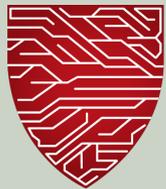


Barriers & Opportunities for ML in Astronomy

Michelle Ntampaka



HDSI | Harvard Data
Science Initiative

CENTER FOR

ASTROPHYSICS

HARVARD & SMITHSONIAN



Barriers to ML adoption

Barrier #1: Missing common vocabulary.

Data Science is

but is often misinterpreted

Statistics &
Astrostats

Applied Math

Machine
Learning

Computer
Science

- big data
- data storage
- data dissemination
- data analysis

Barrier #2: Distrust of the unknown.



Barrier #2: Distrust of the unknown.

What is machine learning?

ML research explores the development and application of algorithms that find patterns in data.

In the context of astronomy, ML algorithms can be used to:

- describe complicated relationships
- identify data clusters and data outliers
- reduce scatter by using complex or subtle signals
- generate simulated data
- classify objects
- address sparse data
- understand error bars
- interpret the physical underpinning

Barrier #3: Steep learning curve.



Shit Academics Say

@AcademicsSay

Following



Let's be honest: we are all two drinks away from being interdisciplinary.

10:37 AM - 22 Mar 2018

1,833 Retweets 5,820 Likes



39 1.8K 5.8K



Tweet your reply



Brian Berry @BrianDavidBerry · 22 Mar 2018

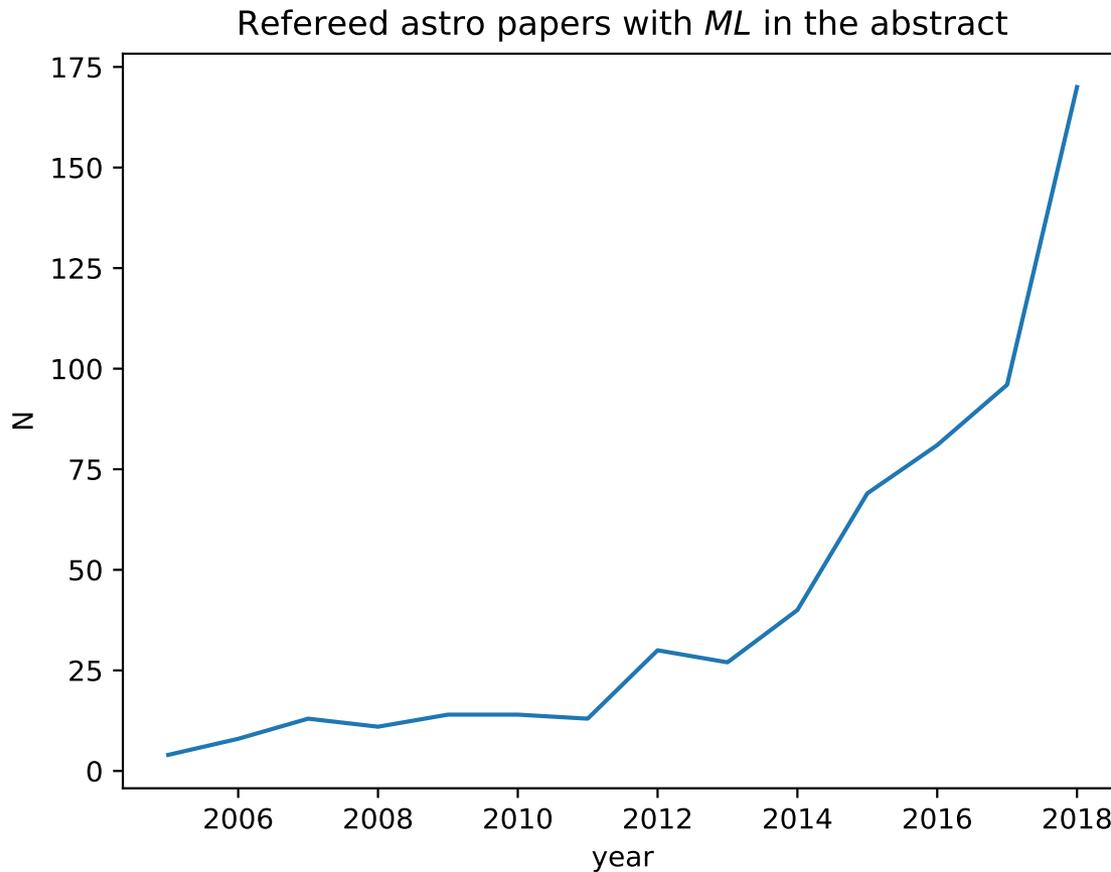


Replying to @AcademicsSay

I assure you, once you are interdisciplinary... you'll need many more drinks.

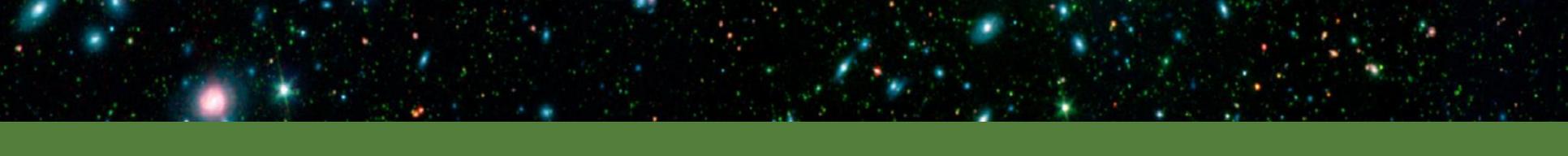
1 3 67

Barrier #4: Unmet demand for interdisciplinary leadership.



Barrier #5: Need to establish standards for interpretability and failure modes.





What role will ML play
in the future of astronomy?

SPACE

Faced with a Data Deluge, Astronomers Turn to Automation

For better or worse, machine learning and big data are poised to transform the study of
the heavens

By Anil Ananthaswamy on August 21, 2019

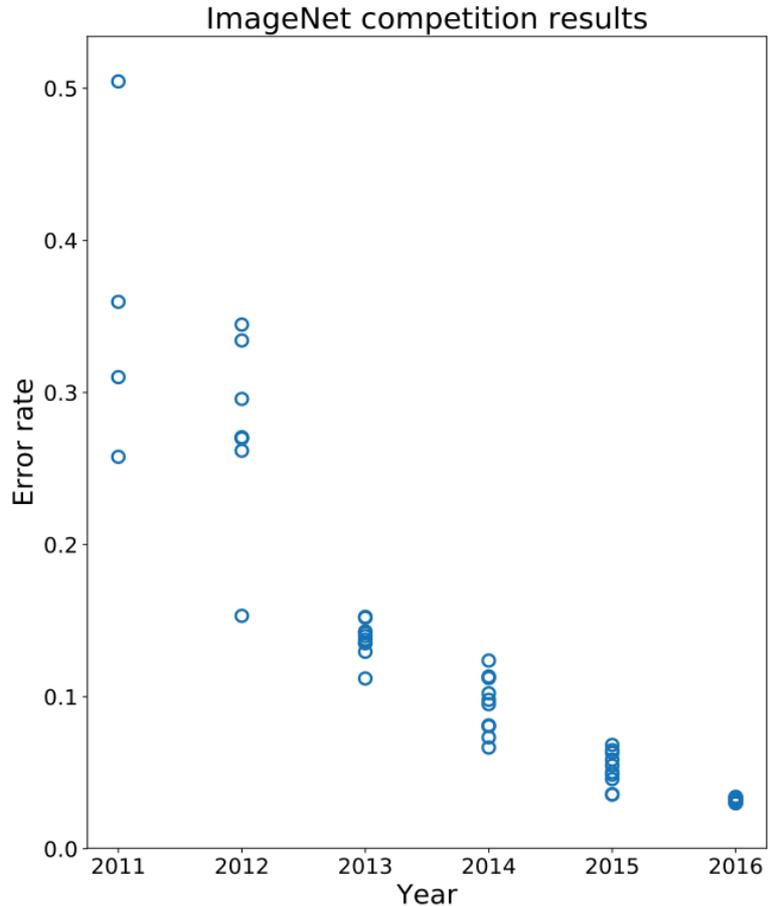
Astronomy is the ideal sandbox for data science.

- Minimal privacy concerns.
- Culture of sharing data.
- Well-posed questions.
- Non-monetizable.

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

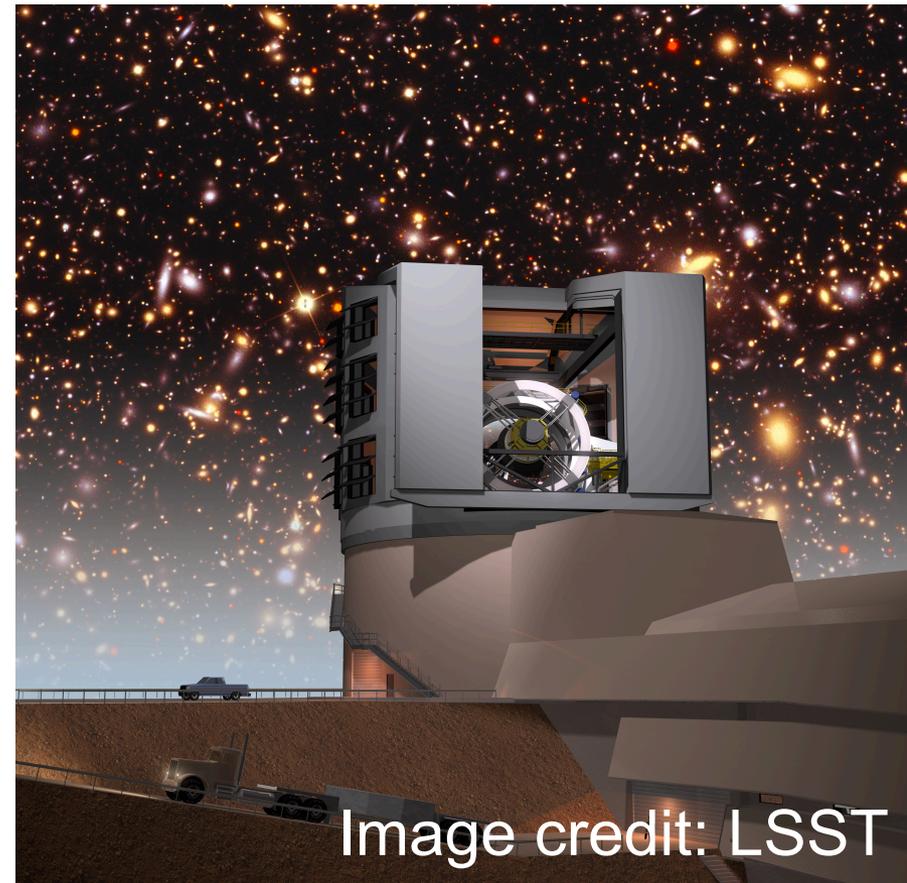
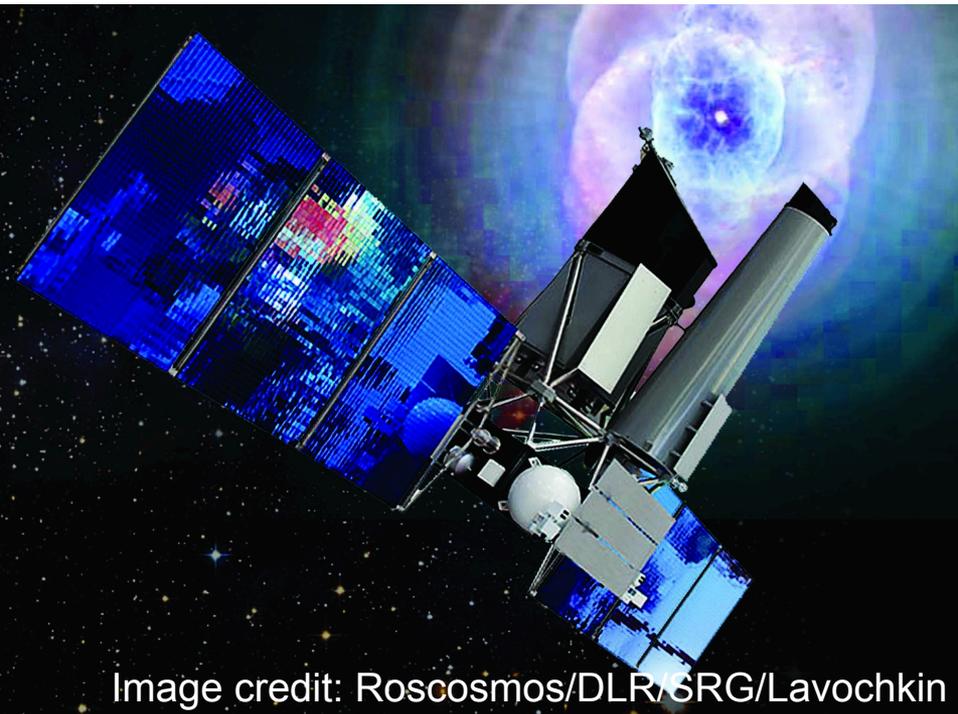


XKCD, 2014



Plot credit: Gkrusze, public domain

Opportunities



There will always be tasks that
require human ingenuity.
(These are not them.)



Barriers to interdisciplinary research:

1. Missing common vocabulary.
2. Distrust of the unknown.
3. Steep learning curve.
4. Unmet demand for interdisciplinary leadership.
5. Need to establish standards for interpretability and failure modes.

The Role of Machine Learning in the Next Decade
of Cosmology 1902.10159

Discussion Questions

1. Do you agree with the barriers list? What other barriers do you see to astro+ML research? What issues did I miss? How do we address these barriers and overcome them?
2. What opportunities are you most excited about? Data sets? Methodological advancements? What tasks might never be automated?
3. How can we be good advocates for interdisciplinary research? How should we deal with opponents of ML? How do we know that ML is not just a phase?

In 20 minutes: 3 sentence, “food for thought” conclusion.
Table scribe = person with the next birthday.