

Active Learning for Supernova Photometric Classification

*LPC Clermont – 23 April 2018
1st meeting of Machine Learning at Cezeaux*

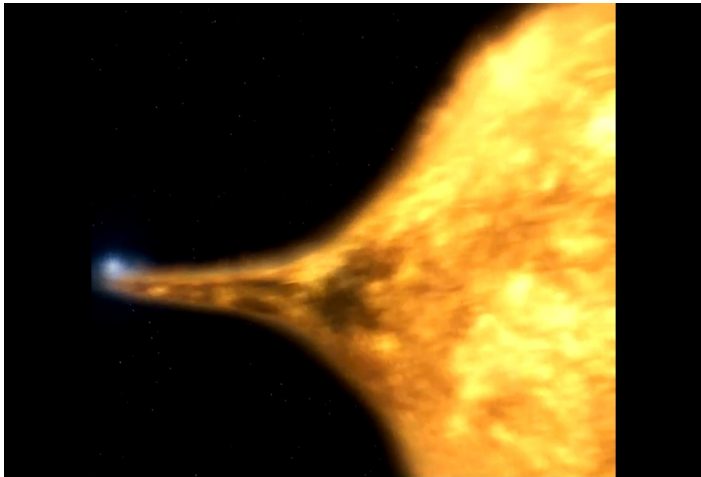
Emille E. O. Ishida

*CNRS/Laboratoire de Physique de Clermont - Université Clermont Auvergne
Clermont Ferrand, France*

Supernovae

Supernovae

Supernovae: Type Ia



<https://www.spacetelescope.org/videos/heic0415b/>

Supernovae

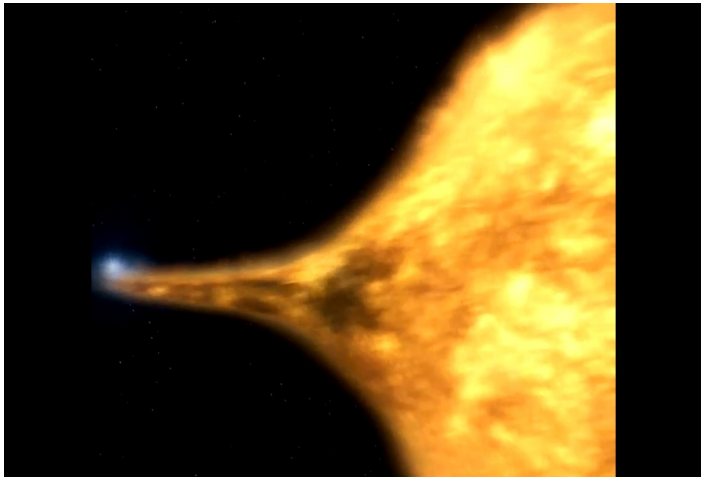
Supernovae: Type II



<https://www.youtube.com/watch?v=ePJAK9zVUw0>

Supernovae

Supernovae: Type Ia

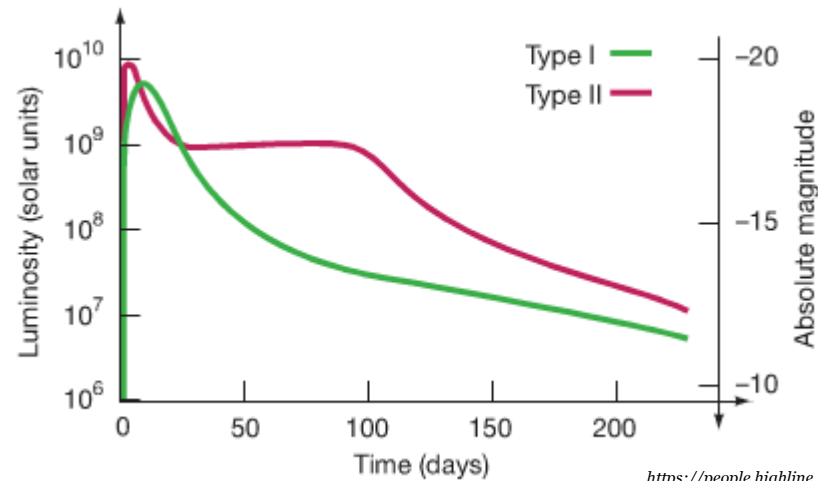


Supernovae: Type II



<https://www.spacetelescope.org/videos/heic0415b/>

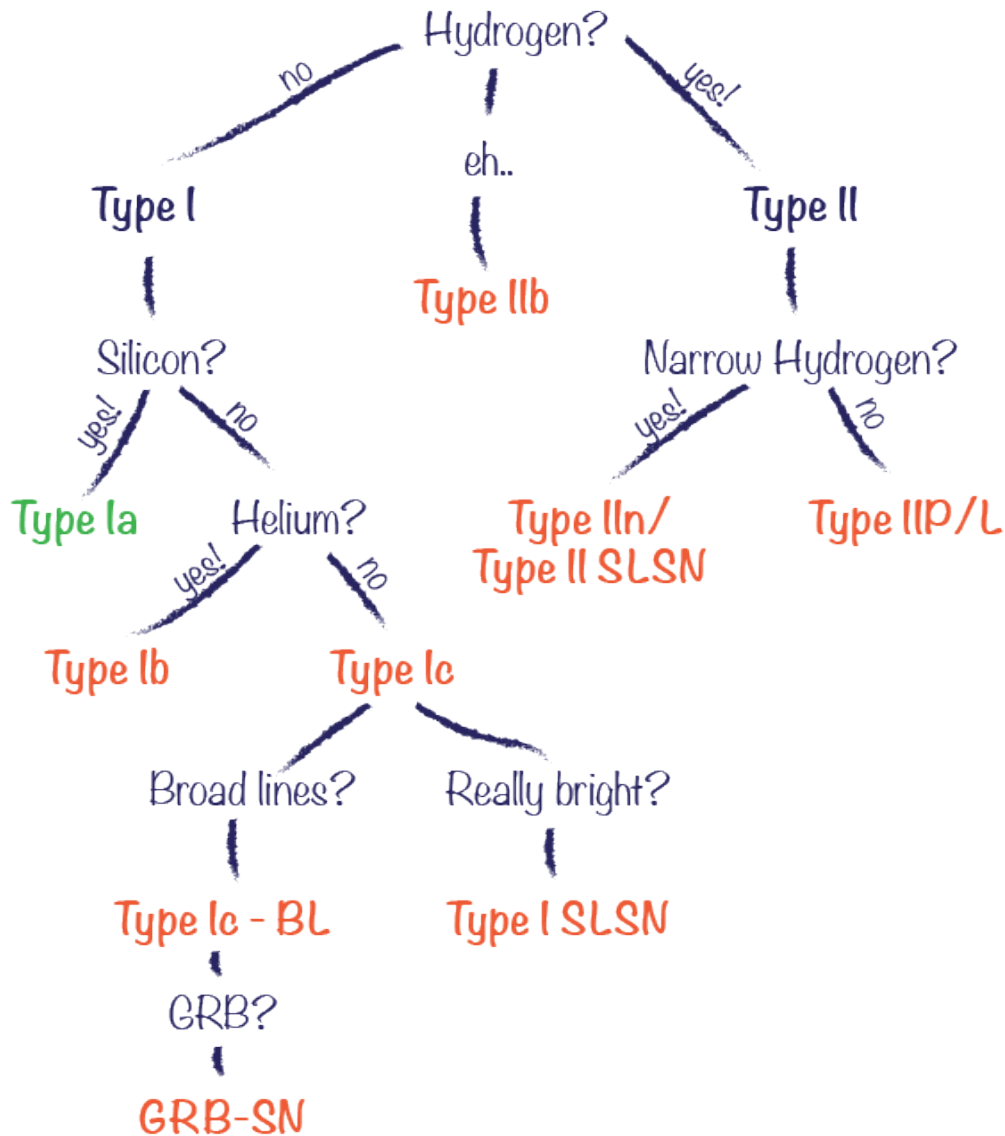
<https://www.youtube.com/watch?v=ePJAK9zVUw0>



<https://people.highline.edu/iglozman/classes/astronotes/media/sncurves.gif>

Supernovae

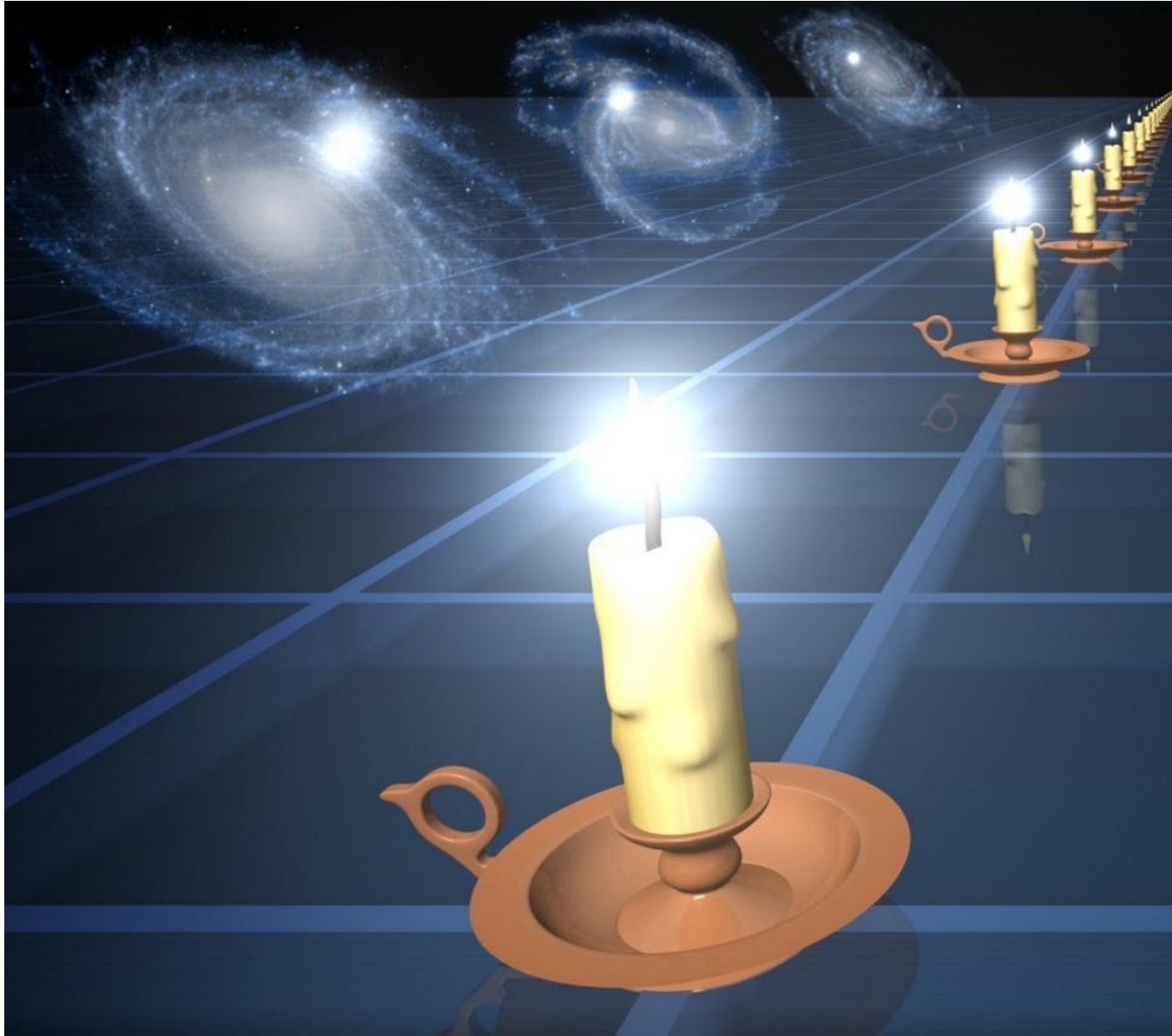
The Supernova Zoo



it's complicated

Supernovae: cosmological application

Only Ia can be used as distance indicators



2011



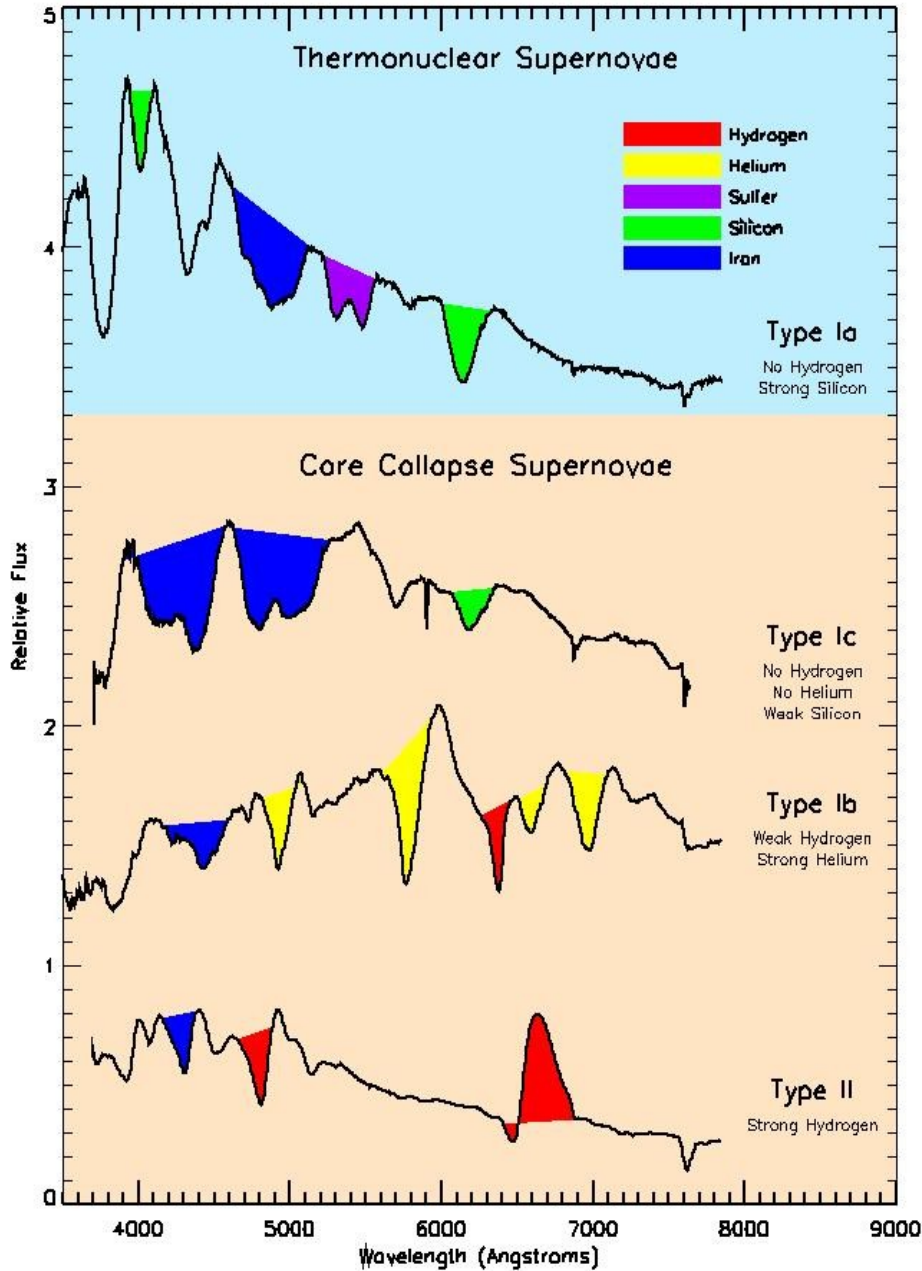
How are they
classified?

Spectroscopy



 TIME
CONSUMING

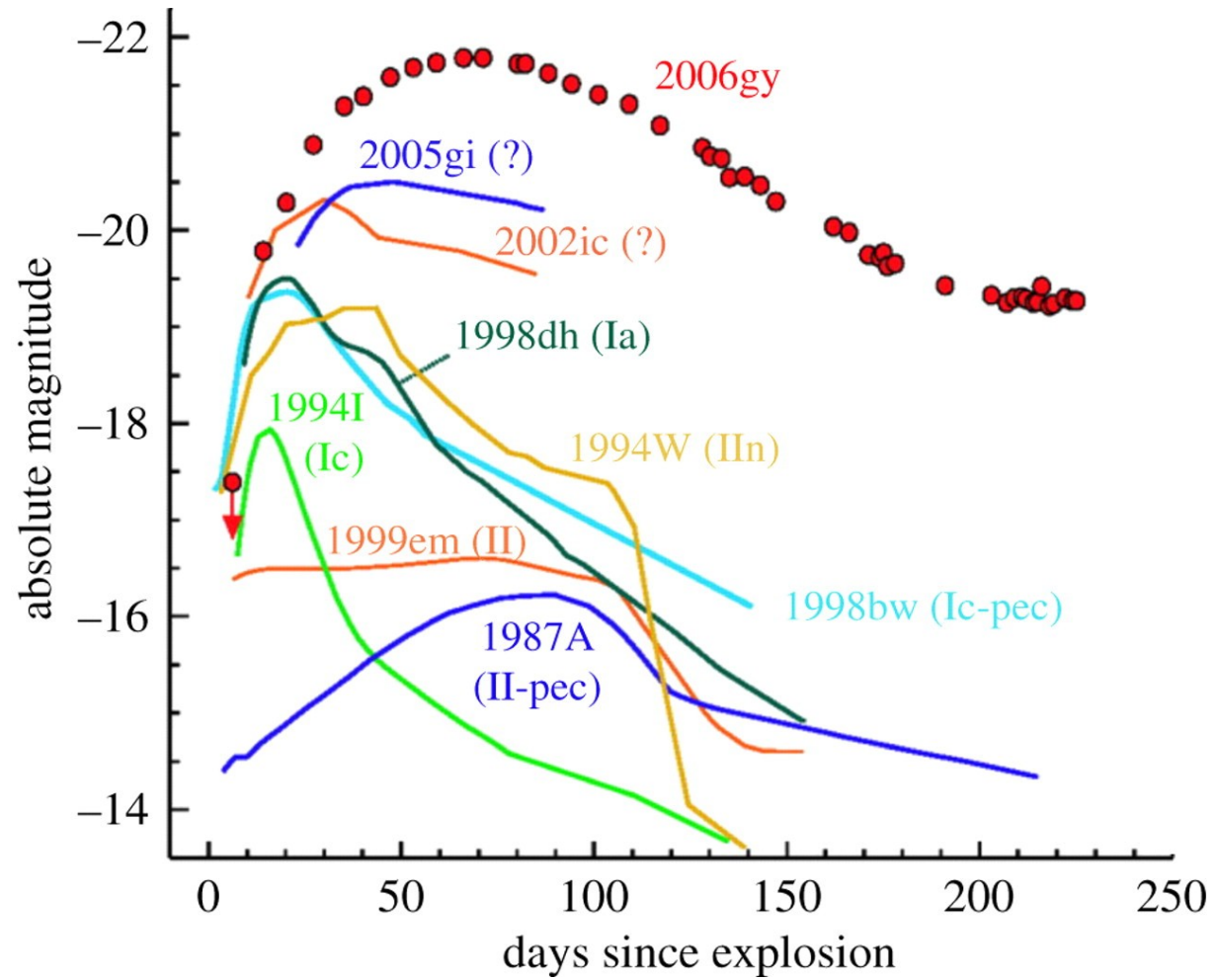
Spectra (high resolution snap-shot)



<http://supernova.lbl.gov/~dnkasen/tutorial/>

Photometry

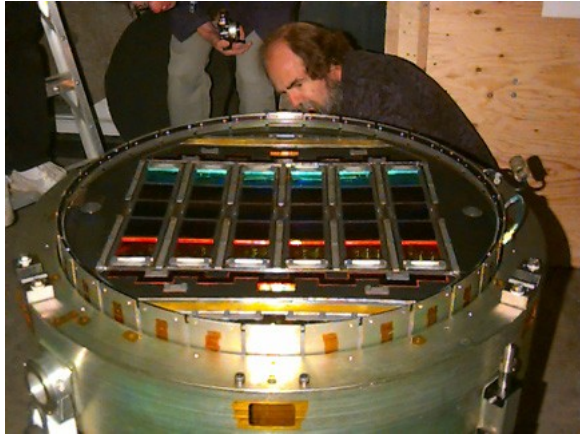
Light Curves (low resolution evolution)



<http://rsta.royalsocietypublishing.org/content/370/1960/774>

Photometry x Spectroscopy

An example from SDSS

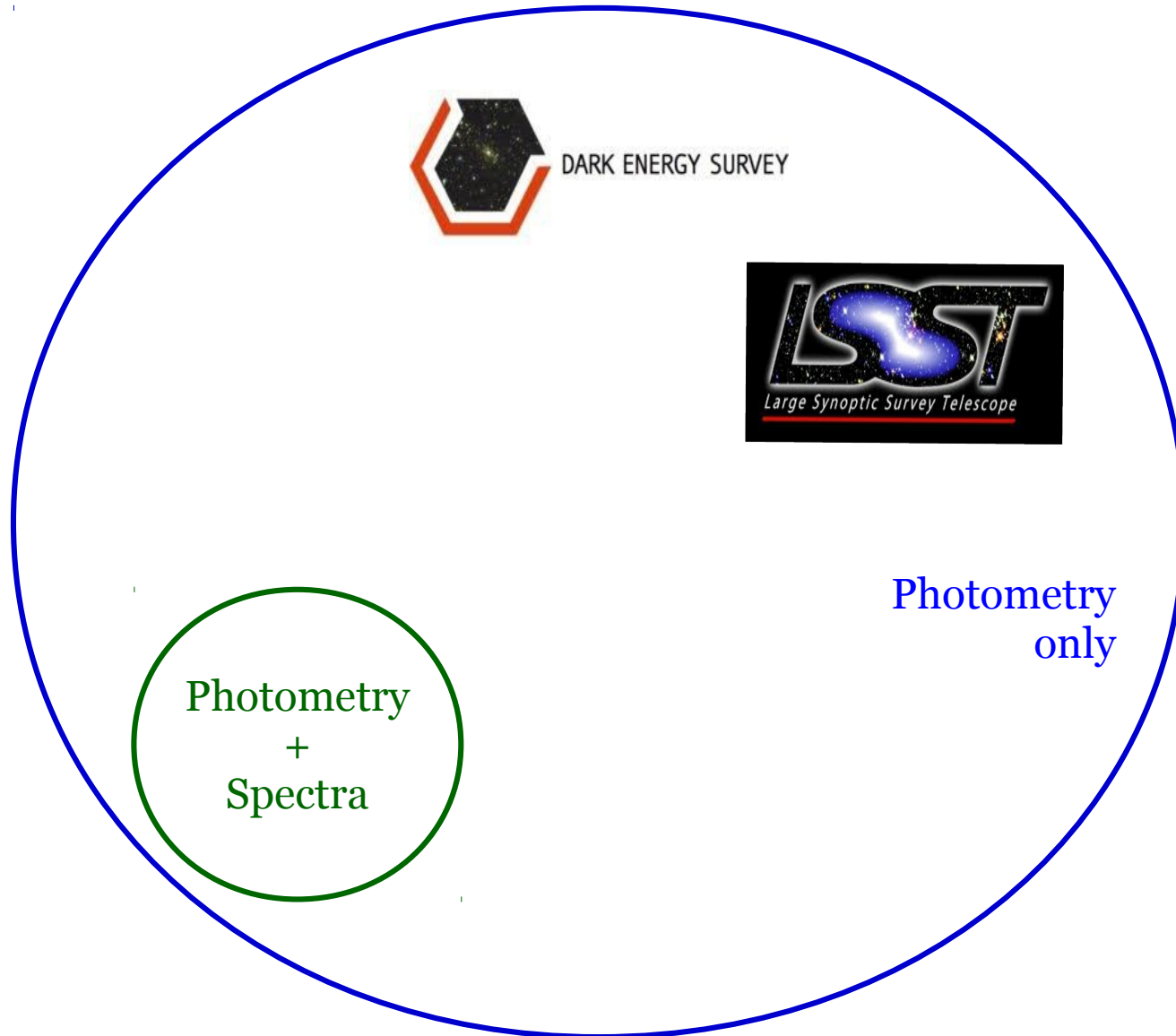


Exposure time 2 x 54s

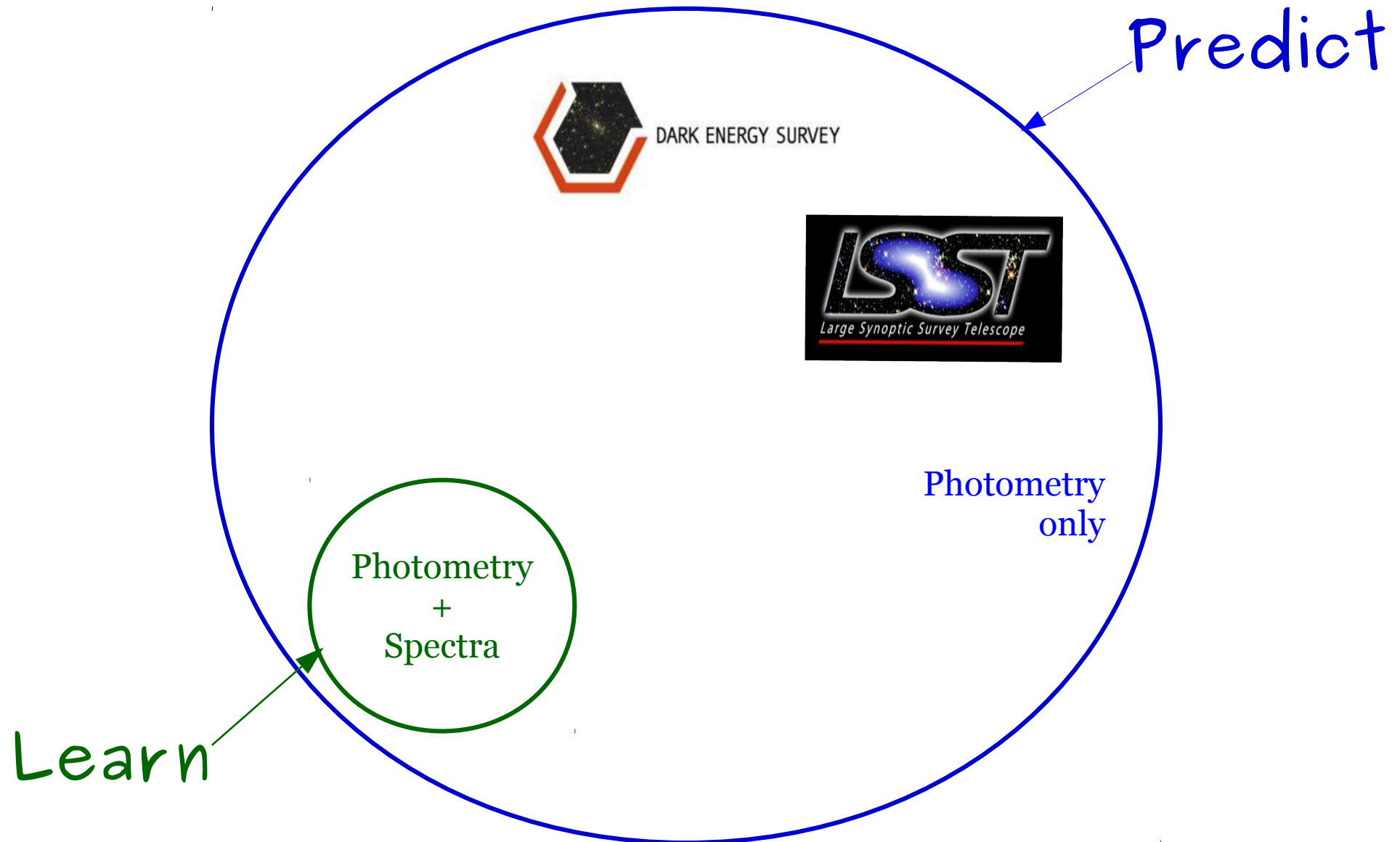


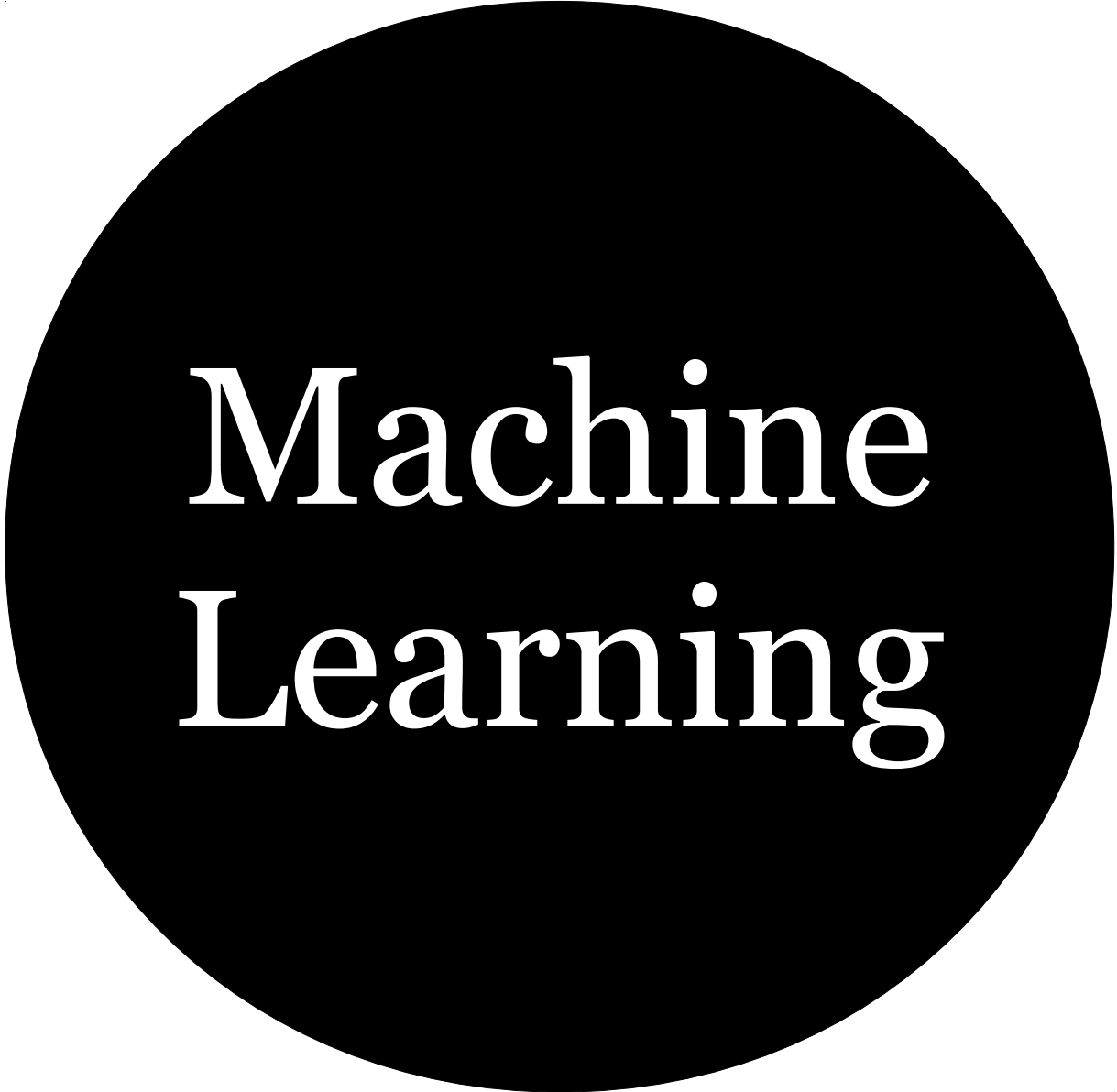
~ Integration at least 45 minutes

Data Situation



Data Situation





Machine Learning

Machine Learning: a definition



Machine learning is a field of computer science that gives computer systems the ability to "learn" (i.e., progressively improve performance on a specific task) with data, without being explicitly programmed.

Machine Learning: a definition

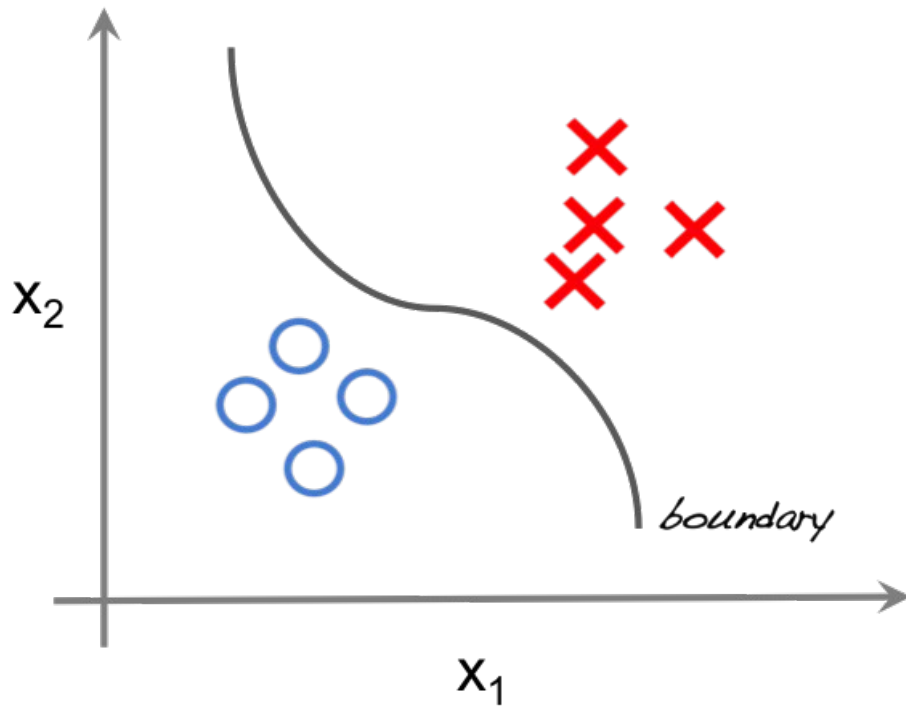


Machine learning is a field of computer science that gives computer systems the ability to "learn" (i.e., progressively improve performance on a specific task) with data, without being explicitly programmed.

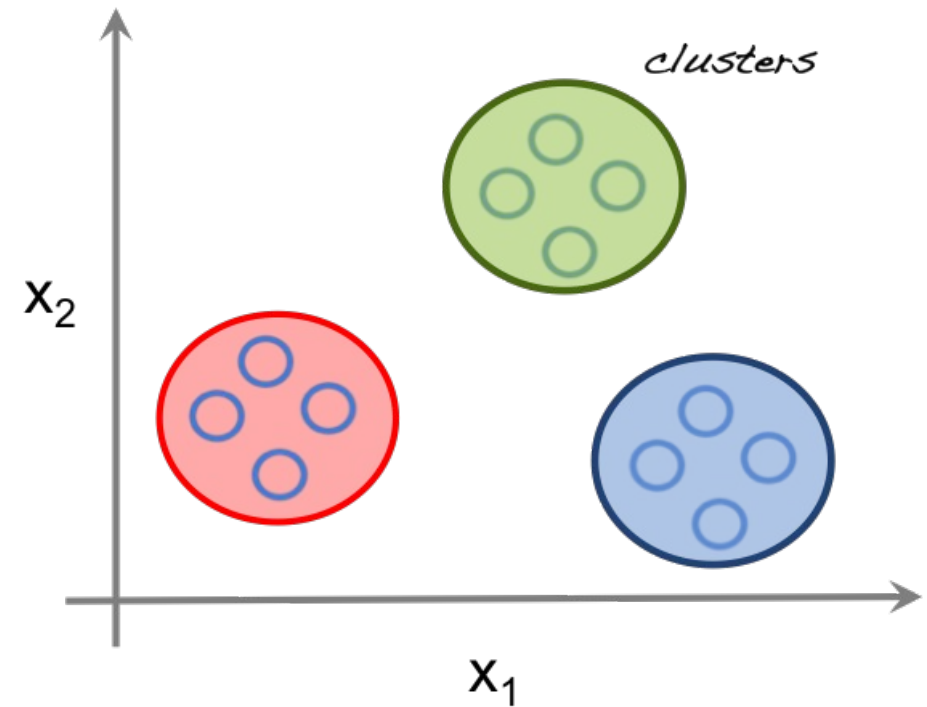
Machine Learning

Flavors

Supervised Learning
learn by example



Unsupervised Learning
identify similarities



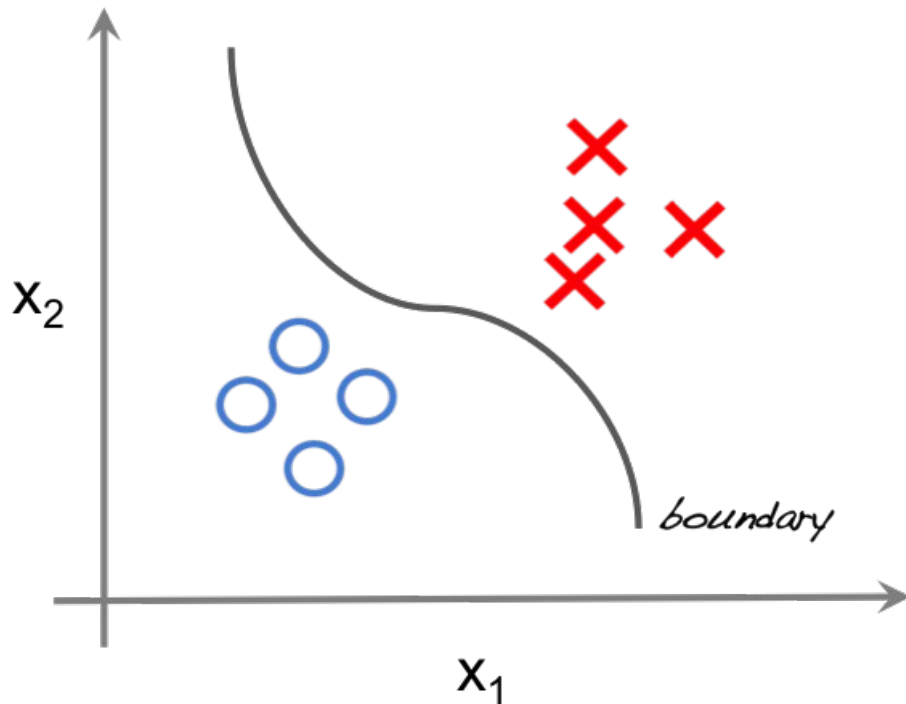
x_1, x_2 —> features

○ × —> classes

Machine Learning

Flavors

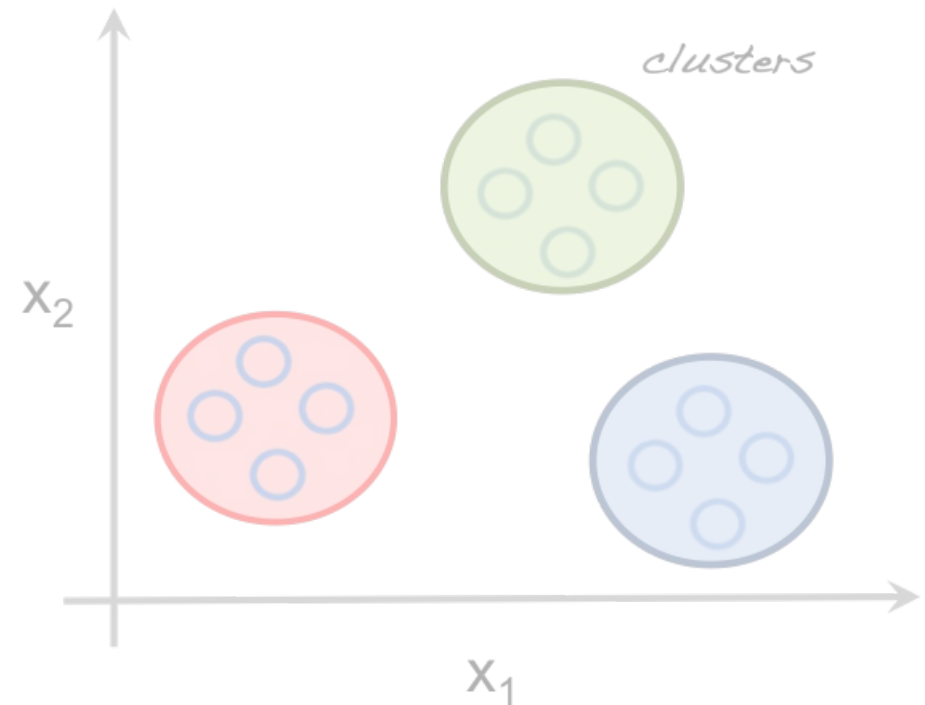
Supervised Learning
learn by example



x_1, x_2 —> features

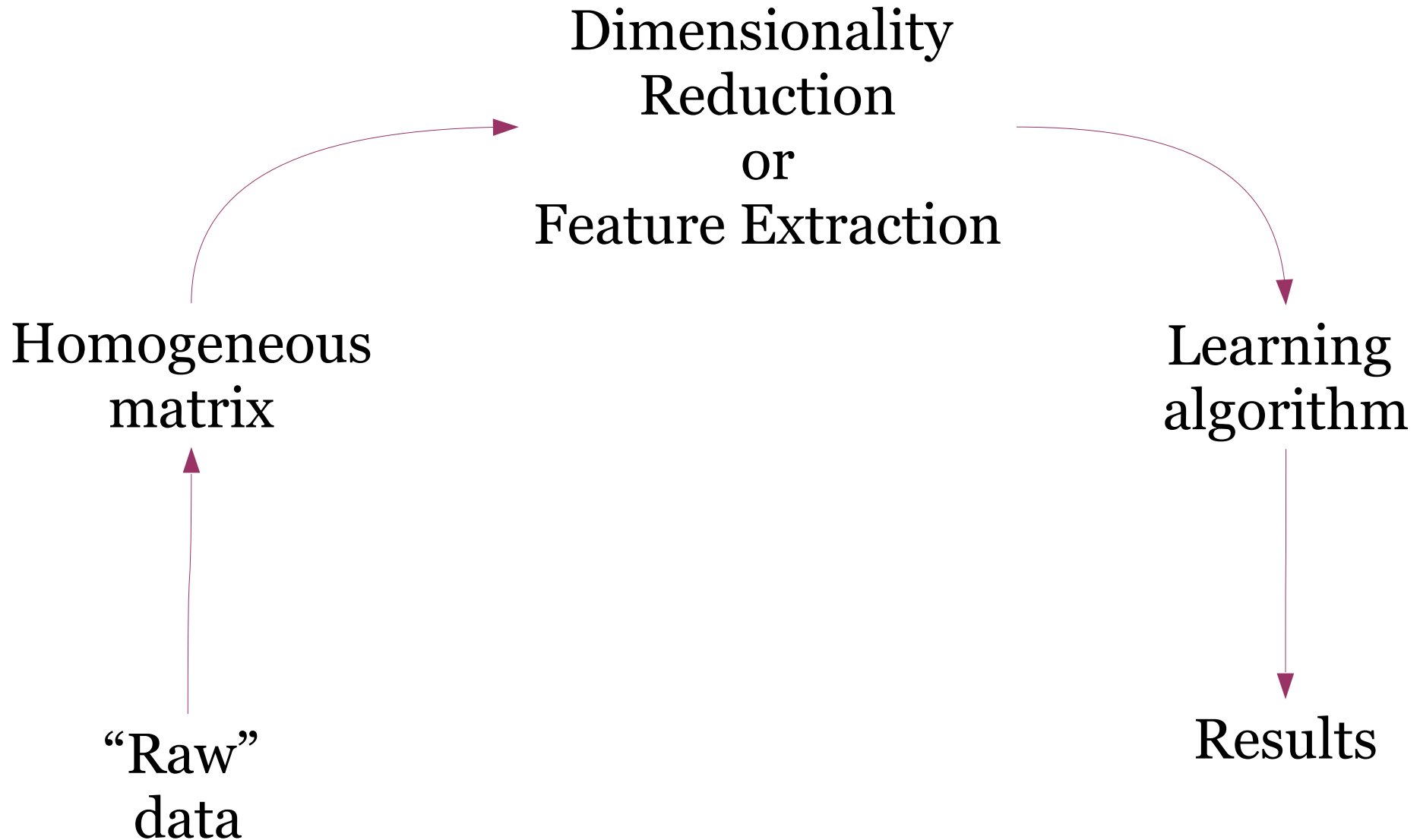
○ × —> classes

Unsupervised Learning
identify similarities

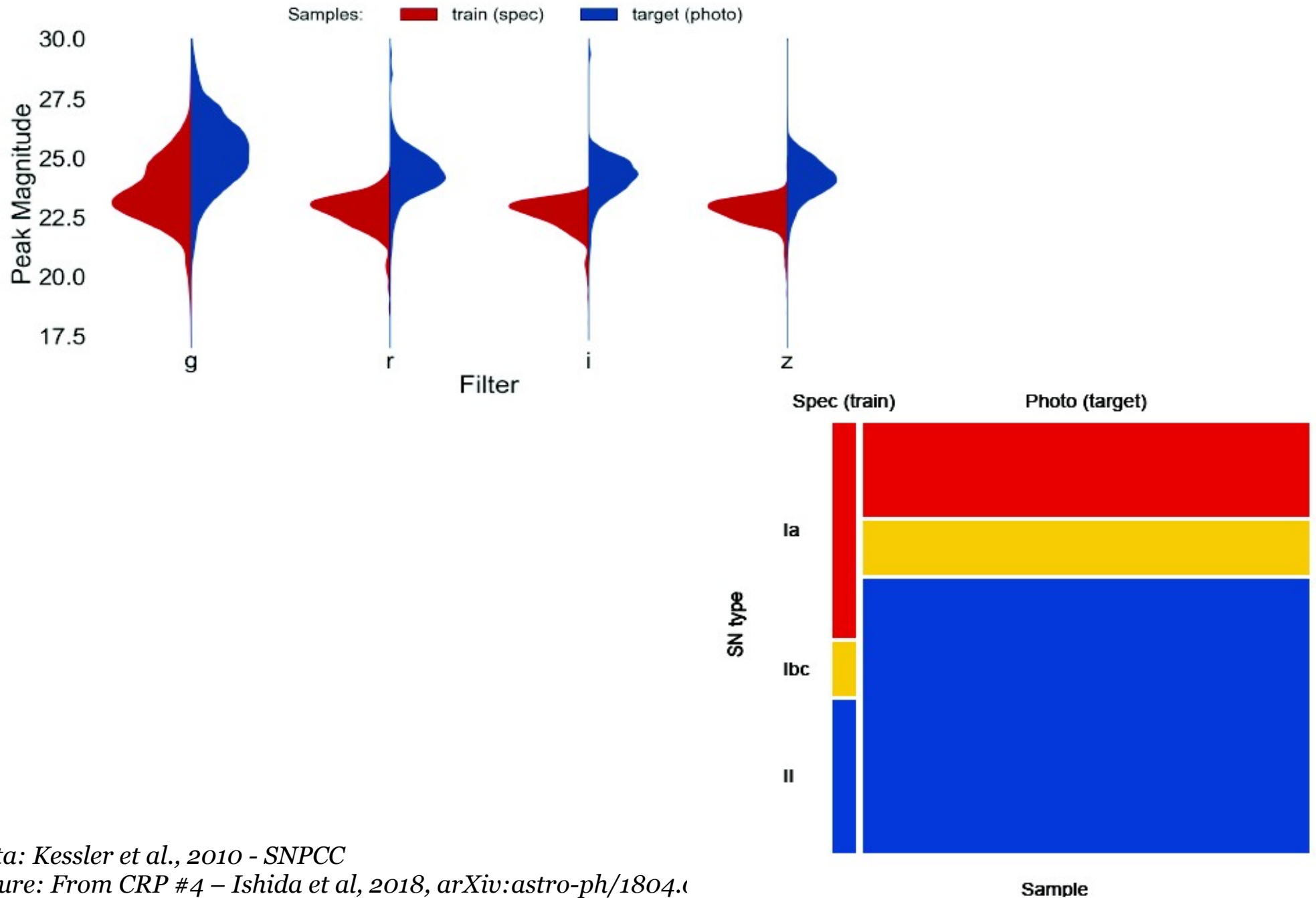


Machine Learning

Work-flow



Representativeness

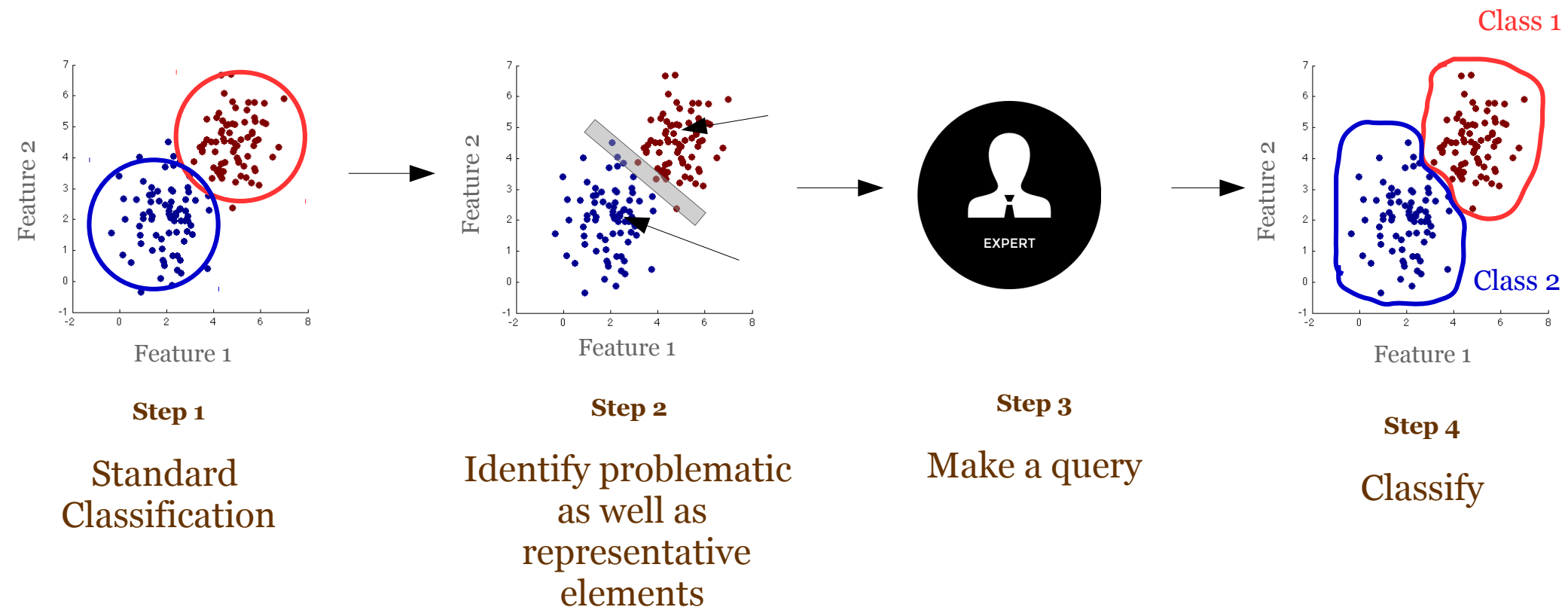


Data: Kessler et al., 2010 - SNPCC

Figure: From CRP #4 - Ishida et al, 2018, arXiv:astro-ph/1804.0

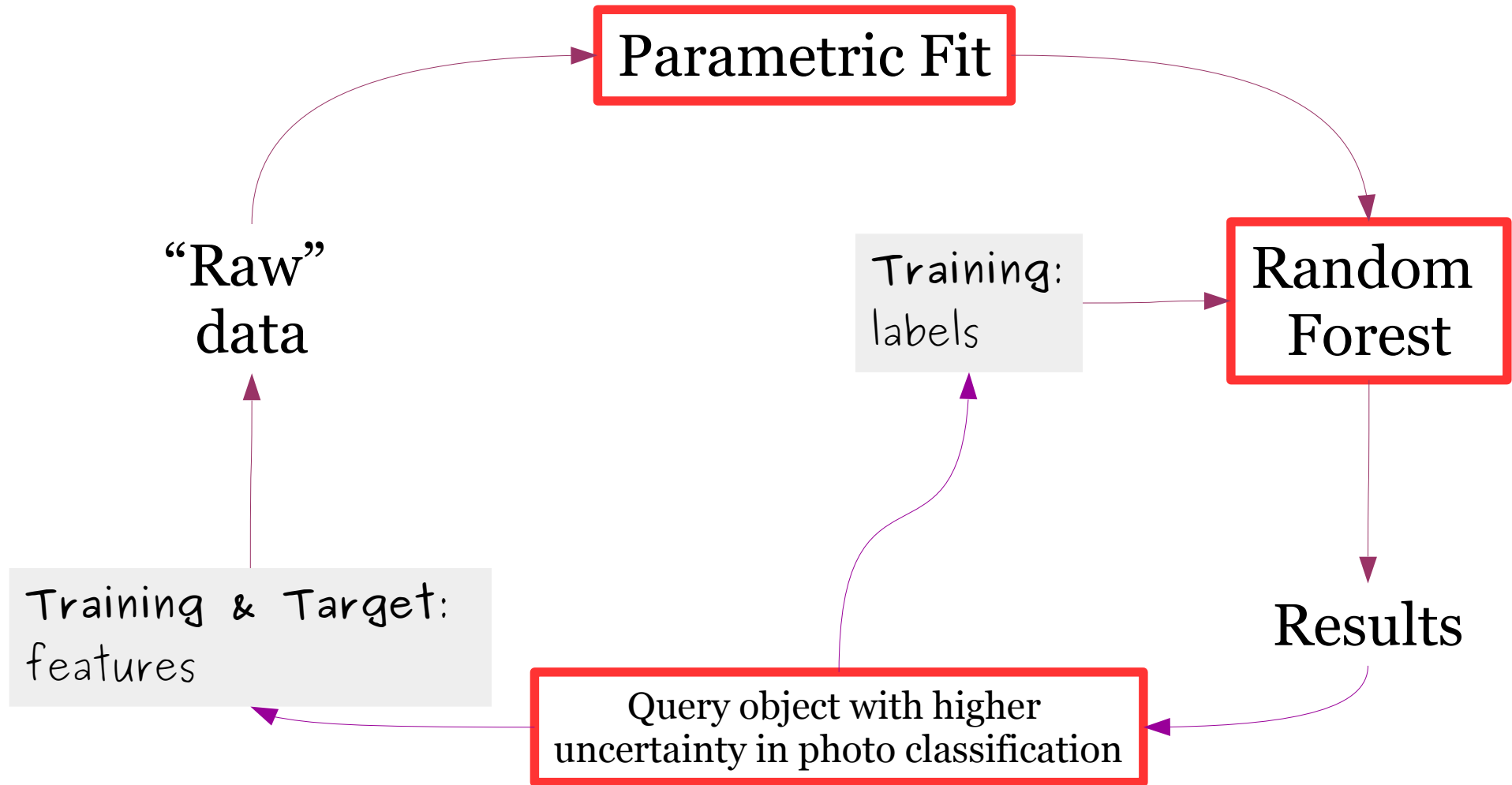
Active Learning *or Optimal Experimental Design*

“Can machines learn with fewer labeled training instances if they are allowed to ask questions?”



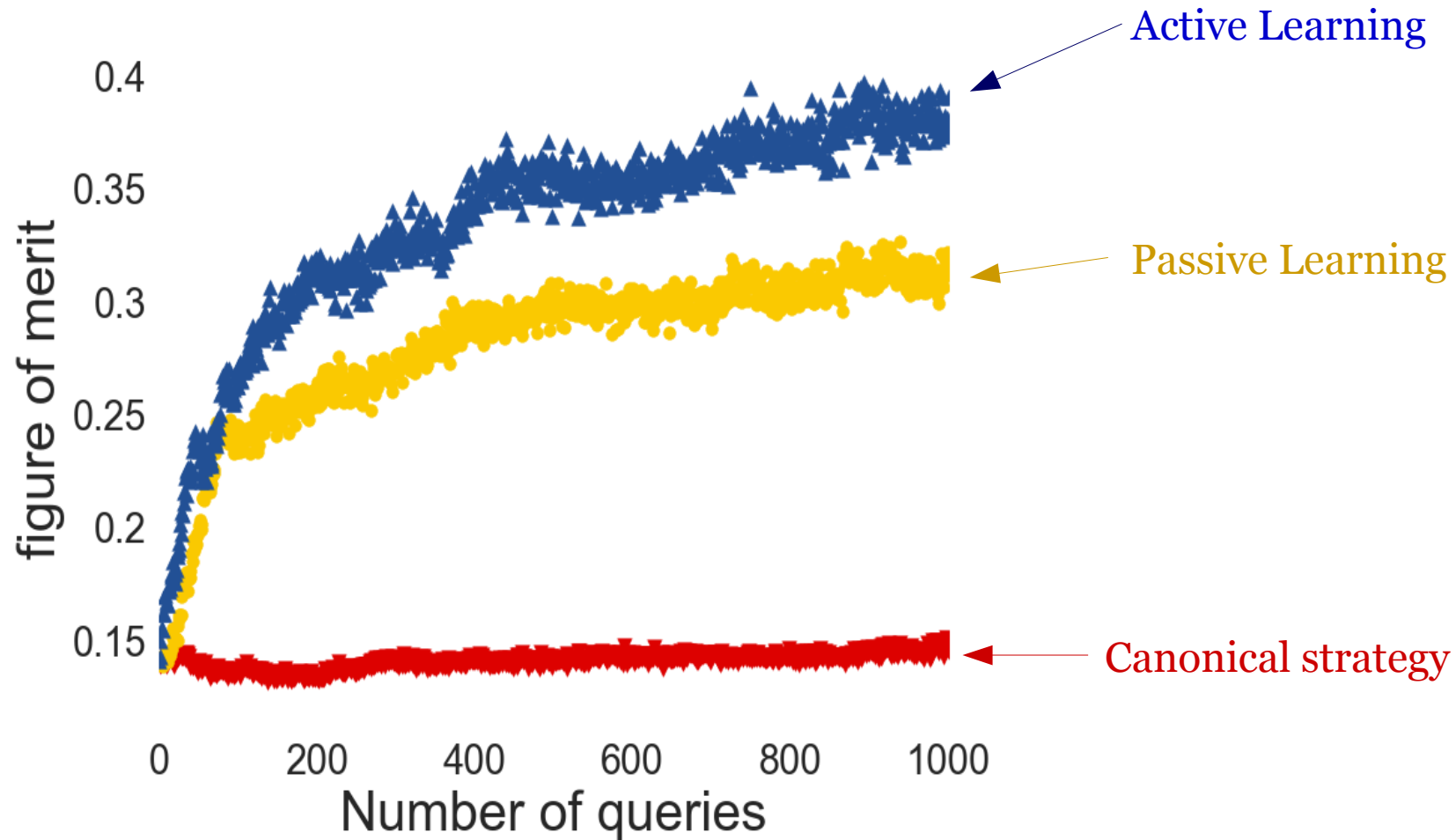
ML for Supernova classification

Our strategy



AL for Supernova Classification

Diagnostic plots

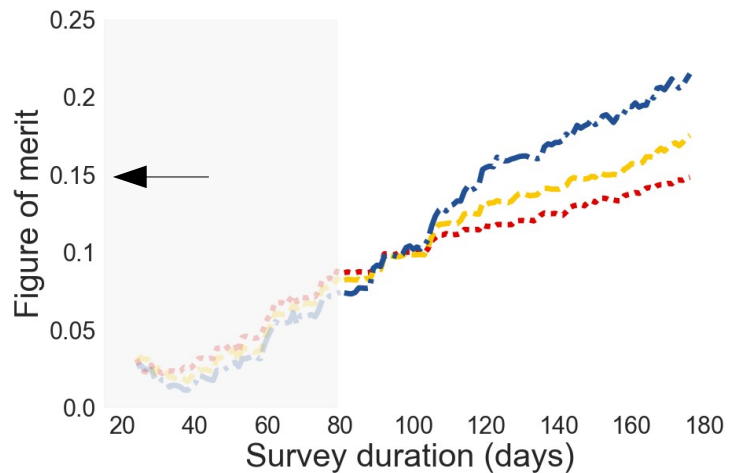
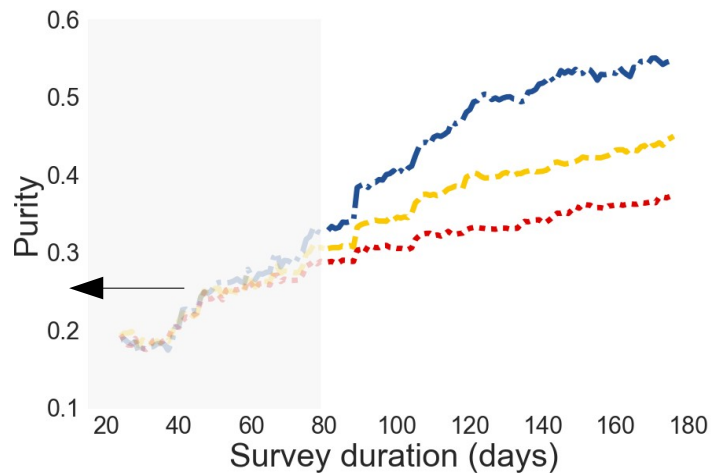
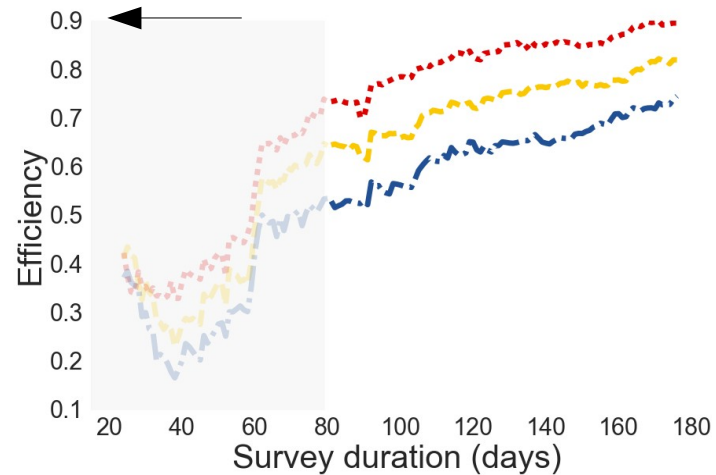
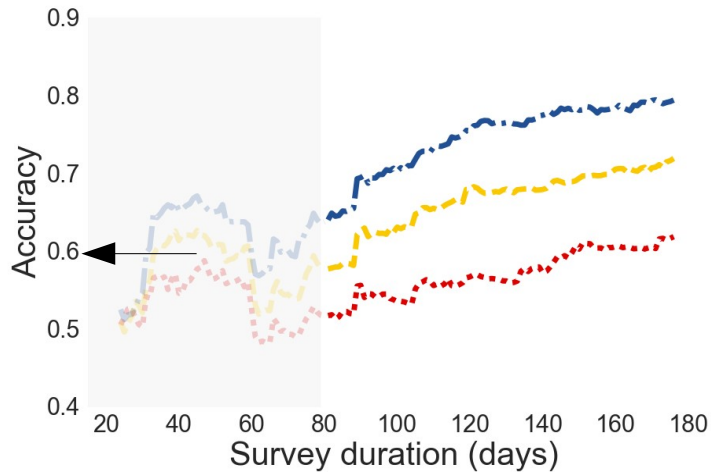


From COIN Residence Program #4, **Ishida et al., 2018**, [arXiv:astro-ph/1804.03765](https://arxiv.org/abs/1804.03765)

AL for Supernova Classification

Real time, partial light curve, restricted query sample

Strategy: - - - Canonical - - - Passive Learning - - - AL: Uncertainty sampling



The arrow
Shows *traditional*
Full light-curve
results with full
SNPCC spec

From COIN Residence Program #4, **Ishida et al., 2018, arXiv:astro-ph/1804.03765**

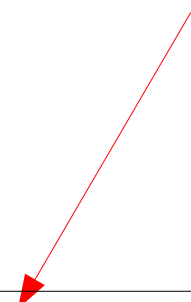
Summary

Active Learning *designed for astronomical data*



What we need

What we have



How do we optimize machine learning results with a minimum number of labeled training instances?"

