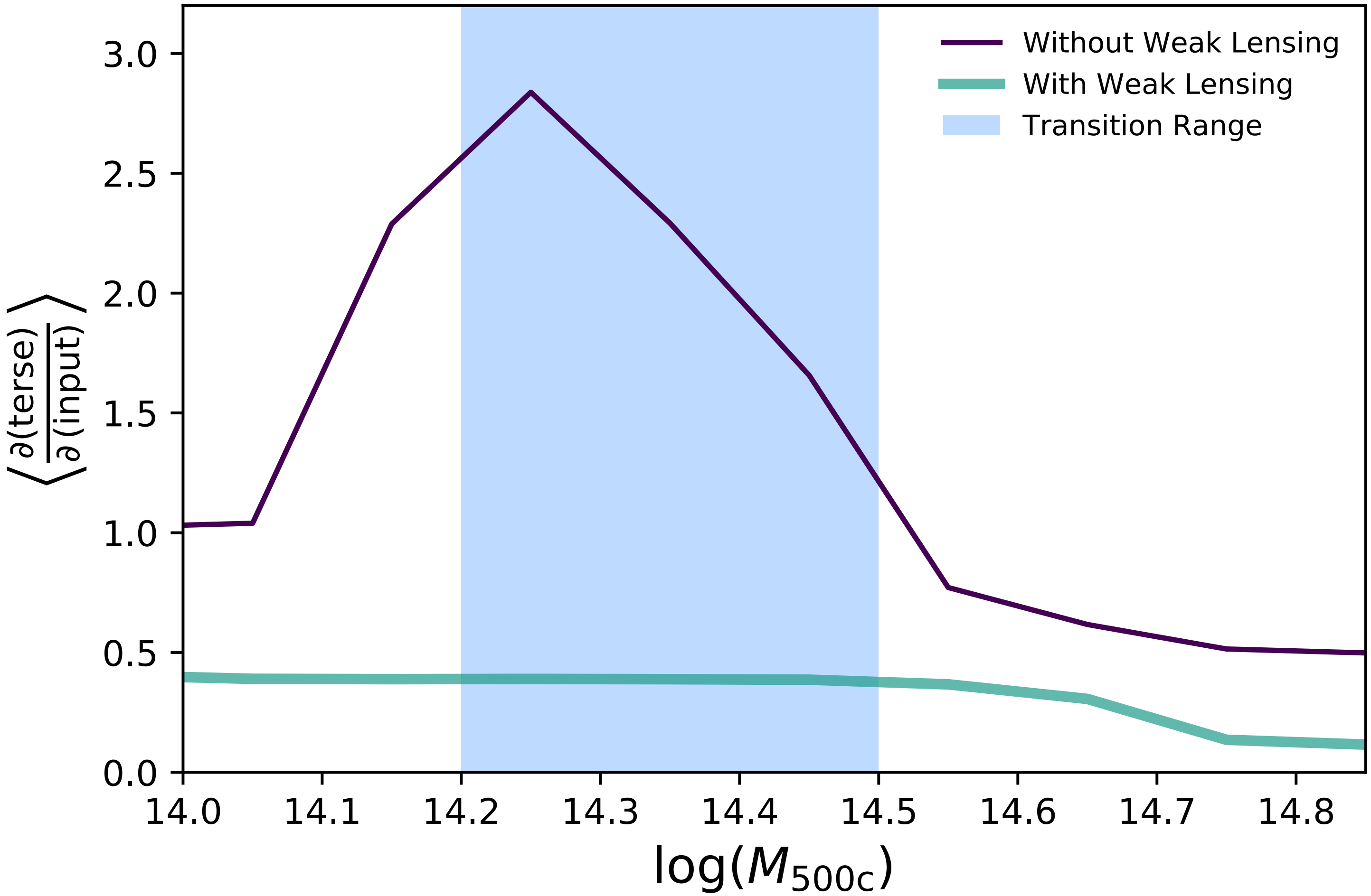
Saliency Trend: A Smoking Gun



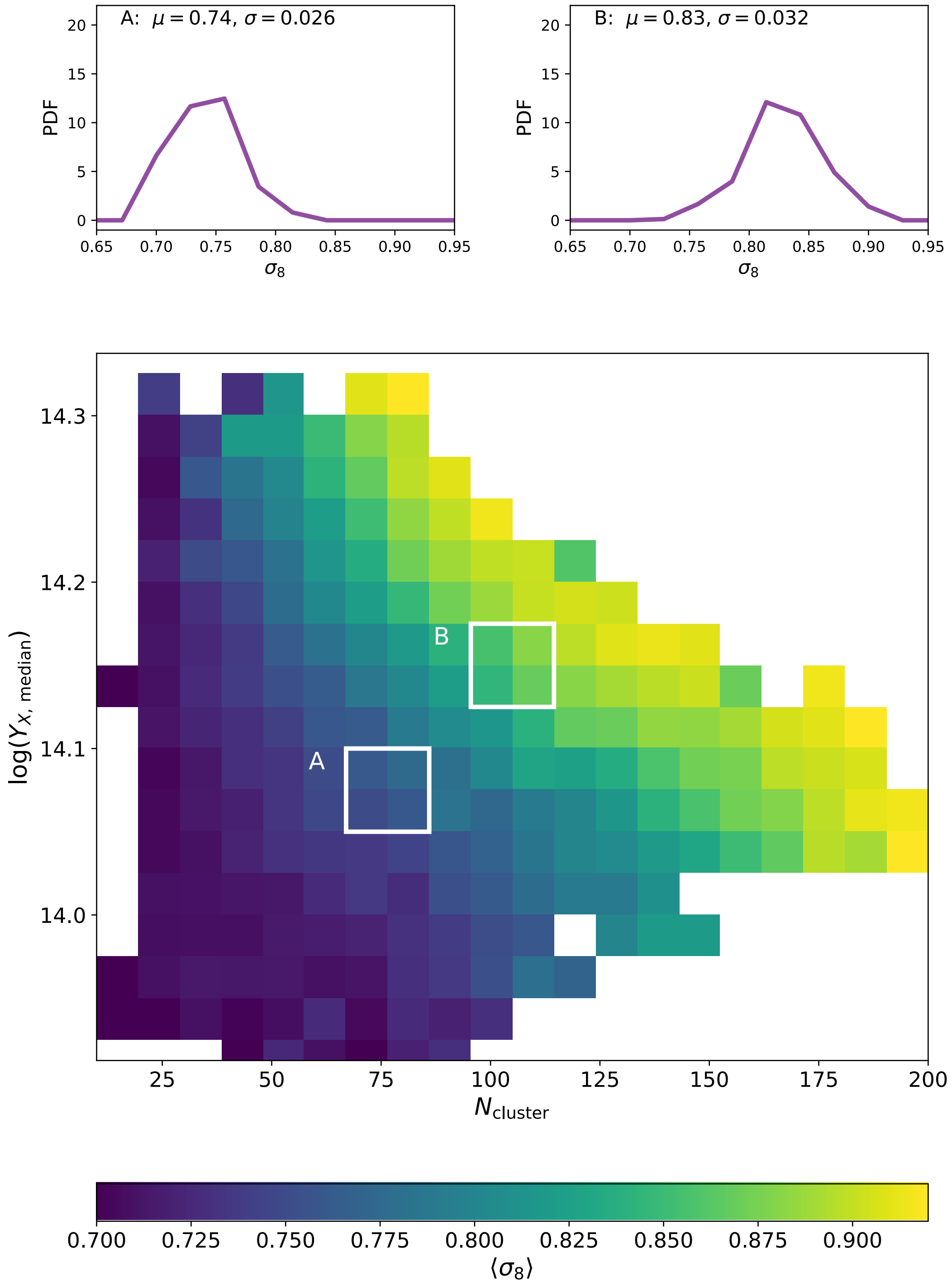
Many Cosmologies from One Simulation



Saliency Trend: A Smoking Gun

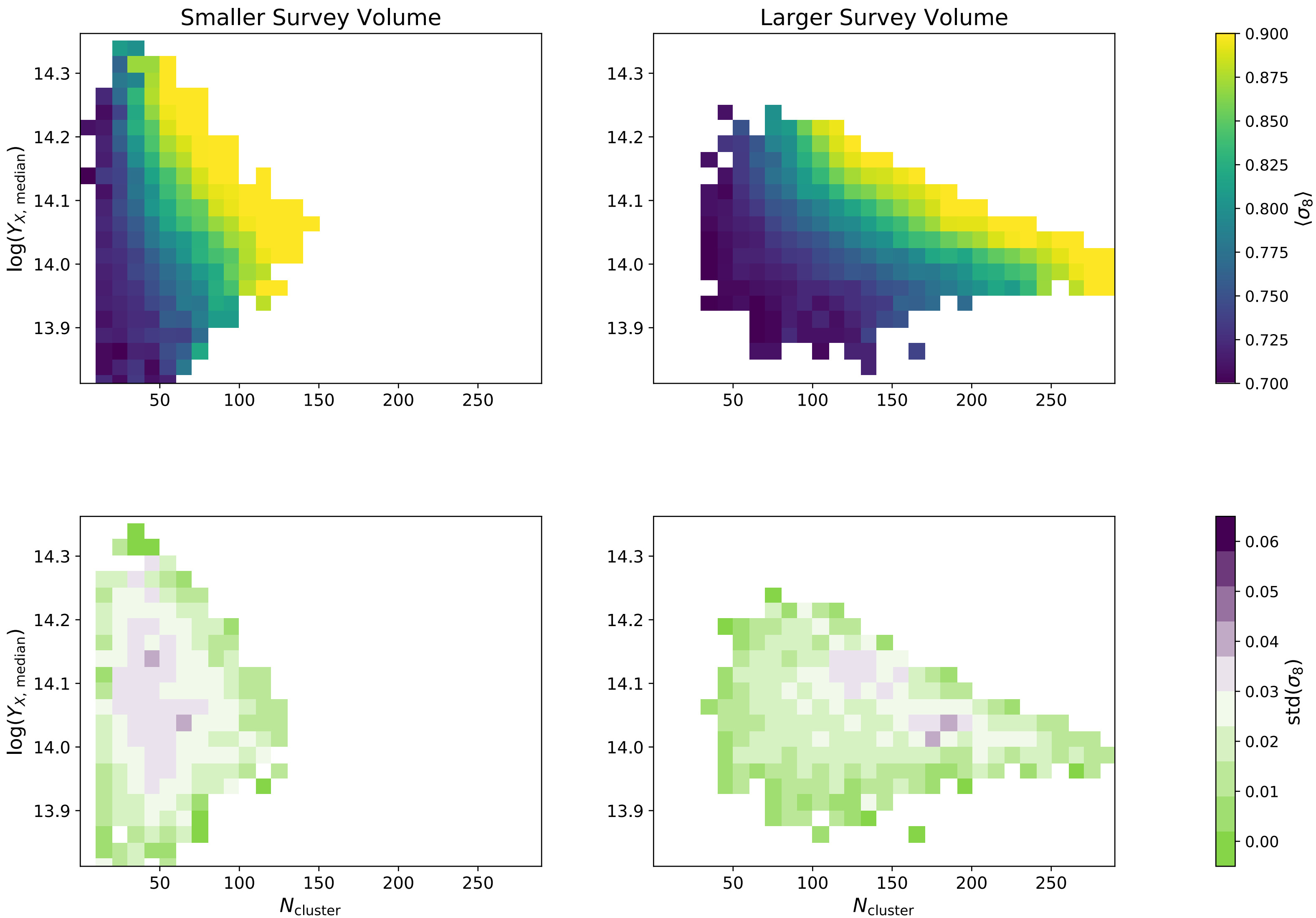


Is the model finding the transition point between the flux- and volume-limited samples to calibrate cluster masses without weak lensing?

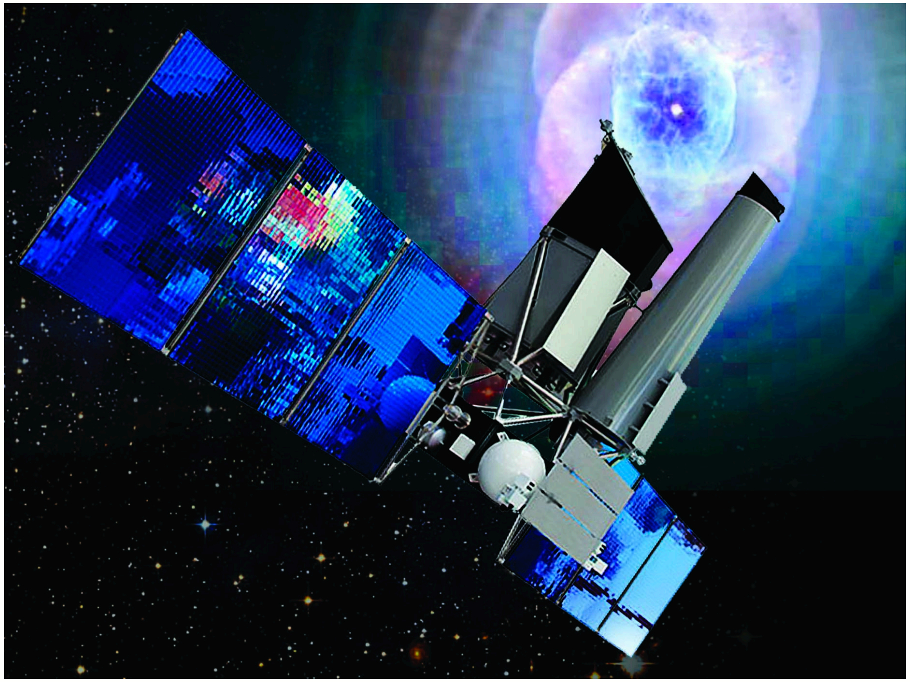


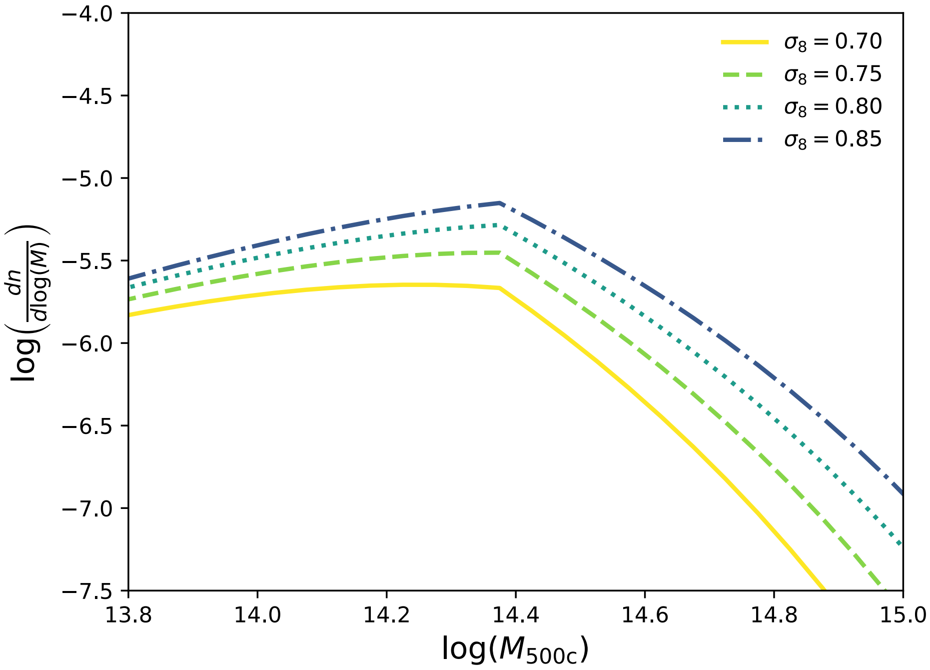
A: $\mu = 0.74$, $\sigma = 0.026$
B: $\mu = 0.83$, $\sigma = 0.032$
PDF
20
15
10
5
0
0.65
0.70
0.75
0.80
0.85
0.90
0.95
σ_8
$\log(Y_{X,\,\mathrm{median}})$
14.3
14.2
14.1
14.0
A
B
25
50
75
100
125
150
175
200
N_{cluster}
0.700
0.725
0.750
0.775
0.800
0.825
0.850
0.875
0.900
$\langle\sigma_8\rangle$

Survey-Dependent



ML-aided discovery of a self-calibration mode for eROSITA



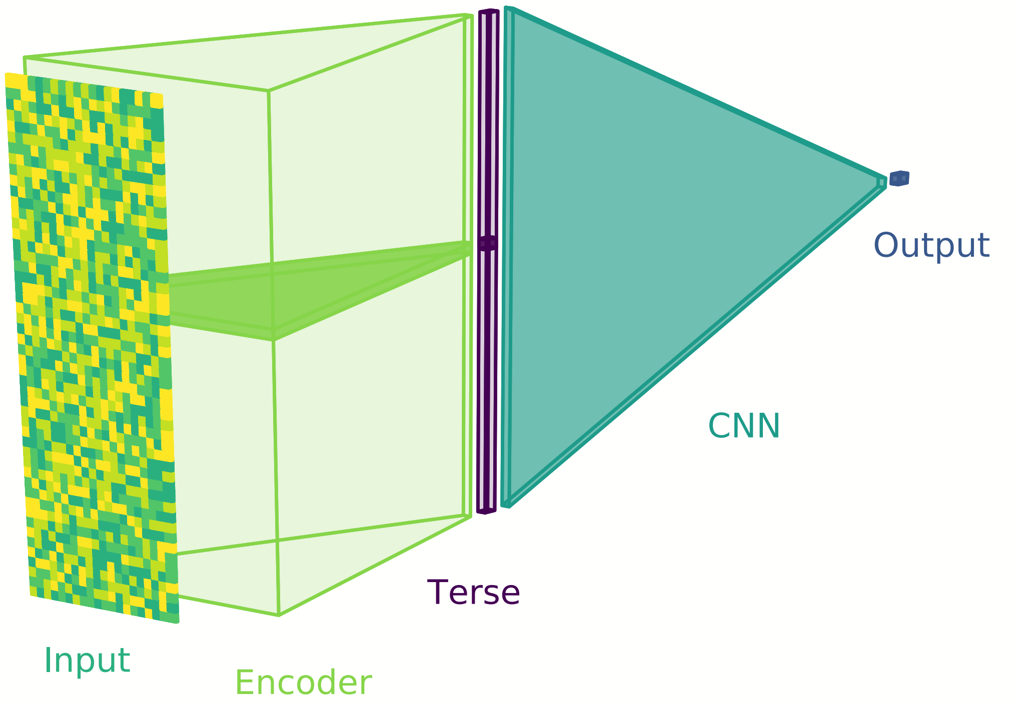


Saliency Maps

Big idea: assess the importance of each cluster feature by looking at gradients in the encoder network. “How does changing the input change the terse value?”

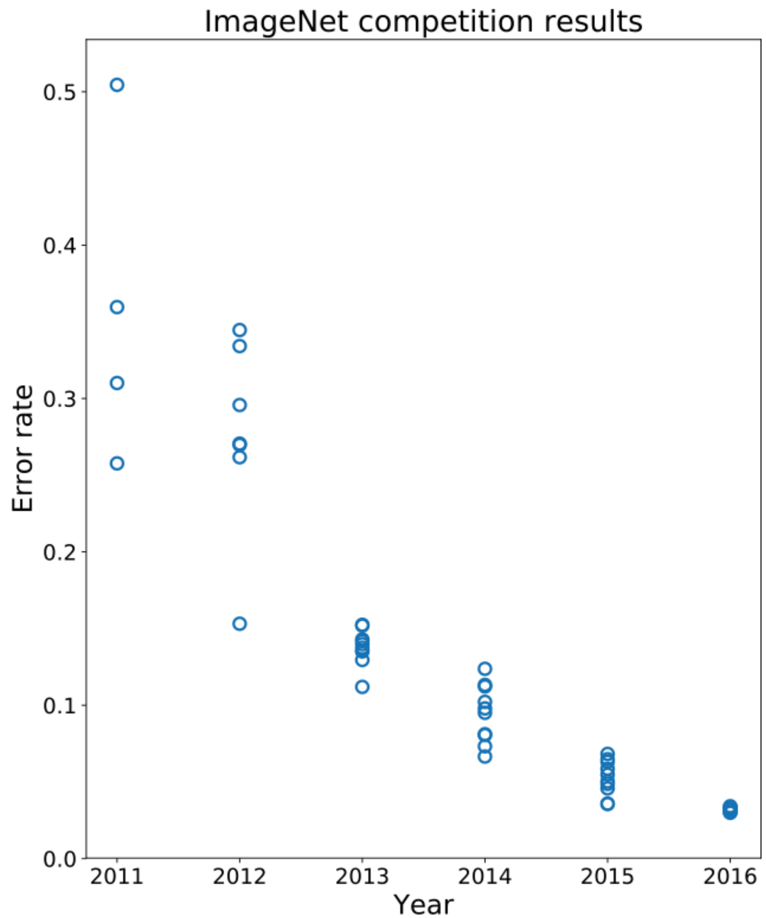
Develop a physical framework for understanding surprising results.

Simonyan+ 2014



What role will ML play in the future of astronomy?

ML & Astronomy can – and should! – move forward together.



Plot credit: Gkrusze, public domain

ML & Astronomy can – and should! – move forward together.

HELLO UNIVERSE

archive.stsci.edu/hello-universe/

A framework for testing and benchmarking machine learning methods on astronomical data

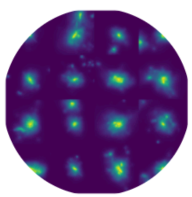
Hello Universe is a new project at MAST designed to help astronomers develop machine learning (ML) methods for astronomical discovery. ML will be an essential tool for analyzing the rich data sets of the upcoming decade, and *Hello Universe* provides a framework for testing ML algorithms and new techniques. Each entry in the *Hello Universe* collection includes:

- **Data**: a high-level science product (HLSP) data set for testing and benchmarking ML algorithms
- **Code**: a tutorial Jupyter notebook that provides step-by-step examples of how to apply an ML technique to the data

Though these data sets are motivated by the needs of a novice data science learner, they are sufficient for a wide range of tasks. *Hello Universe* entries include examples of:

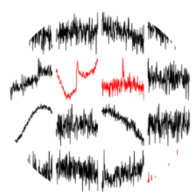
- analyzing 2D (image) and 1D (vector or light curve) data sets.
- applying techniques for regression and for classification.
- developing supervised and unsupervised learning models.
- using best practices for training and optimizing models.
- selecting metrics for assessing model performance.

HELLO UNIVERSE



Classifying JWST/HST galaxy mergers with CNNs

neural networks | 2d data | classification | overfitting | confusion matrix

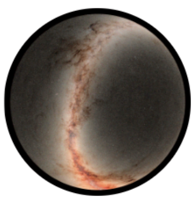


Classifying TESS stellar flares with CNNs

neural networks | 1d data | classification | prediction

Predicting 3D-HST redshift with decision trees

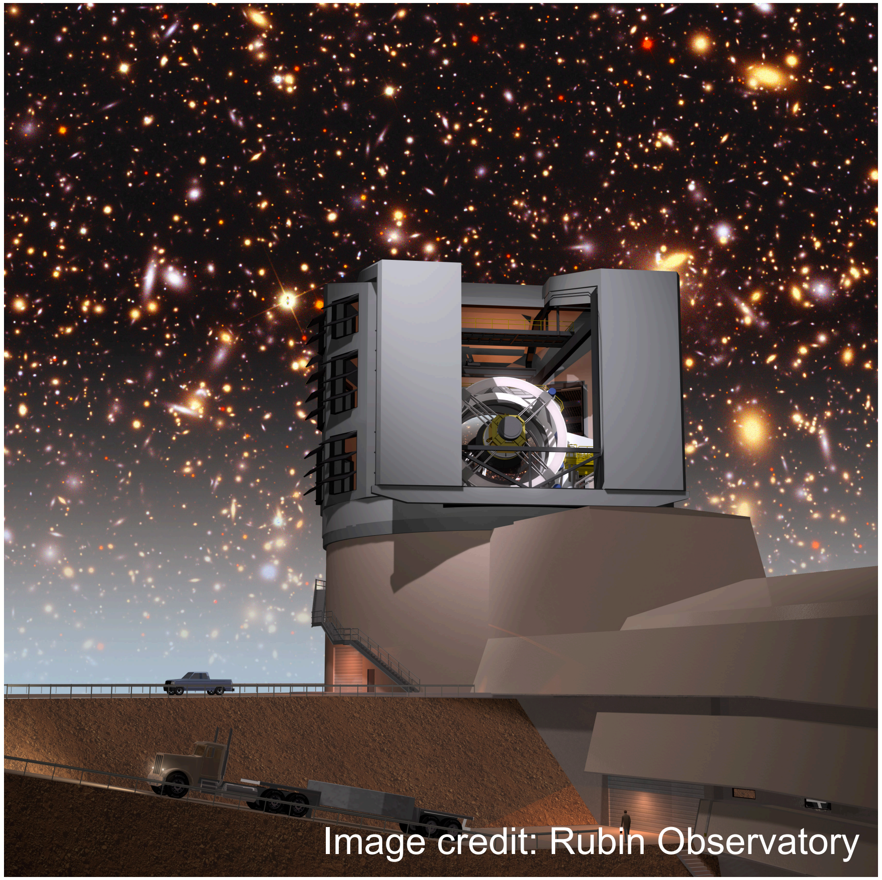
decision trees | 1d data | regression | cross-validation

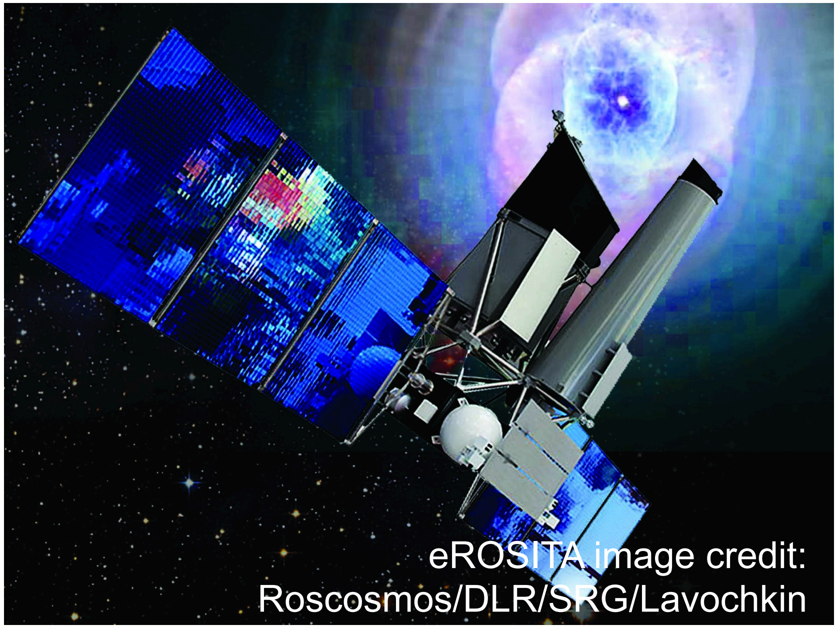


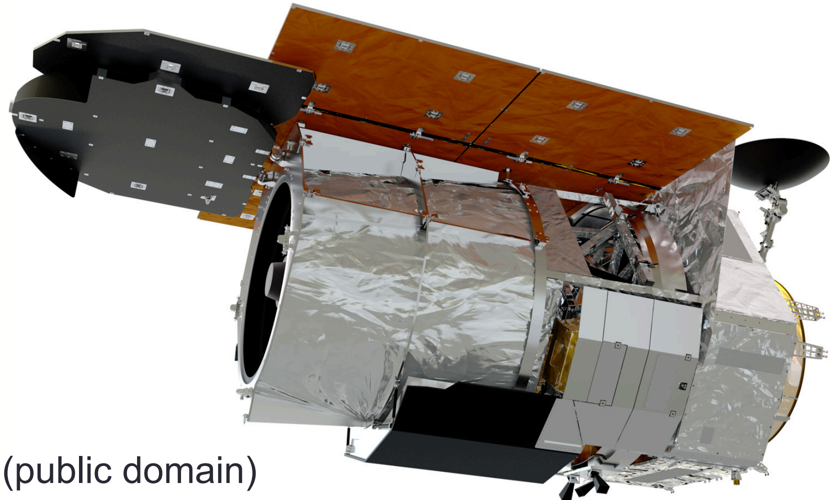
Classifying Pan-STARRS with (un)supervised learning

classification | 1d data | PCA | tSNE | k-means | SGD | unsupervised | supervised

Rich Upcoming Data Sets







Roman image credit: NASA (public domain)

Style Transfer

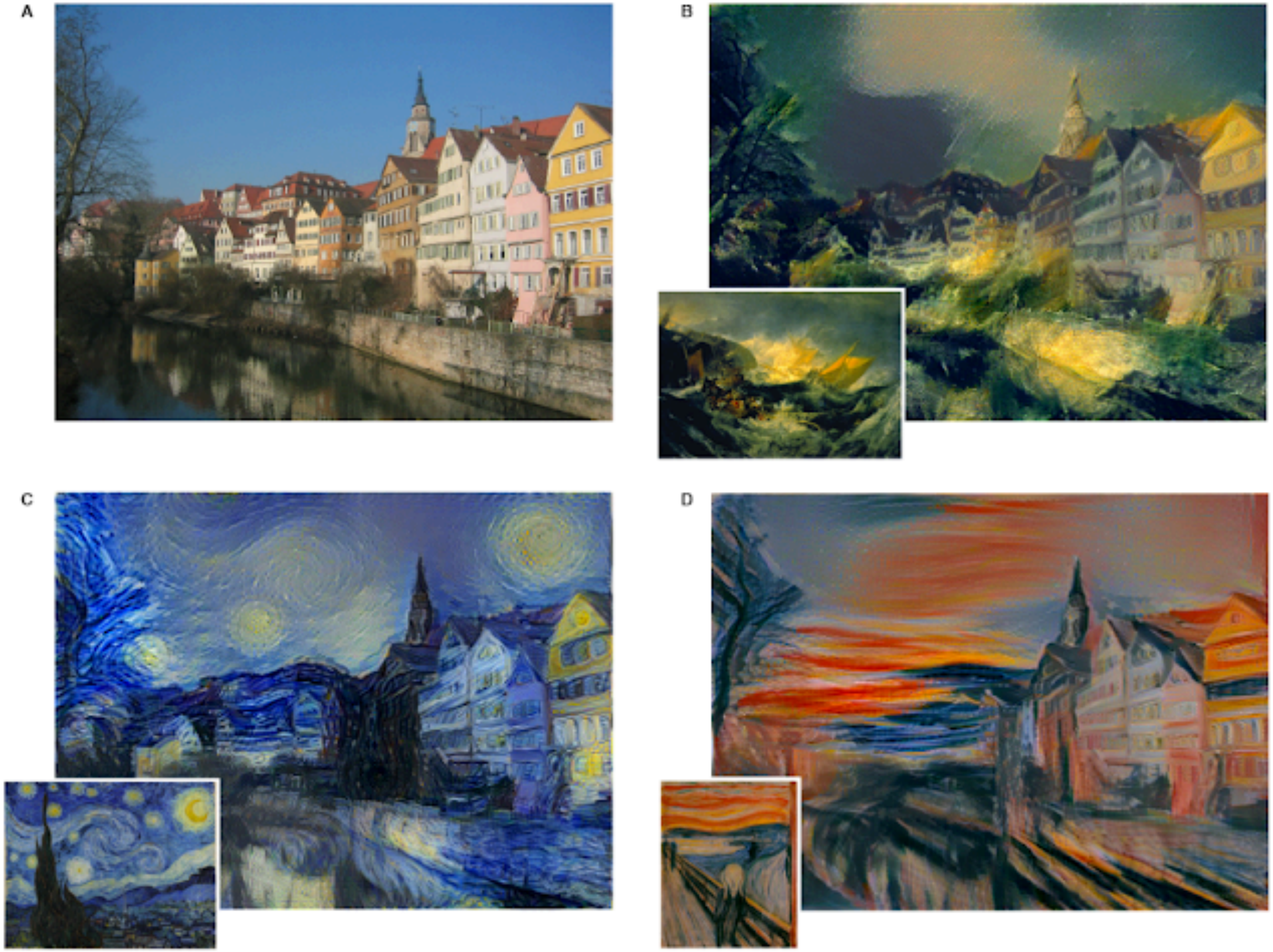


Image credit: Google AI blog, adapted from Gatys+ 2015

Pixel Recursive Super Resolution

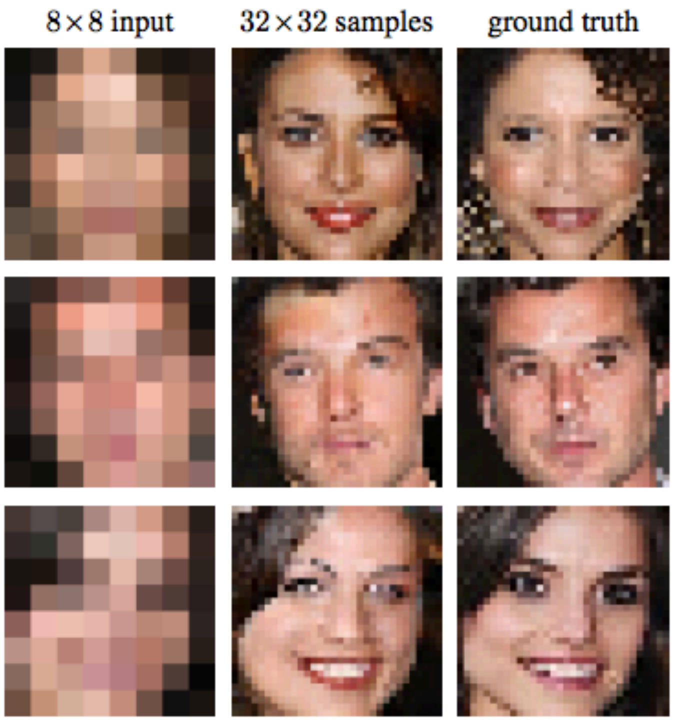


Figure 1: Illustration of our probabilistic pixel recursive super resolution model trained end-to-end on a dataset of celebrity faces. The left column shows 8×8 low resolution inputs from the test set. The middle and last columns show 32×32 images as predicted by our model *vs.* the ground truth. Our model incorporates strong face priors to synthesize realistic hair and skin details.

Dahl+ 2017

Bias in Training Data

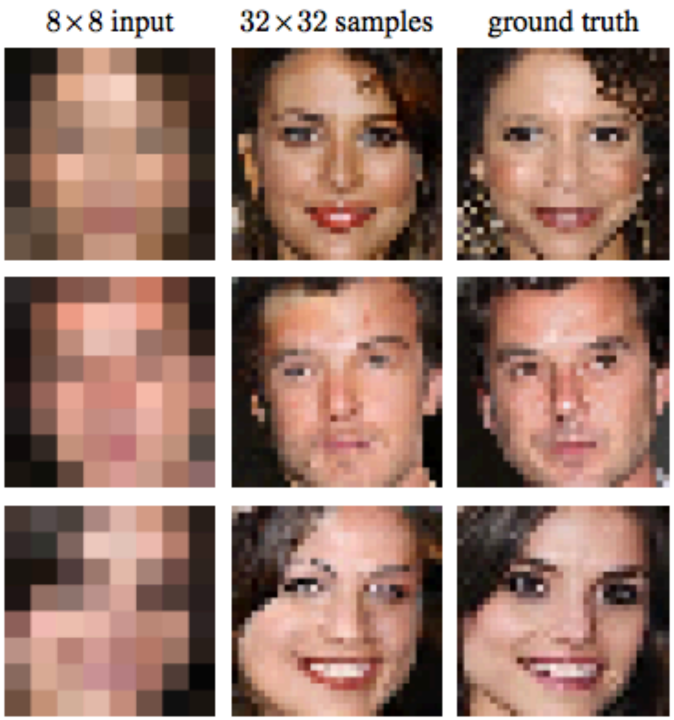
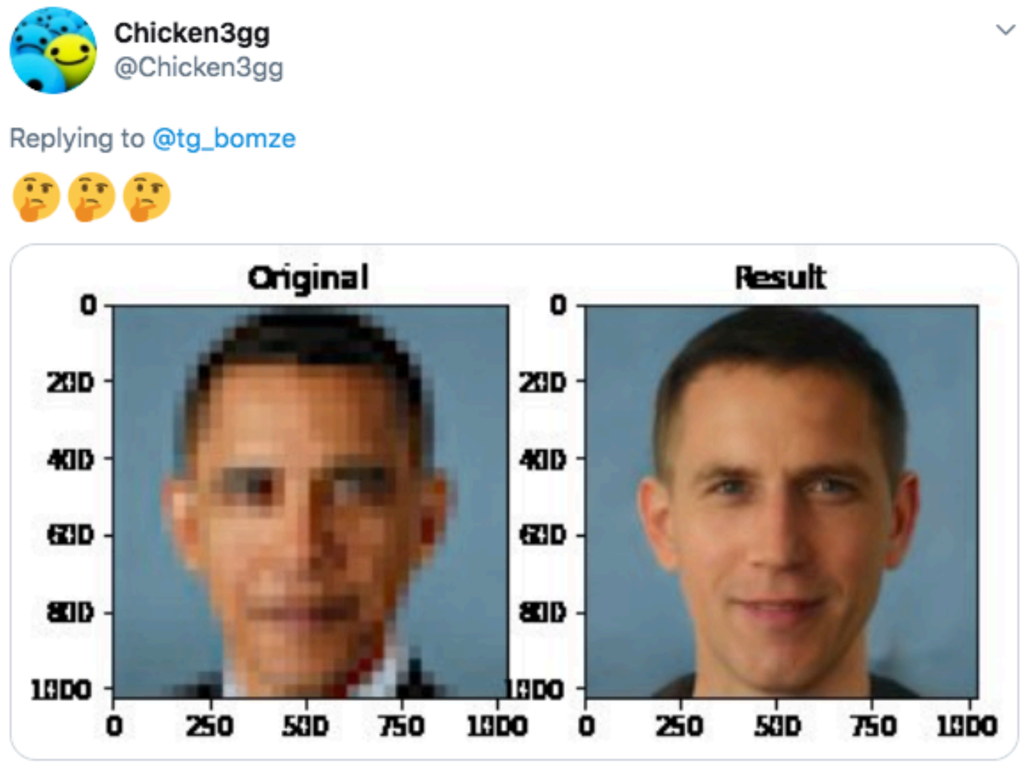
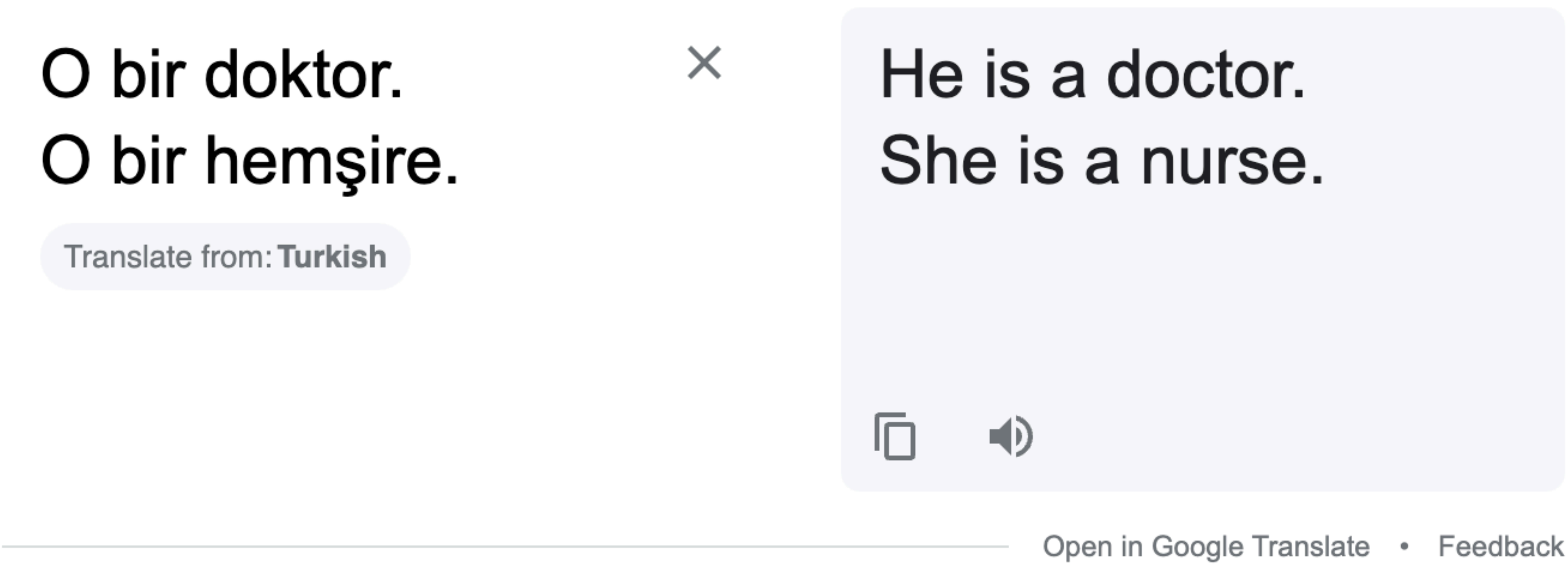


Figure 1: Illustration of our probabilistic pixel recursive super resolution model trained end-to-end on a dataset of celebrity faces. The left column shows 8×8 low resolution inputs from the test set. The middle and last columns show 32×32 images as predicted by our model *vs.* the ground truth. Our model incorporates strong face priors to synthesize realistic hair and skin details.



Dahl+ 2017

Bias in Language Translation



Caliskan+ 2017

Automating Human Bias

RETAIL OCTOBER 10, 2018 / 7:04 PM / UPDATED 4 YEARS AGO

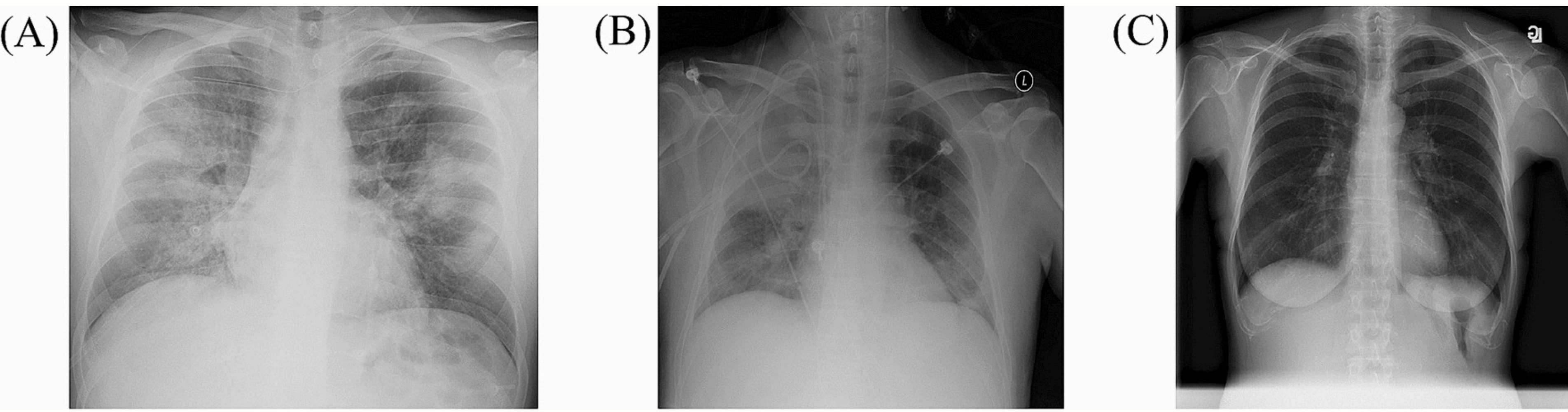
Amazon scraps secret AI recruiting tool that showed bias against women

By Jeffrey Dastin

8 MIN READ

SAN FRANCISCO (Reuters) - Amazon.com Inc's AMZN.O machine-learning specialists uncovered a big problem: their new recruiting engine did not like women.

Undesirable Learning Behaviors



Covid Pneumonia Healthy

image credit: Nishia et al., 2020

ARTIFICIAL INTELLIGENCE

Hundreds of AI tools have been built to catch covid. None of them helped.

Some have been used in hospitals, despite not being properly tested. But the pandemic could help make medical AI better.

By Will Douglas Heaven

July 30, 2021

Is ML the right tool for astronomy?

Machine Learning *can be* the right tool for astronomy:

- Engineer models that follow a human approach with checkpoints to make sure that the model is learning something sensible.

- Interrogate and interpret models.

- Approach high-accuracy results with scrutiny.

- Treat ML as a tool to be used in partnership with traditional statistical methods & human exploration.

1. Can ML be trusted?

Terse value correlations: to assess whether the model will generalize.

Building Trustworthy ML Models for Astronomy

Ntampaka, Ho, & Nord 2021, 2111.14566

2. Can it drive physical understanding?

Saliency maps: to identify what part of the cluster carries the most cosmological information.

The Importance of Being Interpretable

Ntampaka & Vikhlinin 2022, 2112.05768

3. What role will ML play in the future of astronomy?

ML in partnership – not in competition! – with traditional methods.

The Role of ML in the Next Decade of Cosmology

Ntampaka+ 2019, 1902.10159

Hello Universe

archive.stsci.edu/hello-universe

mntampaka@stsci.edu