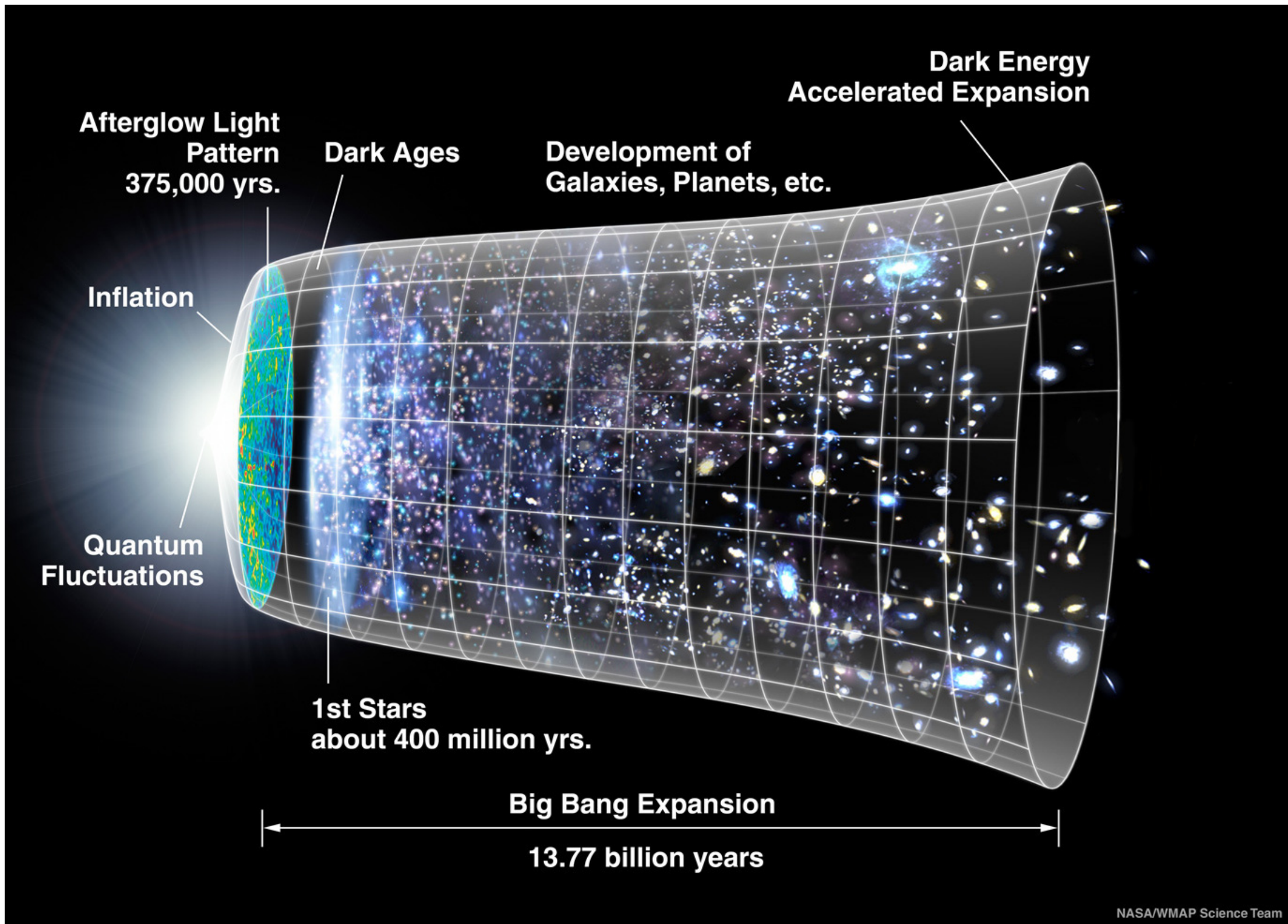


Machine learning classification and
statistical analyses:
challenges in supernova cosmology

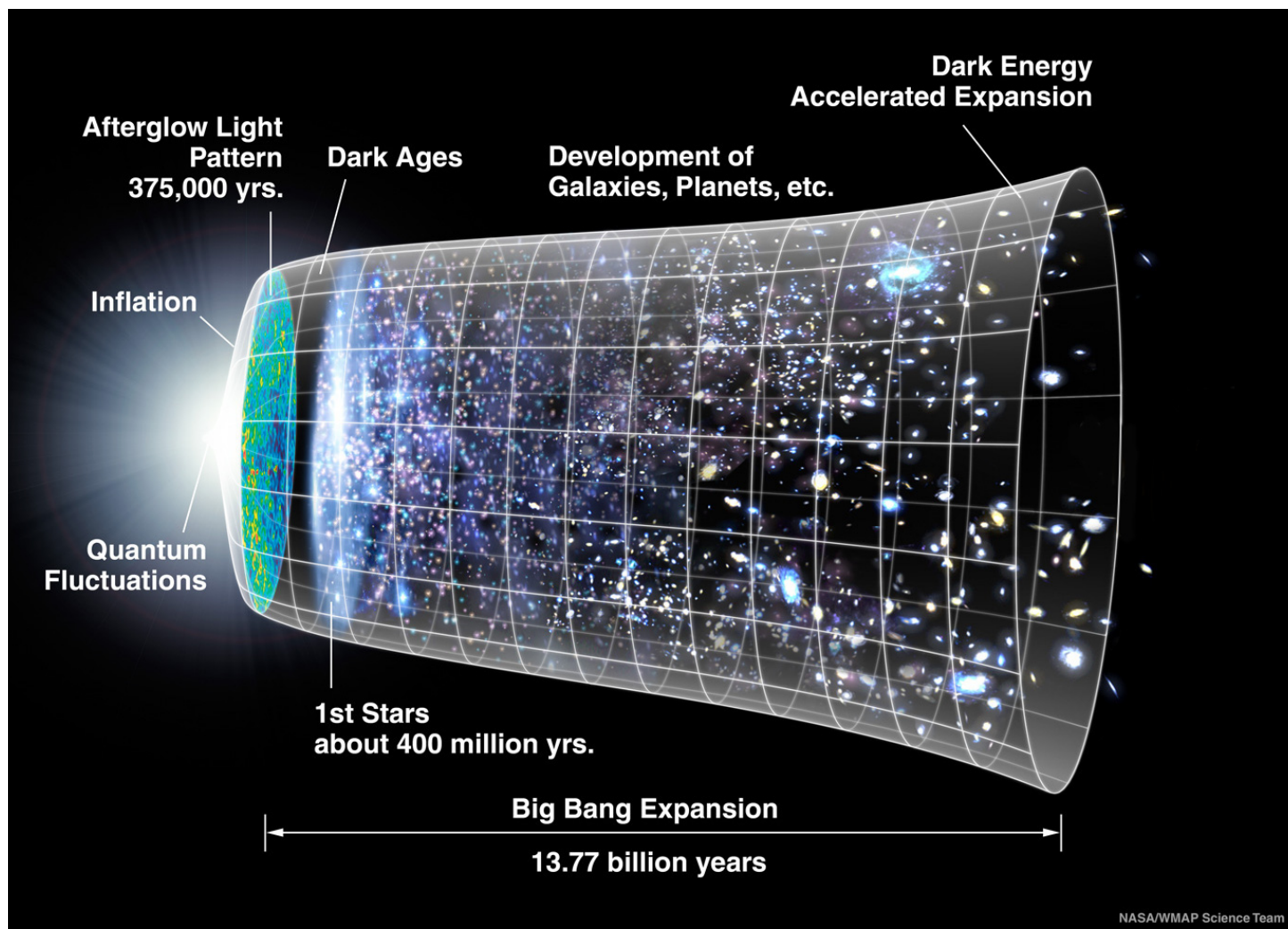
Anais Möller
CNRS / LPC Clermont

Advanced Pattern Recognition workshop
October 23rd, 2019

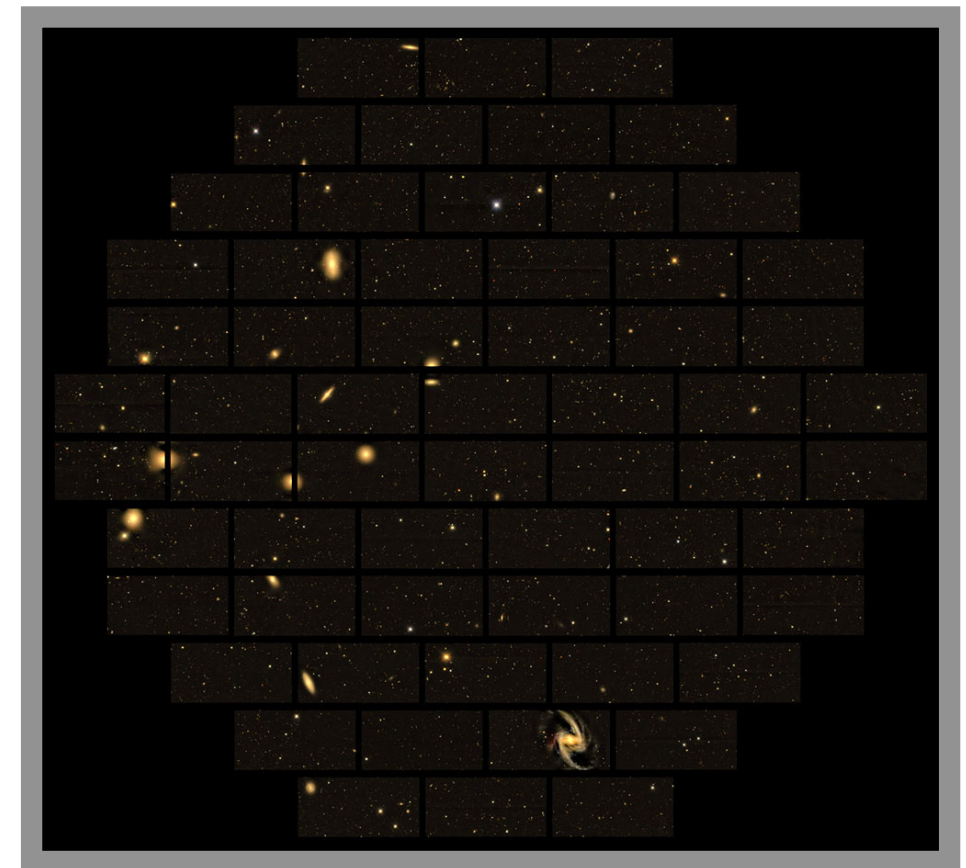
LambdaCDM universe



LambdaCDM universe

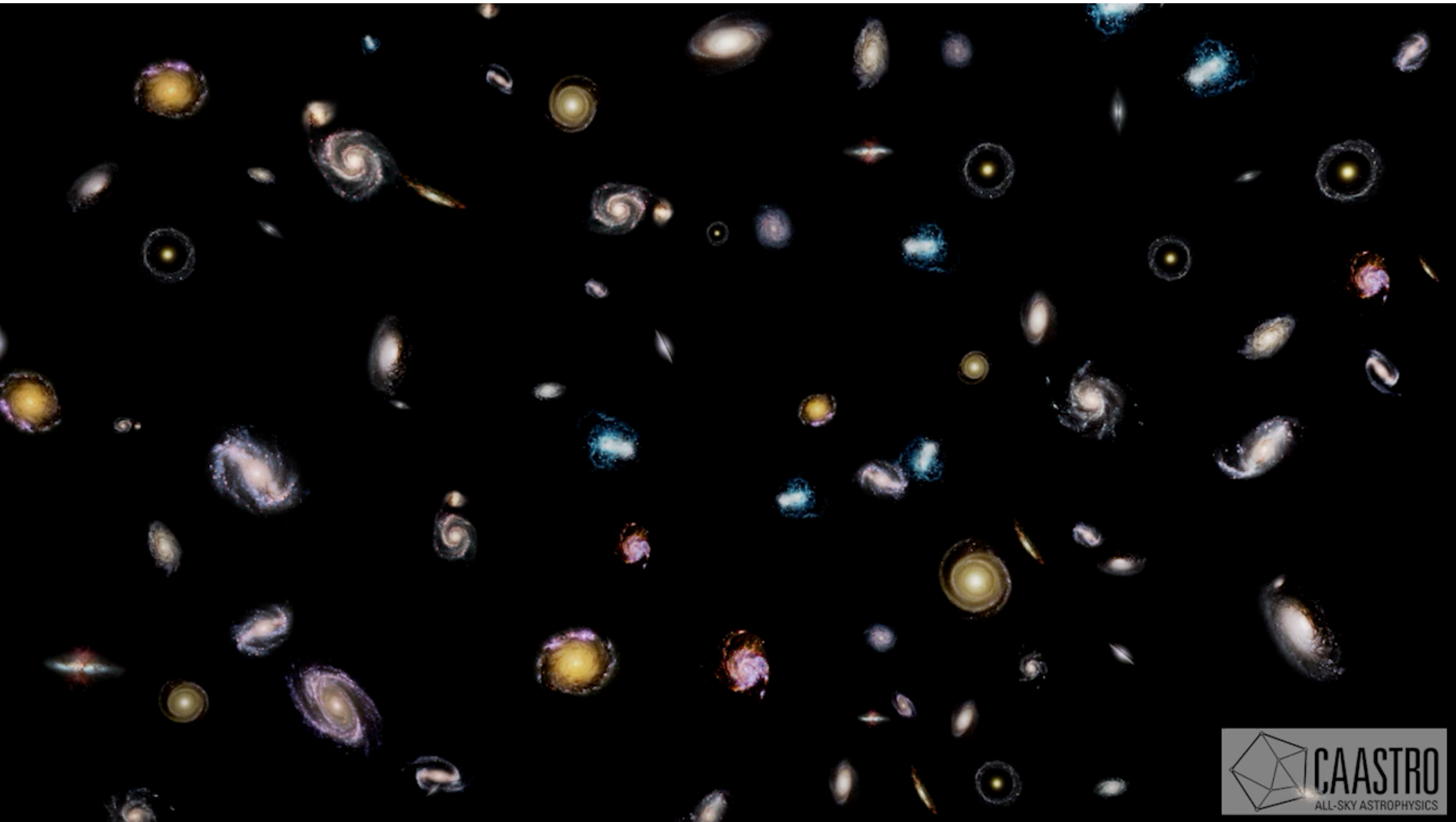


DECam @ Blanco telescope in Chile



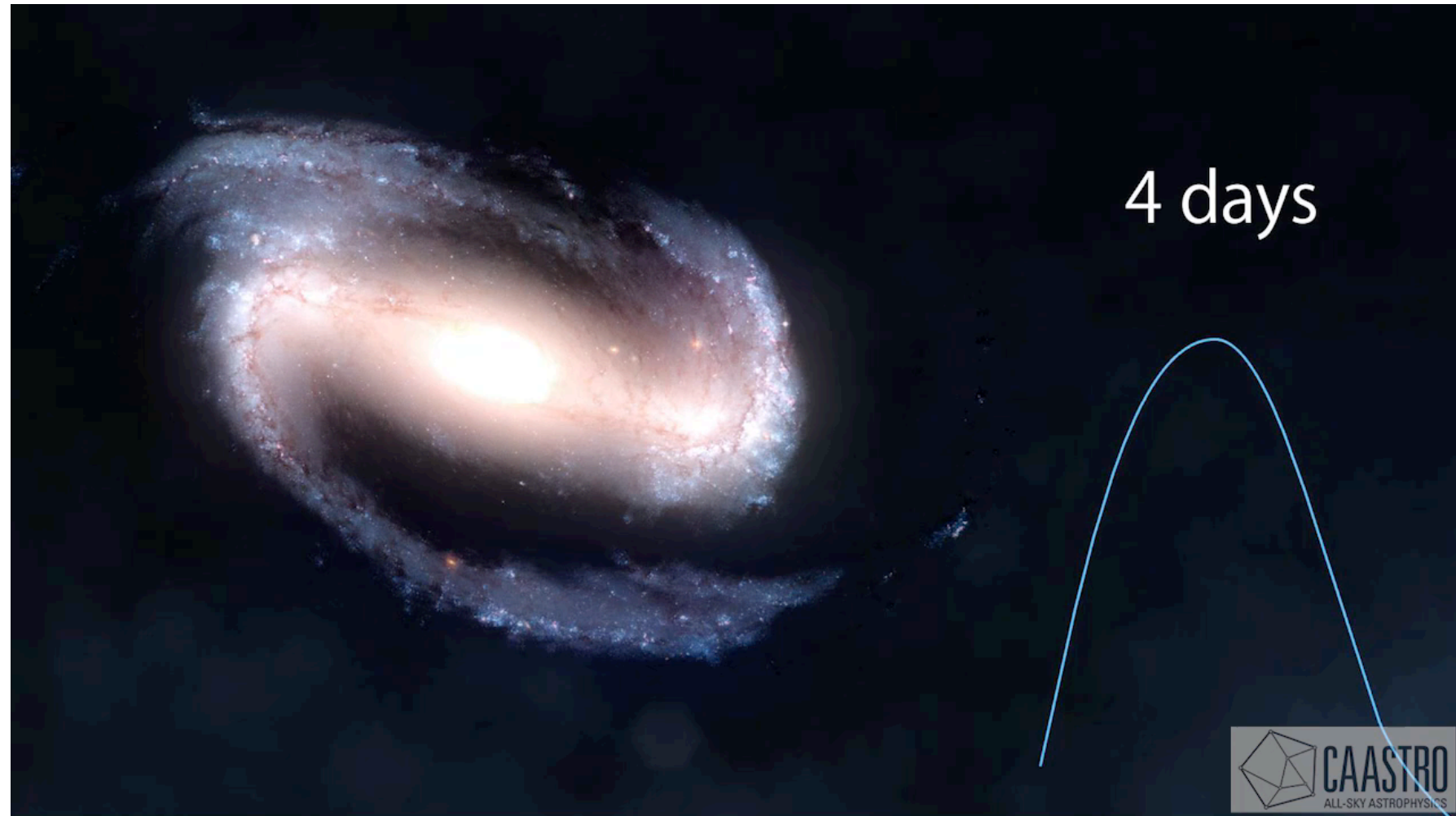
galaxy clusters, weak lensing, large scale structure, type Ia SNe, gravitational waves (kilonovae), ...

Cosmic expansion



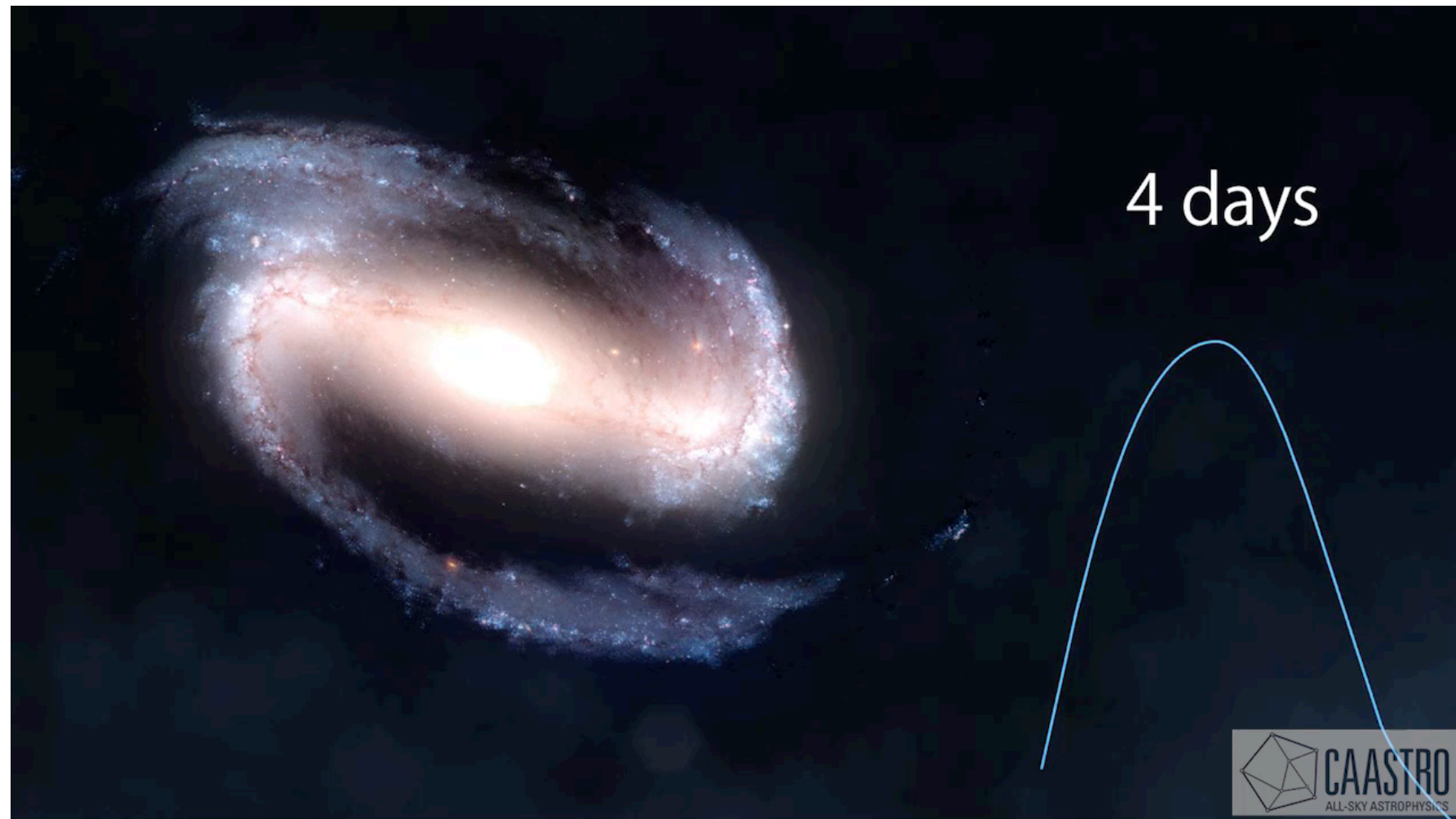
supernovae

- stellar explosions
(transient events)



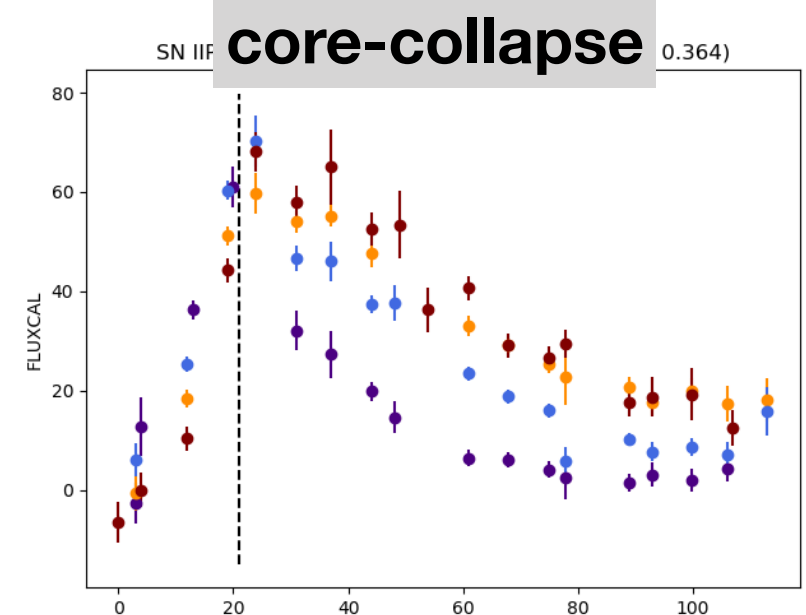
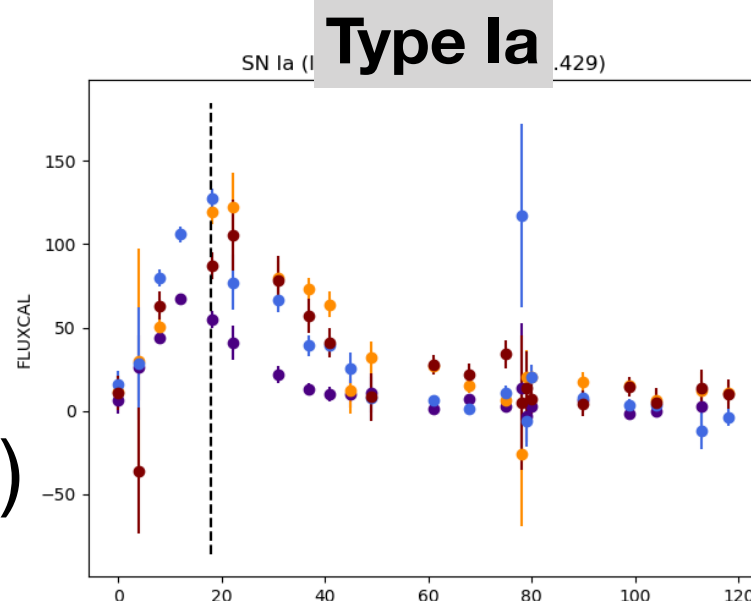
supernovae

- stellar explosions (transient events)



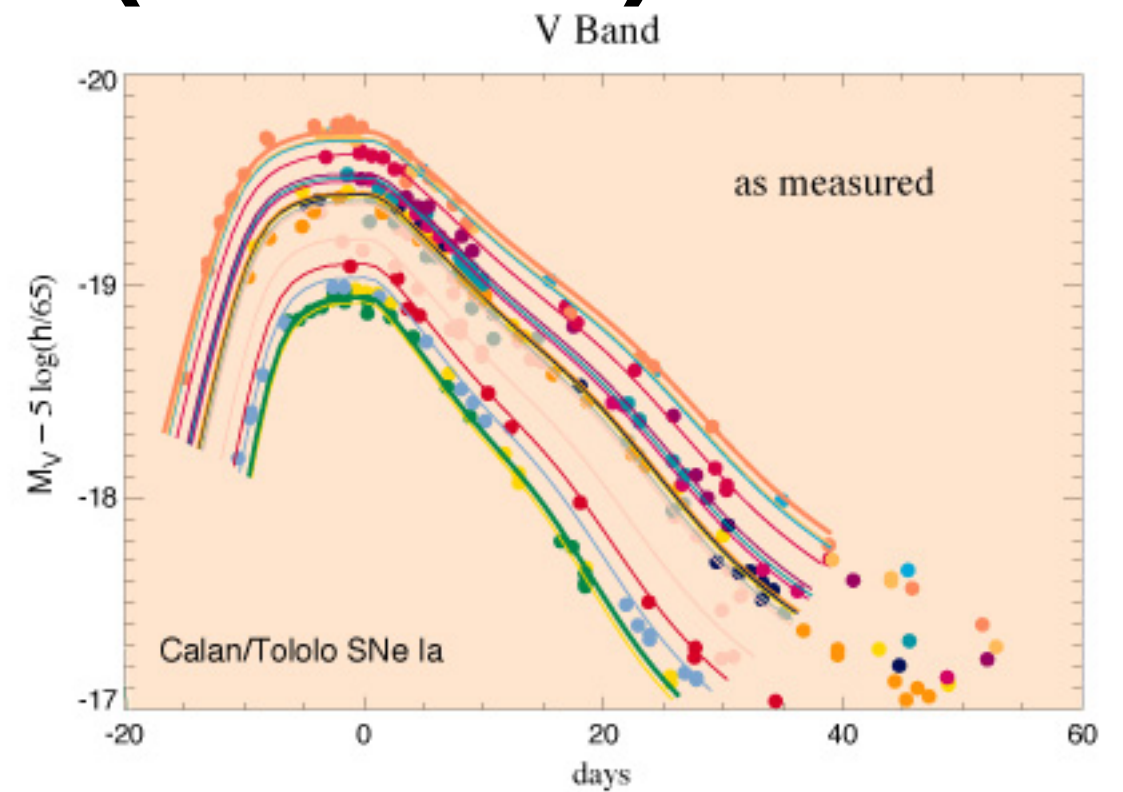
- types:

- Ia (thermonuclear)
- II, Ib, Ic (core-collapse)



type Ia supernovae (SNe Ia)

- very luminous
- homogeneous spectral and photometric properties



type Ia supernovae (SNe Ia)

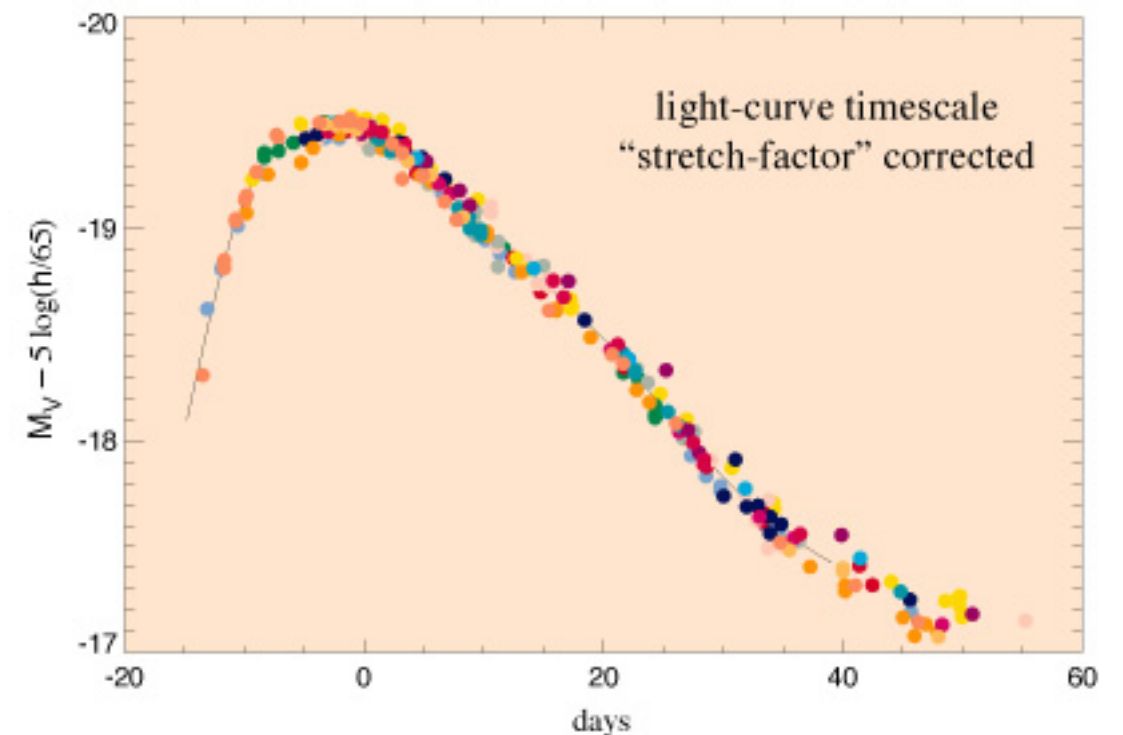
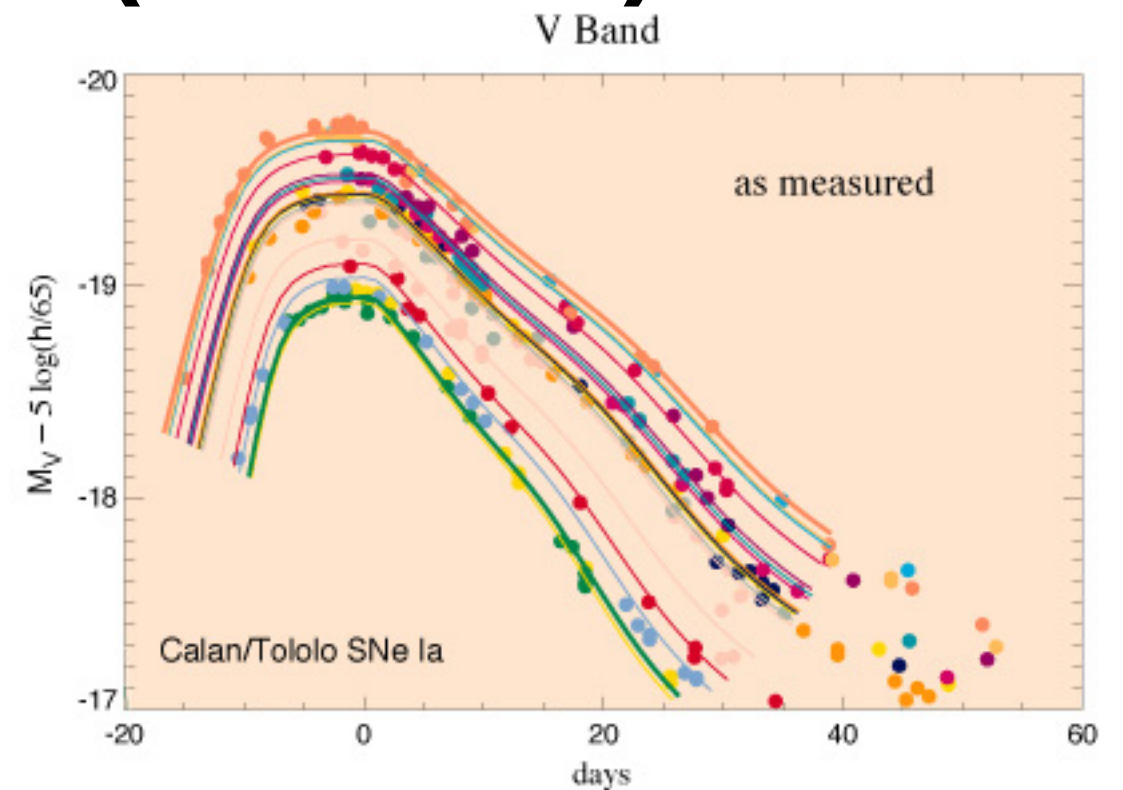
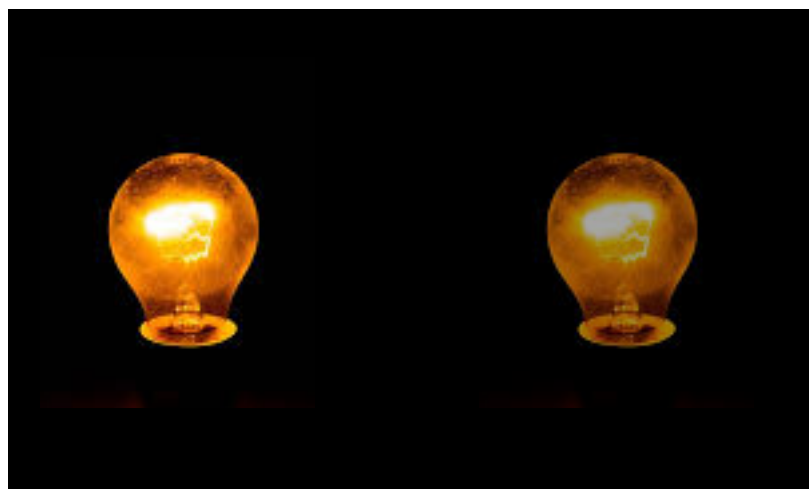
distance modulus:

$$\mu_B = m_B^* - M_B + \underbrace{\alpha x_1 - \beta C}_{\text{correction shape \& color}}$$

observed
brightness

absolute
brightness

correction
shape &
color



current surveys



In numbers:

- * 5-year survey, started 2013

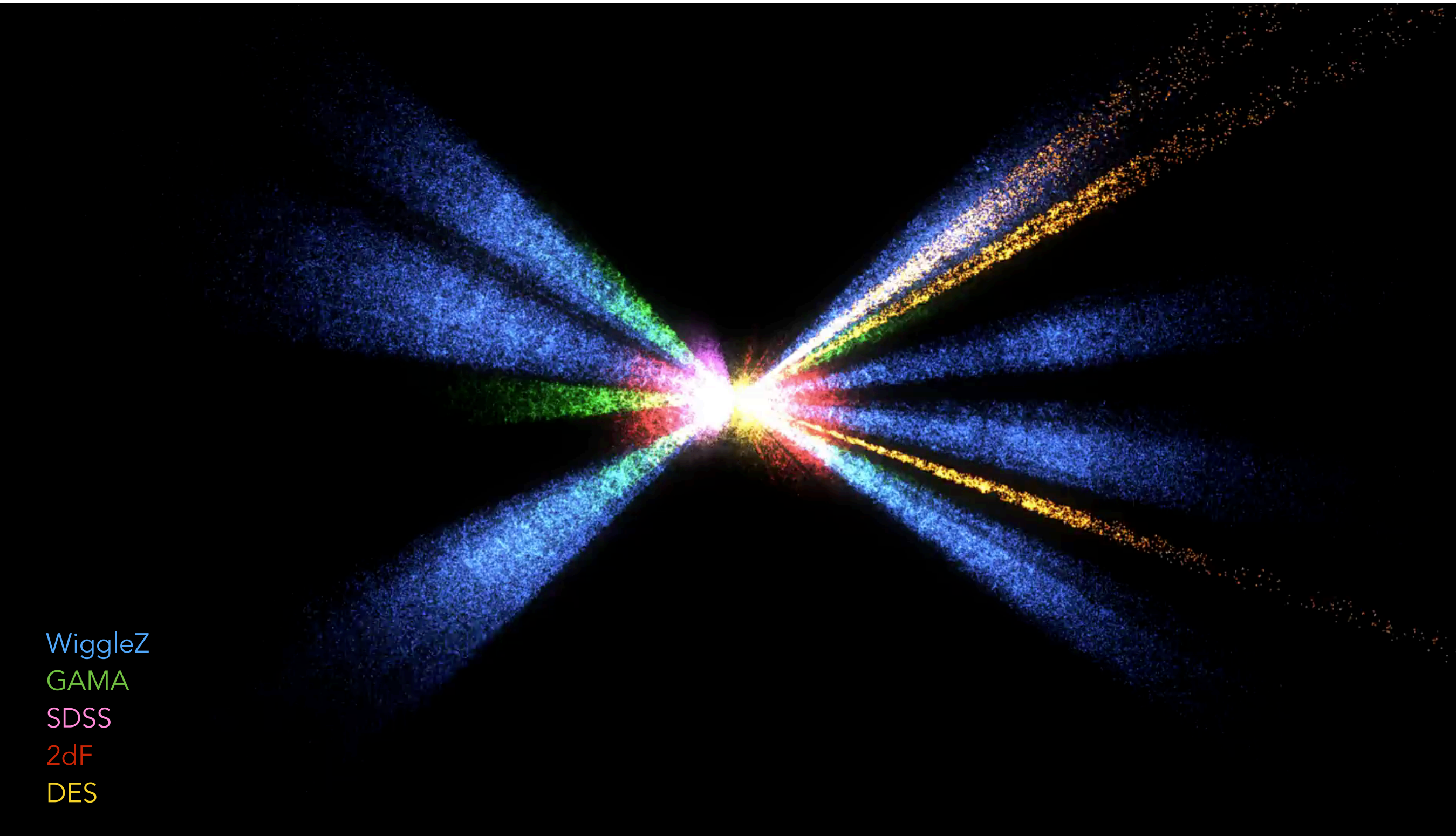
- * 4 primary probes: galaxy clusters, weak lensing, large scale structure, type Ia SNe

- * <2,000 well measured SNe Ia

current surveys



THE DARK ENERGY SURVEY



future surveys:

LSST



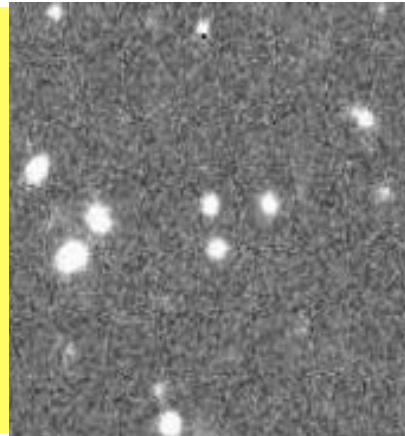
In numbers:

- * 10-year survey, starting 2022
- * 1,000 images/night = 15 TB/night
- * 10,000 alerts/30 seconds = 1 GB / 30 s
- * >4,000 well measured SNe Ia

outline

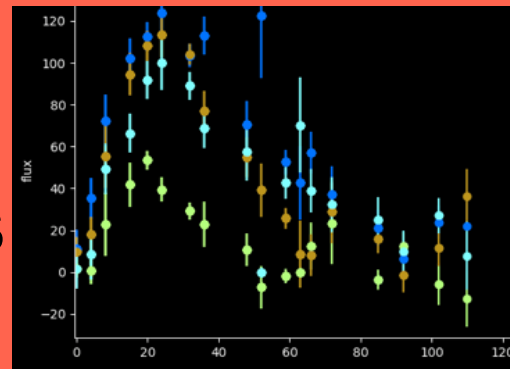
Machine learning in supernova cosmology: classification tasks

Part I: real vs. bogus



Part II: typing with photometry

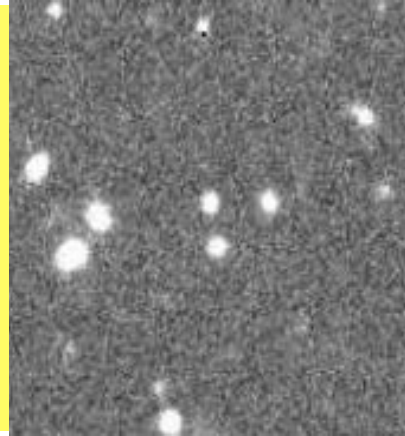
1. Datasets: PLAsTiCC
2. ML classification issues
 1. Representativity
 2. Incompleteness
 3. “probabilities” for cosmology
3. FINK broker

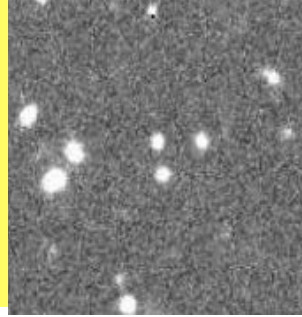


outline

Machine learning in supernova cosmology: classification tasks

Part I: real vs. bogus



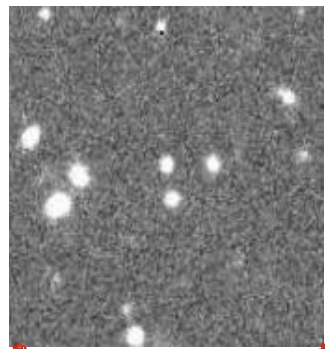


Finding transients

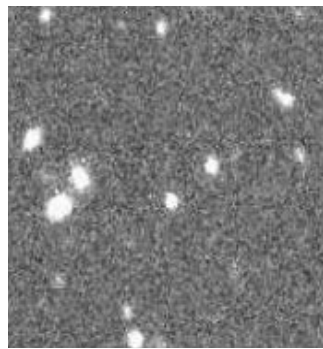
new

template

difference



-



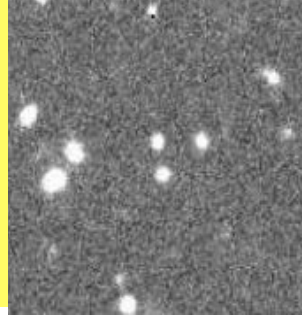
=



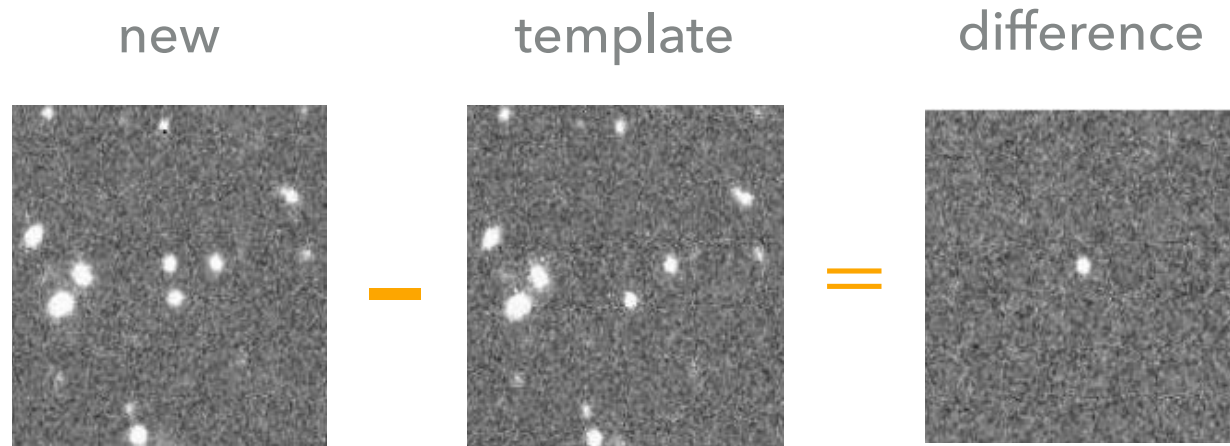
Difference imaging

*Kessler + 2015
Goldstein + 2015*





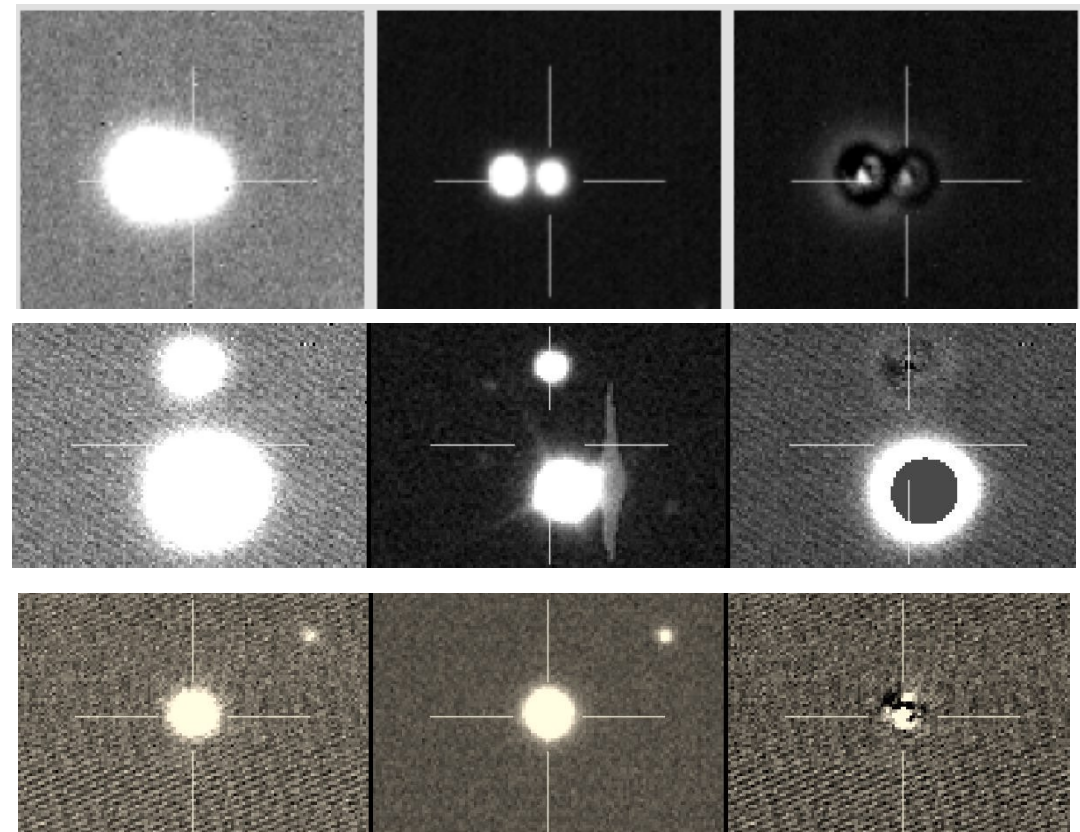
Finding transients

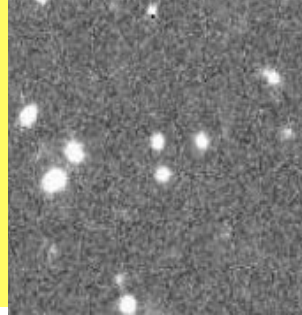


Difference imaging

*Kessler + 2015
Goldstein + 2015*

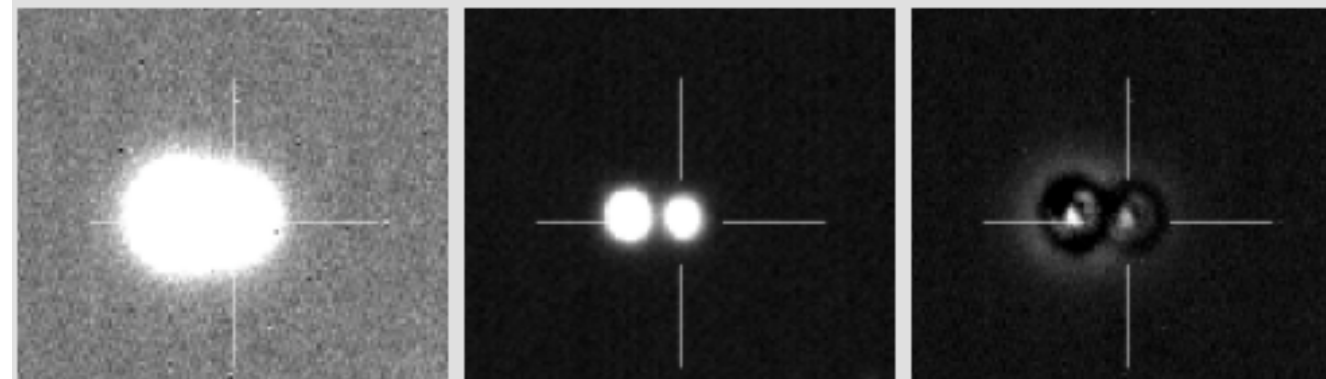
Bogus detections





1. Humans

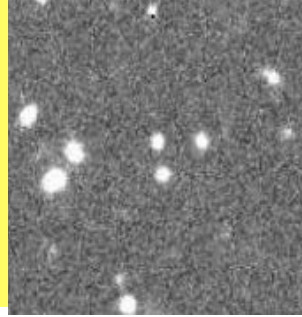
Your team



Citizen scientists

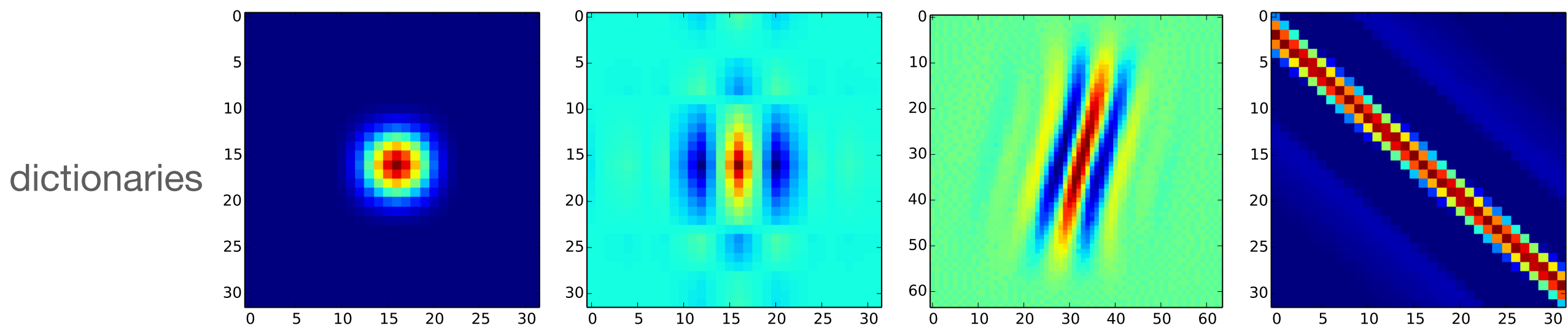


*AM, Tucker, Armstrong & the SkyMapper Transient team
2017-2018*



2. Signal Processing

Möller+2015 dictionaries where chosen based on SN-like and artifact studies:



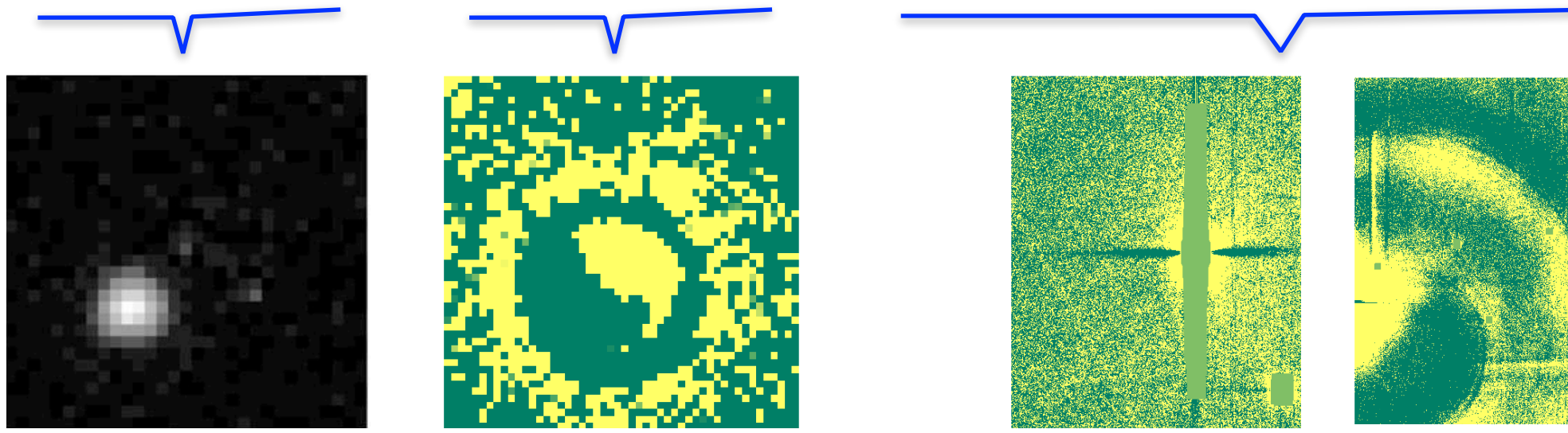
(A) Starlet

(B) Bi-orthogonal

(C) Curvelet

(D) Ridgelet

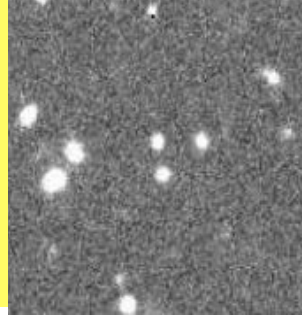
objects



SN-like

dipoles

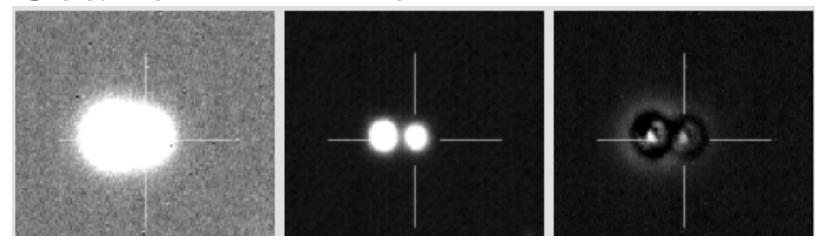
saturated star, optical ghost,...



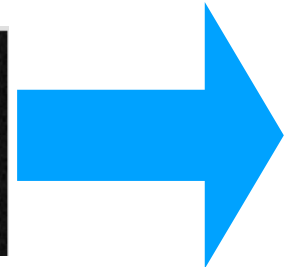
Eliminating bogus

3. Machine learning

Scalzo...AM +2017

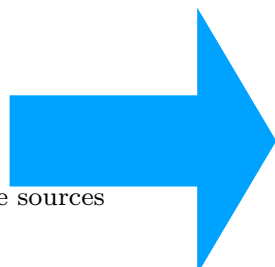


SkyMapper images

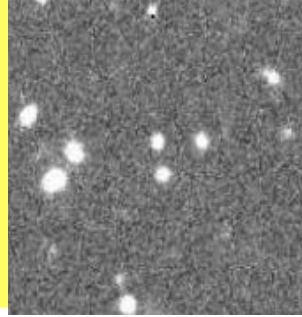


Features

- eref ellipticity of source on template image
- thref semi-major axis of source on template image
- fwhmref full width at half maximum of all template image sources
- f4ref flux within 4-pixel aperture in template image
- flagref SEXTRACTOR source flags in template image



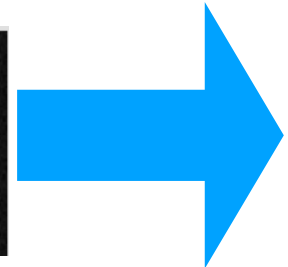
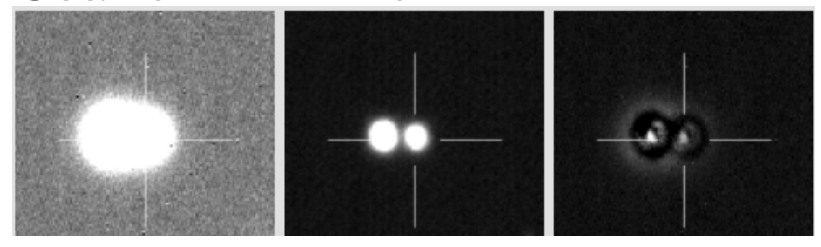
ML classifier
Random Forest



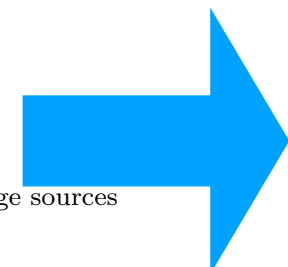
Eliminating bogus

3. Machine learning

Scalzo...AM +2017



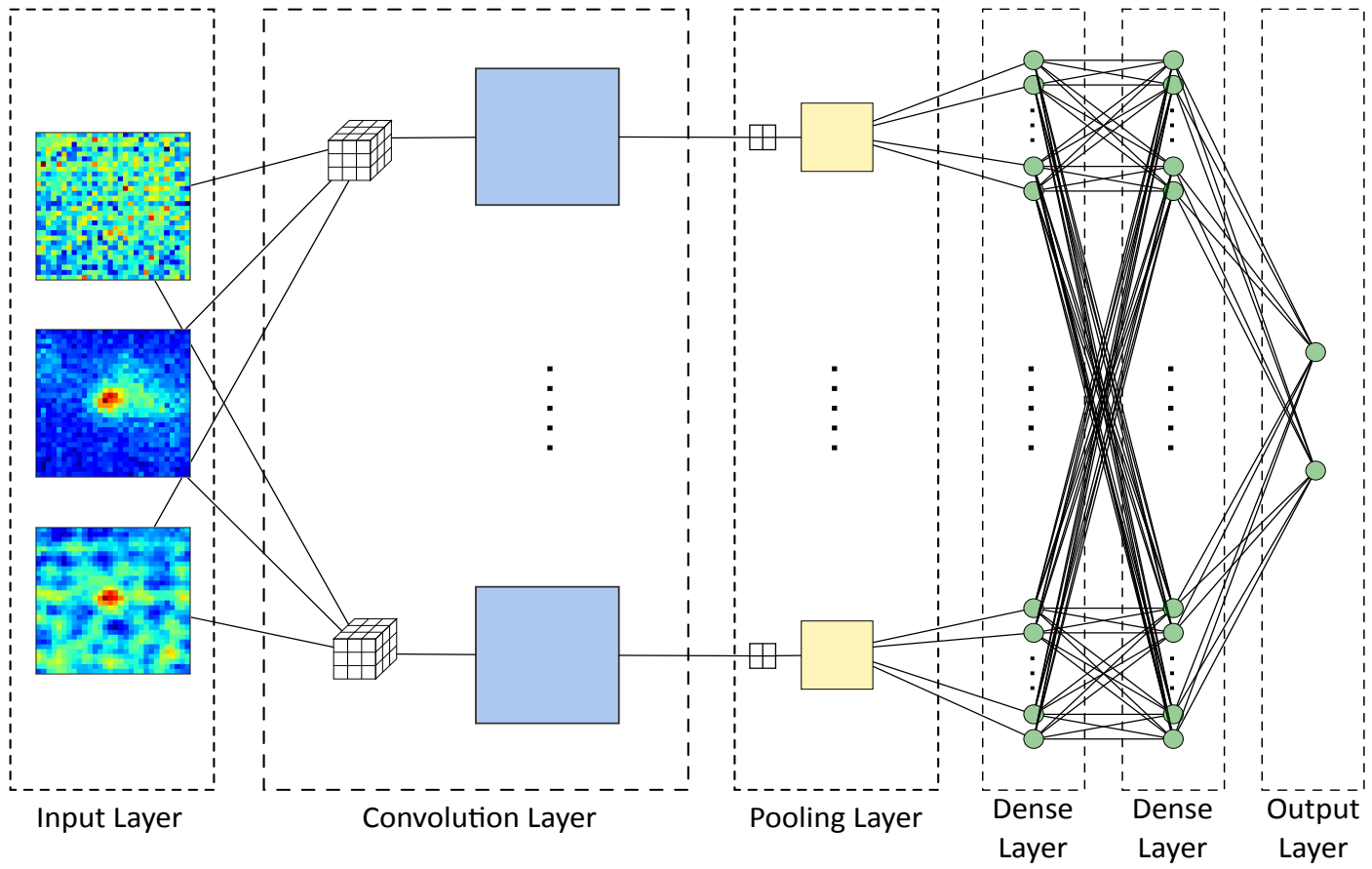
- Features**
- eref ellipticity of source on template image
 - thref semi-major axis of source on template image
 - fwhmref full width at half maximum of all template image sources
 - f4ref flux within 4-pixel aperture in template image
 - flagref SEXTRACTOR source flags in template image



ML classifier
Random Forest

SkyMapper images

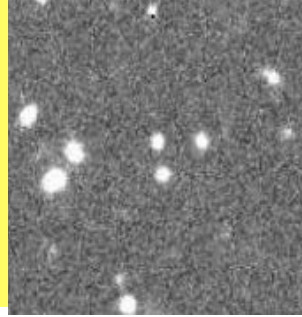
Gieseke...AM+ 2017



SkyMapper images augmented for training

Model	AUC
Random forest	0.9907
Net1(32,64)	0.9914
Net3	0.9972
E2	0.9946

Part I: real vs. bogus

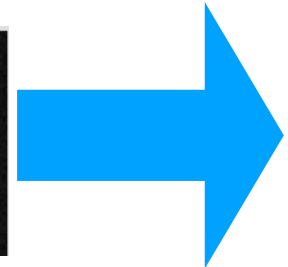
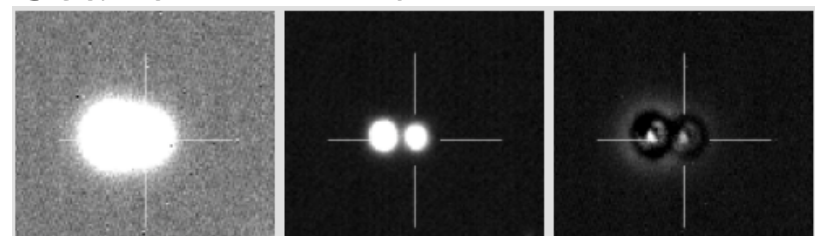


Eliminating bogus

limitations:
training sets!
feature extraction

3. Machine learning

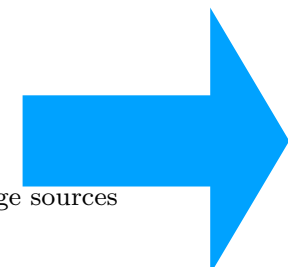
Scalzo...AM +2017



eref
thref
fwhmref
f4ref
flagref

Features

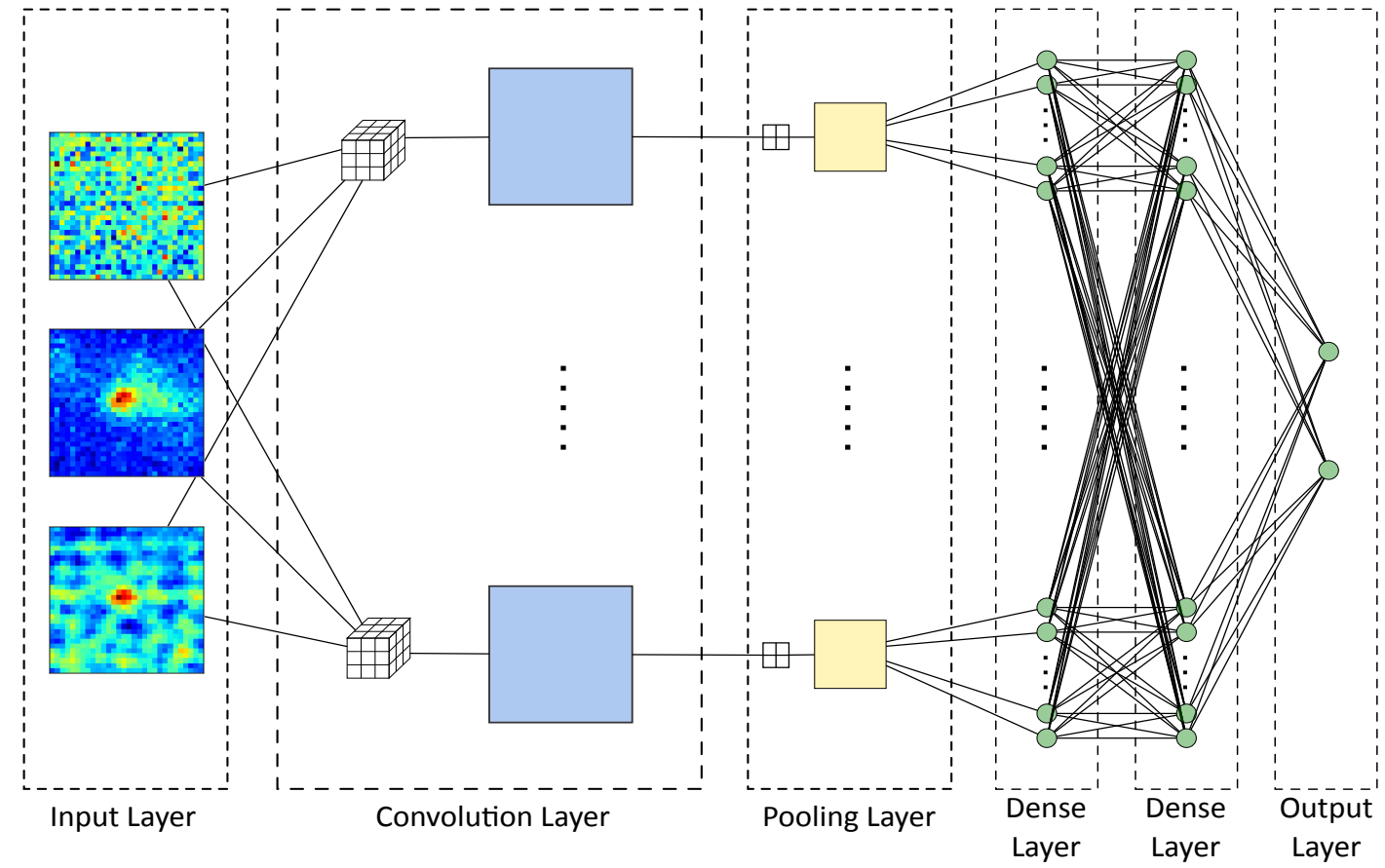
ellipticity of source on template image
semi-major axis of source on template image
full width at half maximum of all template image sources
flux within 4-pixel aperture in template image
SEXTRACTOR source flags in template image



ML classifier
Random Forest

SkyMapper images

Gieseke...AM+ 2017



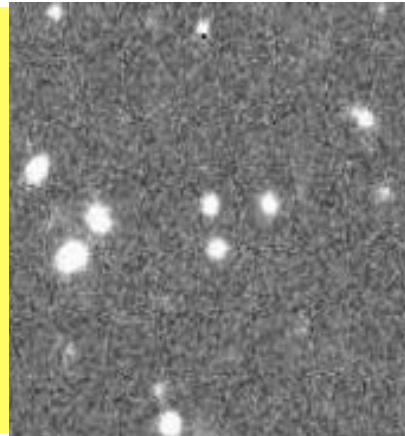
SkyMapper images
augmented for
training

Model	AUC
Random forest	0.9907
Net1(32,64)	0.9914
Net3	0.9972
E2	0.9946

outline

Machine learning in supernova cosmology

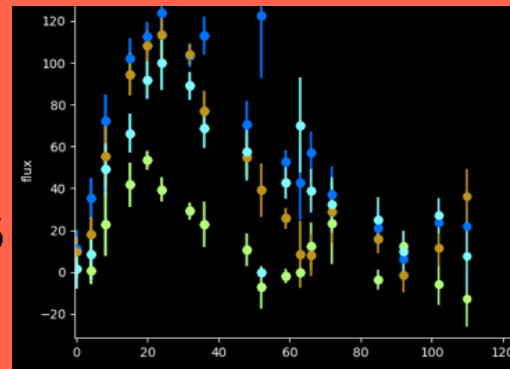
Part I: real vs. bogus



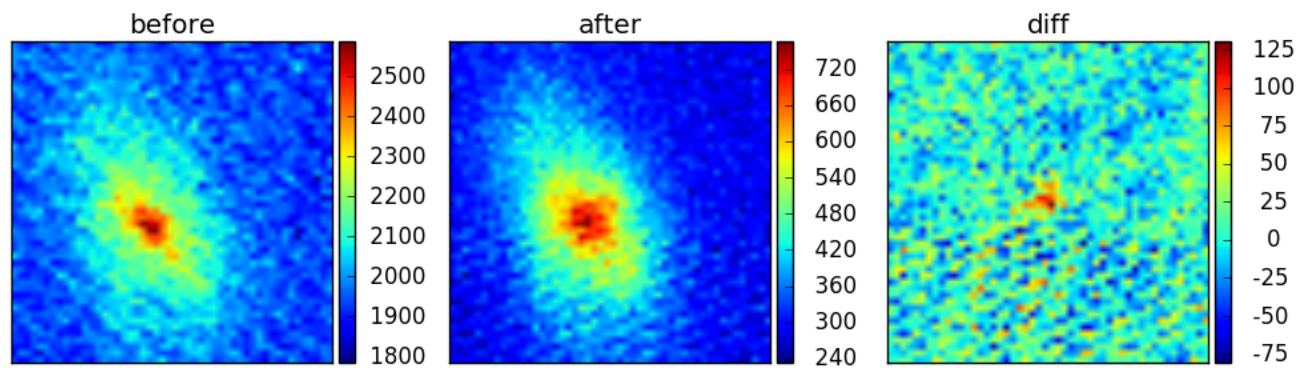
Is it a supernova? Which type?

Part II: typing with photometry

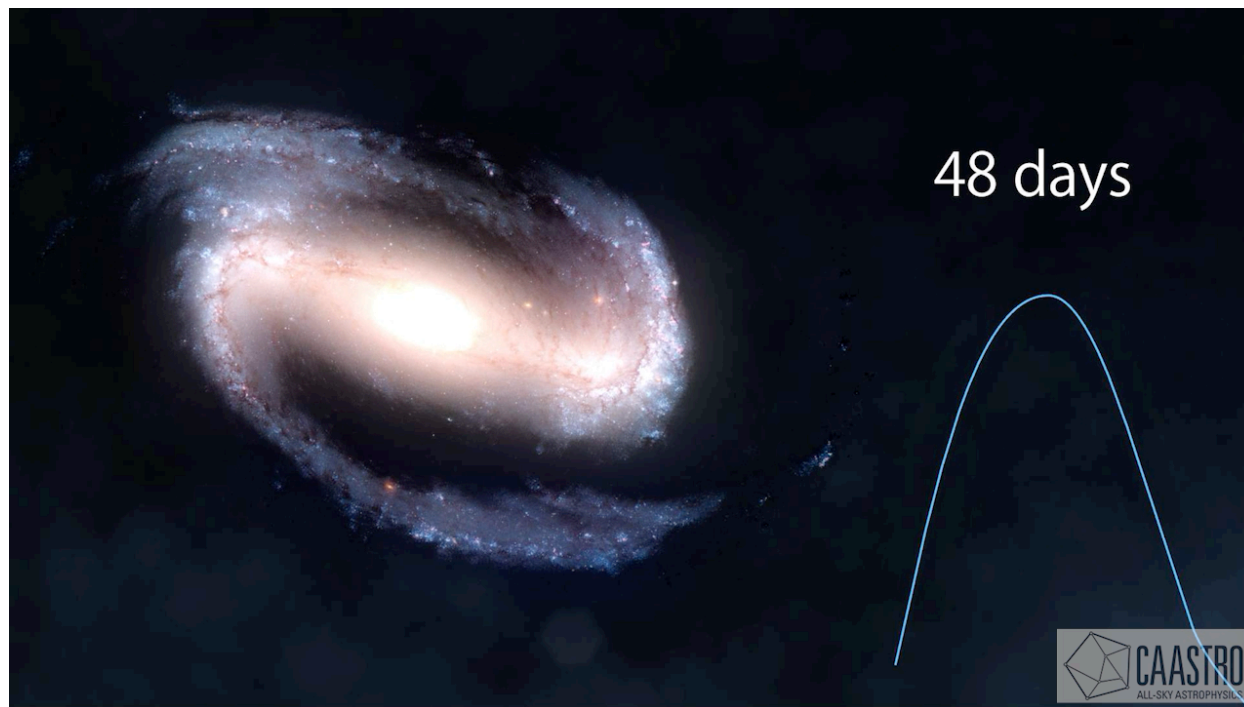
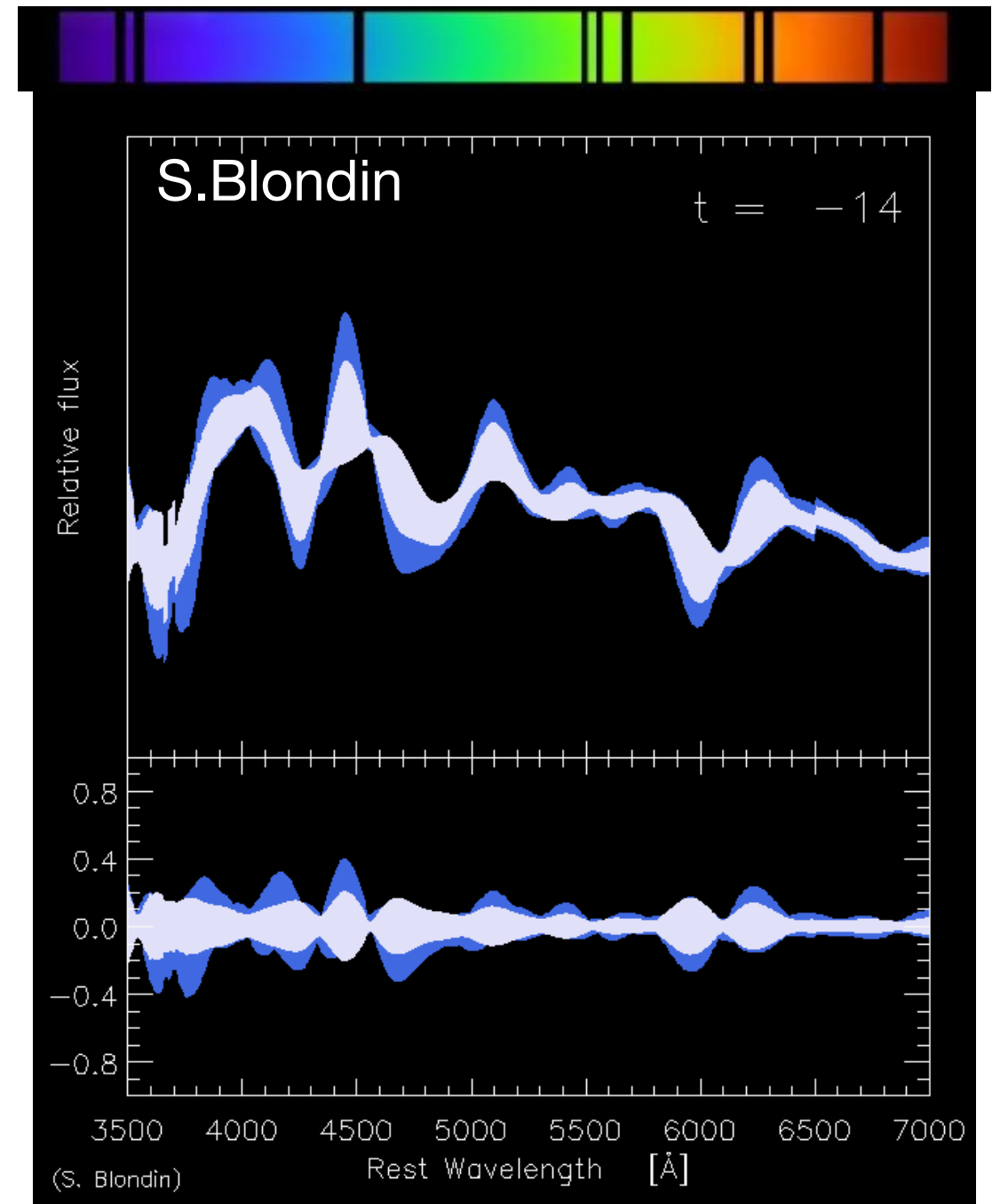
1. Datasets: PLAsTiCC
2. ML classification issues
 1. Representativity
 2. Incompleteness
 3. “probabilities” for cosmology
3. FINK broker



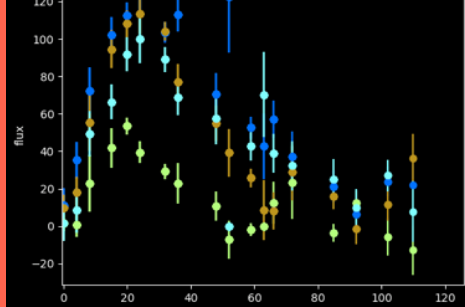
typing supernovae with spectroscopy



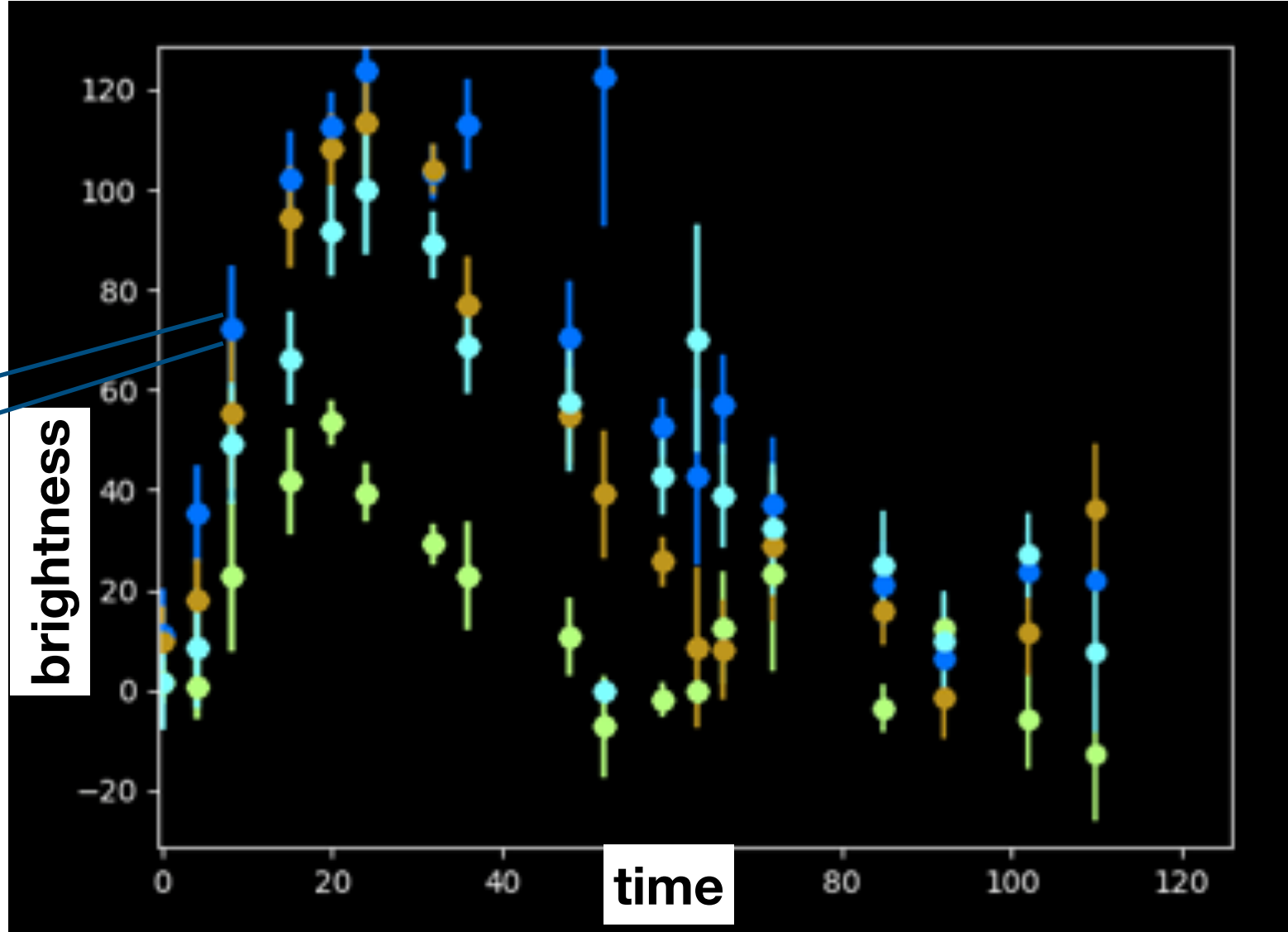
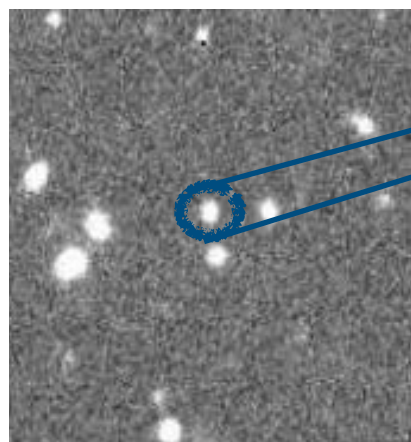
Is it a supernova? Which type?



Part II: typing with photometry



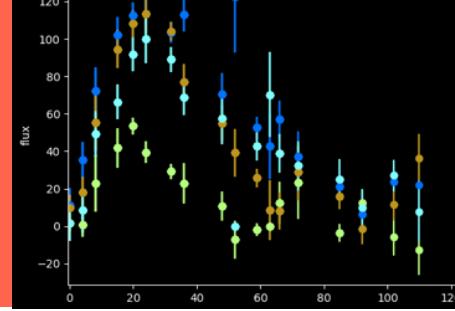
Our goal



Instead of spectra, we just see the evolution on brightness in some wavelengths

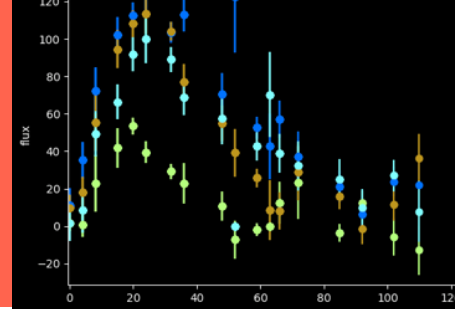
classification of SN types

Part II: typing with photometry



we need “simulated” datasets:

- to evaluate our methods
- to train ML classifiers



we need “simulated” datasets:

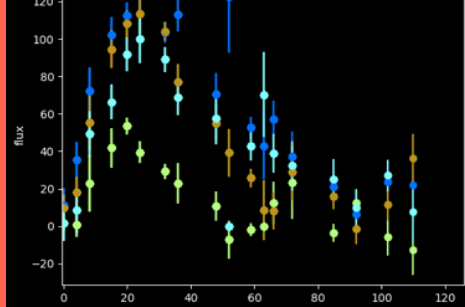
- to evaluate our methods
- to train ML classifiers

The Photometric LSST Astronomical Time-series Classification Challenge (PLAsTiCC): Data set

*The PLAsTiCC team:,¹ Tarek Allam Jr.,² Anita Bahmanyar,³ Rahul Biswas,⁴ Mi Dai,⁵
Lluís Galbany,⁶ Renée Hložek,³ Emille E. O. Ishida,⁷ Saurabh W. Jha,⁵ David O. Jones,⁸
Richard Kessler,⁹ Michelle Lochner,^{10, 11} Ashish A. Mahabal,^{12, 13} Alex I. Malz,^{14, 15}
Kaisey S. Mandel,^{16, 17} Juan Rafael Martínez-Galarza,¹⁸ Jason D. McEwen,²
Daniel Muthukrishna,¹⁶ Gautham Narayan,¹⁹ Hiranya Peiris,^{4, 20} Christina M. Peters,³
Kara Ponder,²¹ and Christian N. Setzer⁴*

(LSST Dark Energy Science Collaboration and the LSST Transients and Variable Stars Science Collaboration)

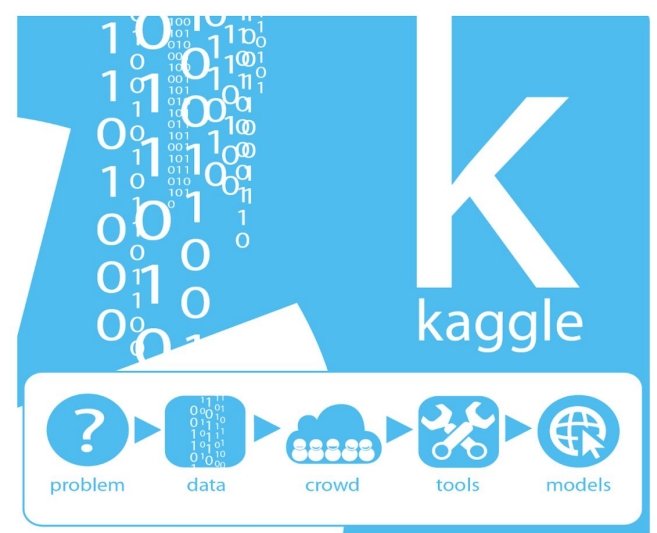
Part II: typing with photometry



Datasets: PLAsTiCC

we need “simulated” datasets:

- to evaluate our methods
- to train ML classifiers



Featured Prediction Competition

PLAsTiCC Astronomical Classification

Can you help make sense of the Universe?

Overview Data Kernels Discussion Leaderboard Rules

3.5K , 14 classes in training

3 M, 14+1 classes in target

Launched 28 Sep 2018

Description

Help some of the world's leading astronomers grasp the deepest properties of the universe.

Evaluation

Prizes

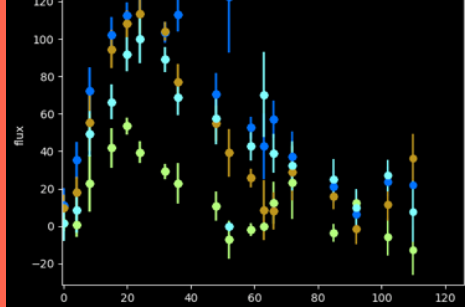
Timeline

PLAsTiCC's Team

The human eye has been the arbiter for the classification of astronomical sources in the night sky for hundreds of years. But a new facility -- the **Large Synoptic Survey Telescope (LSST)** -- is about to revolutionize the field, discovering 10 to 100 times more astronomical sources that vary in the night sky than we've ever known. Some of these sources will be completely unprecedented!

credit: E. Ishida

Part II: typing with photometry

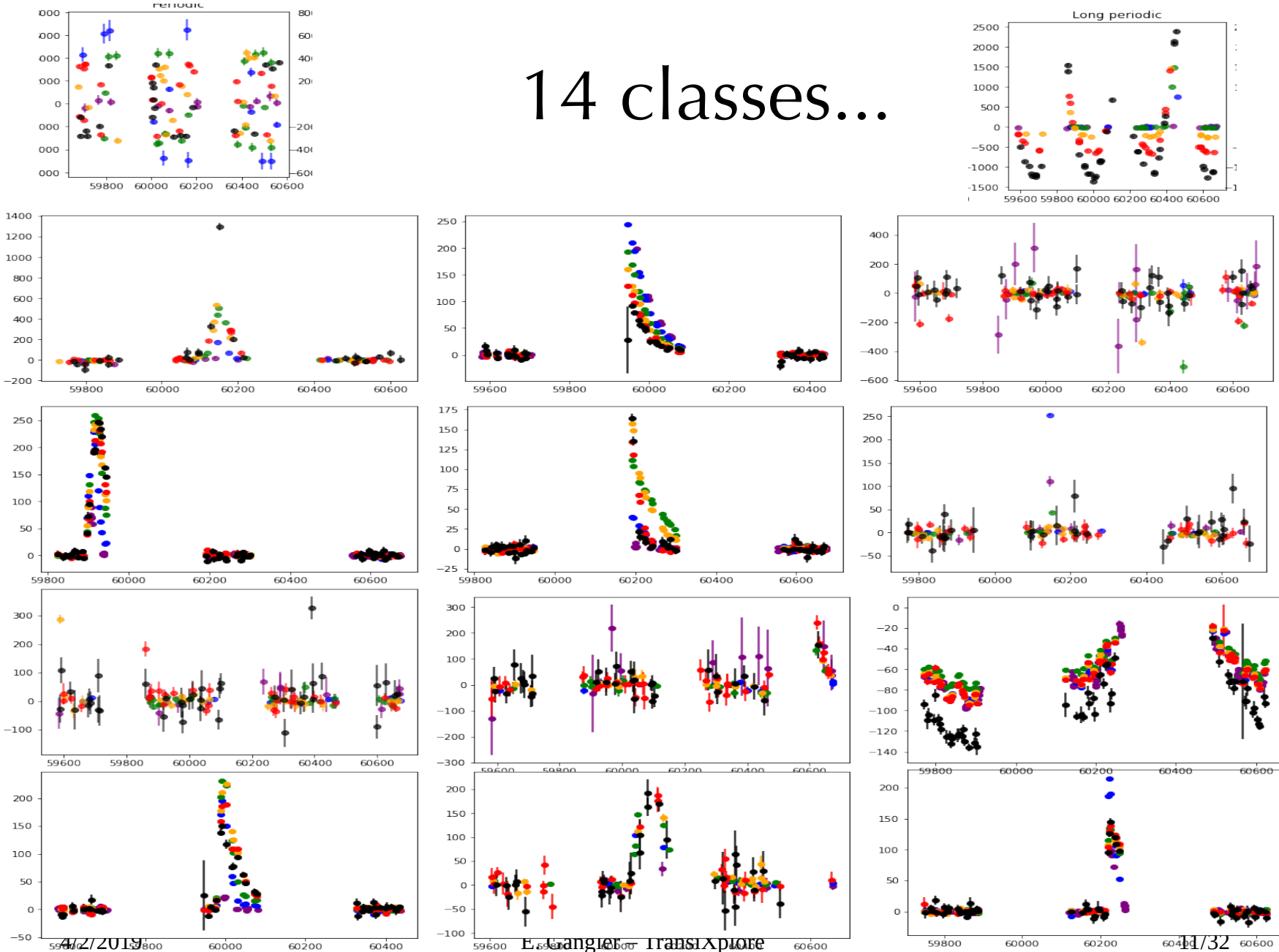


Datasets: PLAsTiCC

14 classes...

metric:
each class is roughly equally important for the final score.

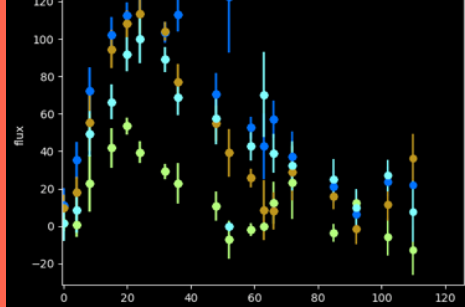
$$\text{Log Loss} = - \left(\frac{\sum_{i=1}^M w_i \cdot \sum_{j=1}^{N_i} \frac{y_{ij}}{N_i} \cdot \ln p_{ij}}{\sum_{i=1}^M w_i} \right)$$



model class num ^a : name	model description
90: SNIa	WD detonation, Type Ia SN
67: SNIa-91bg	Peculiar type Ia: 91bg
52: SNIax	Peculiar SNIax
42: SNI	Core Collapse, Type II SN
62: SNIbc	Core Collapse, Type Ibc SN
95: SLSN-I	Super-Lum. SN (magnetar)
15: TDE	Tidal Disruption Event
64: KN	Kilonova (NS-NS merger)
88: AGN	Active Galactic Nuclei
92: RRL	RR lyrae
65: M-dwarf	M-dwarf stellar flare
16: EB	Eclipsing Binary stars
53: Mira	Pulsating variable stars
6: μ Lens-Single	μ -lens from single lens
991: μ Lens-Binary	μ -lens from binary lens
992: ILOT	Intermed. Lum. Optical Trans.
993: CaRT	Calcium Rich Transient
994: PISN	Pair Instability SN
995: μ Lens-String	μ -lens from cosmic strings
TOTAL	Sum of all models

credit: E. Gangler

Part II: typing with photometry



Datasets: PLAsTiCC

Featured Prediction Competition

PLAsTiCC Astronomical Classification

Can you help make sense of the Universe?

Prize Money: \$25,000

LSST Project · 1,094 teams · 17 days ago

Overview Data Kernels Discussion Leaderboard Rules Team My Submissions New Topic

18 topics Follow Sort by Relevance

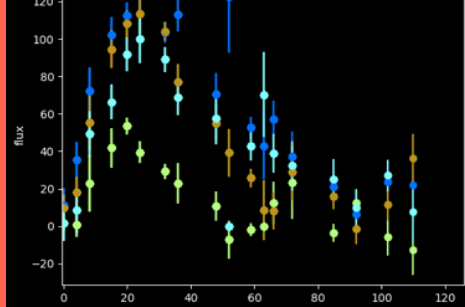
All Mine Upvoted solution

- 31 Source code for a complete solution
JohannesBuchner 2 months ago
last comment by Ivan Petrov 1mo ago
10
- 96 4th Place Solution with Github Repo
AhmetErdem 17 days ago
last comment by Debashish Barua 10d ago
45
- 78 Congrats and 8th place Rapids solution updated!
Jiwei Liu 17 days ago
last comment by Blonde 14d ago
23
- 170 Overview of 1st place solution
Kyle Boone 17 days ago
last comment by Rajesh D 3d ago
81
- 43 5th Place Partial Solution (RNN)
Kun Hao Yeh 17 days ago
last comment by Aryan Pariani 13d ago
11
- 72 Solution #5 tidbits (revised with code)
CPMP 17 days ago
last comment by Blonde 4d ago
37
- 66 14th place solution
Belinda Trotta 17 days ago
last comment by LongYin 2d ago
20
- 61 2nd-Place Solution Notes
Silogram 17 days ago
last comment by SD 6d ago
27
- 51 6th Place Solution Summary
Stefan Stefanov 17 days ago
last comment by olivier 16d ago
10

Solutions posted on Kaggle

- 55 #13 Solution, true story: tries and fails
Blonde 16 days ago
last comment by SooperDoop 8d ago
19
- 15 PostProcess Trick - 21st place Partial Solution
fatihöztürk 16 days ago
last comment by Murat KORKMAZ 16d ago
3
- 22 21st Solution ~super tough road~
takuoko 16 days ago
last comment by takuoko 16d ago
11
- 24 19th Place Solution
ONODERA 16 days ago
last comment by Vig Nam 15d ago
4
- 28 11th solution - very basic but may different methods
SimonChen 16 days ago
last comment by SimonChen 13d ago
15
- 11 A solution and some learnings
Helgi 15 days ago
last comment by Avinash Tayade 14d ago
4
- 17 12th Place Solution
Daniel Bi 15 days ago
last comment by go5paopao 7d ago
4
- 32 20th Place Solution
Giba 15 days ago
last comment by Giba 14d ago
7
- 20 9th place solution
Albert Garreta 14 days ago
last comment by Albert Garreta 11d ago
4

credit: M. Dai



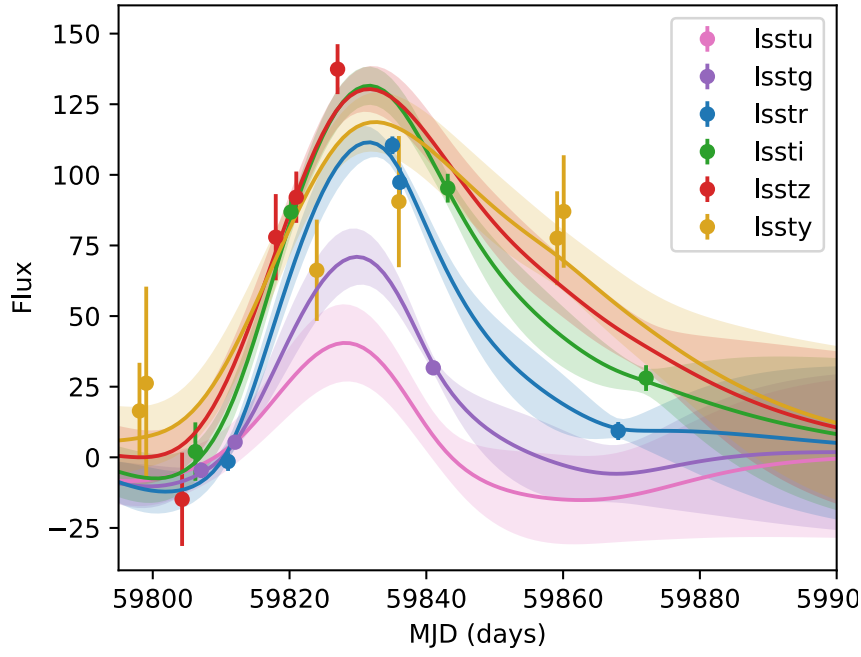
Avocado: Photometric Classification of Astronomical Transients with Gaussian Process Augmentation

Winning solution

KYLE BOONE^{1,2}

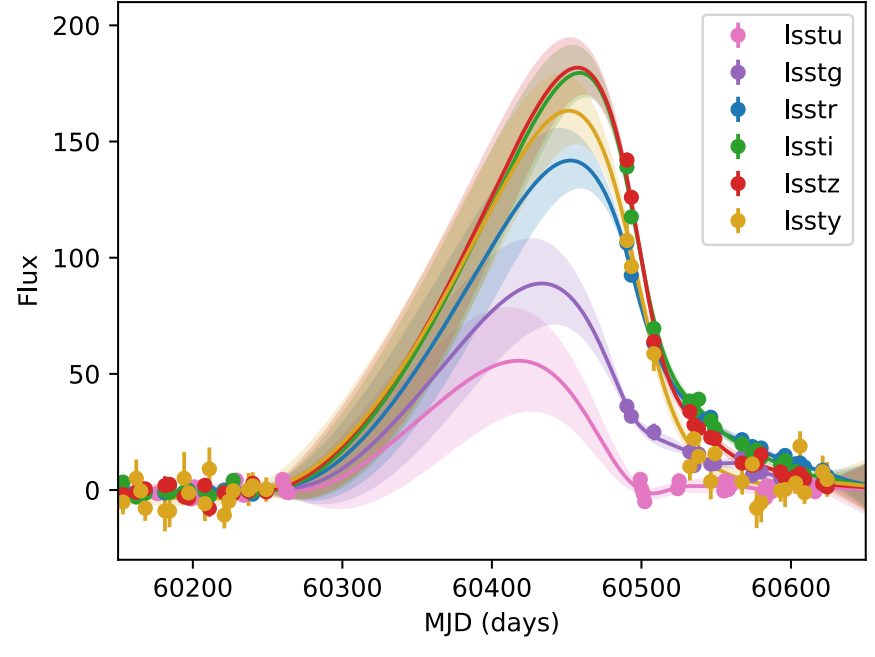
GP modelling

WFD survey, $z=0.35$



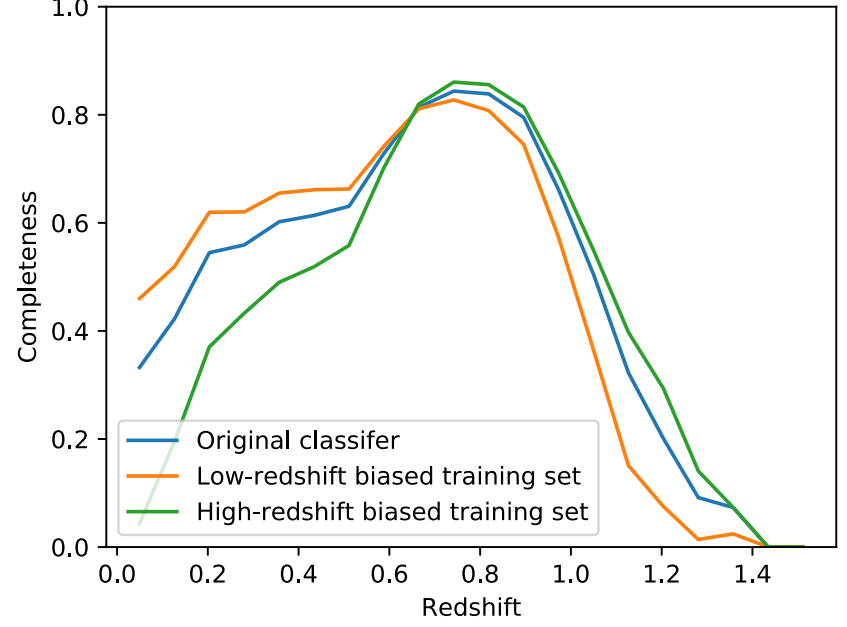
feature engineering

DDF survey, $z=0.19$



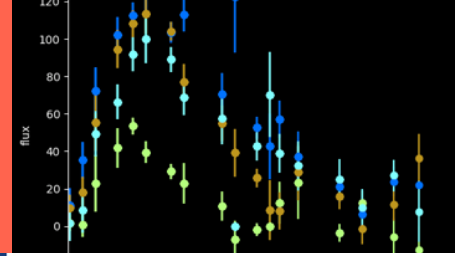
data augmentation

WFD survey, flat-weighted classifier, bias test



+ boosted decision tree

Part II: typing with photometry



Datasets: PLAsTiCC

Featured Prediction Competition

PLAsTiCC Astronomical Classification
Can you help make sense of the Universe?

\$25,000
Prize Money

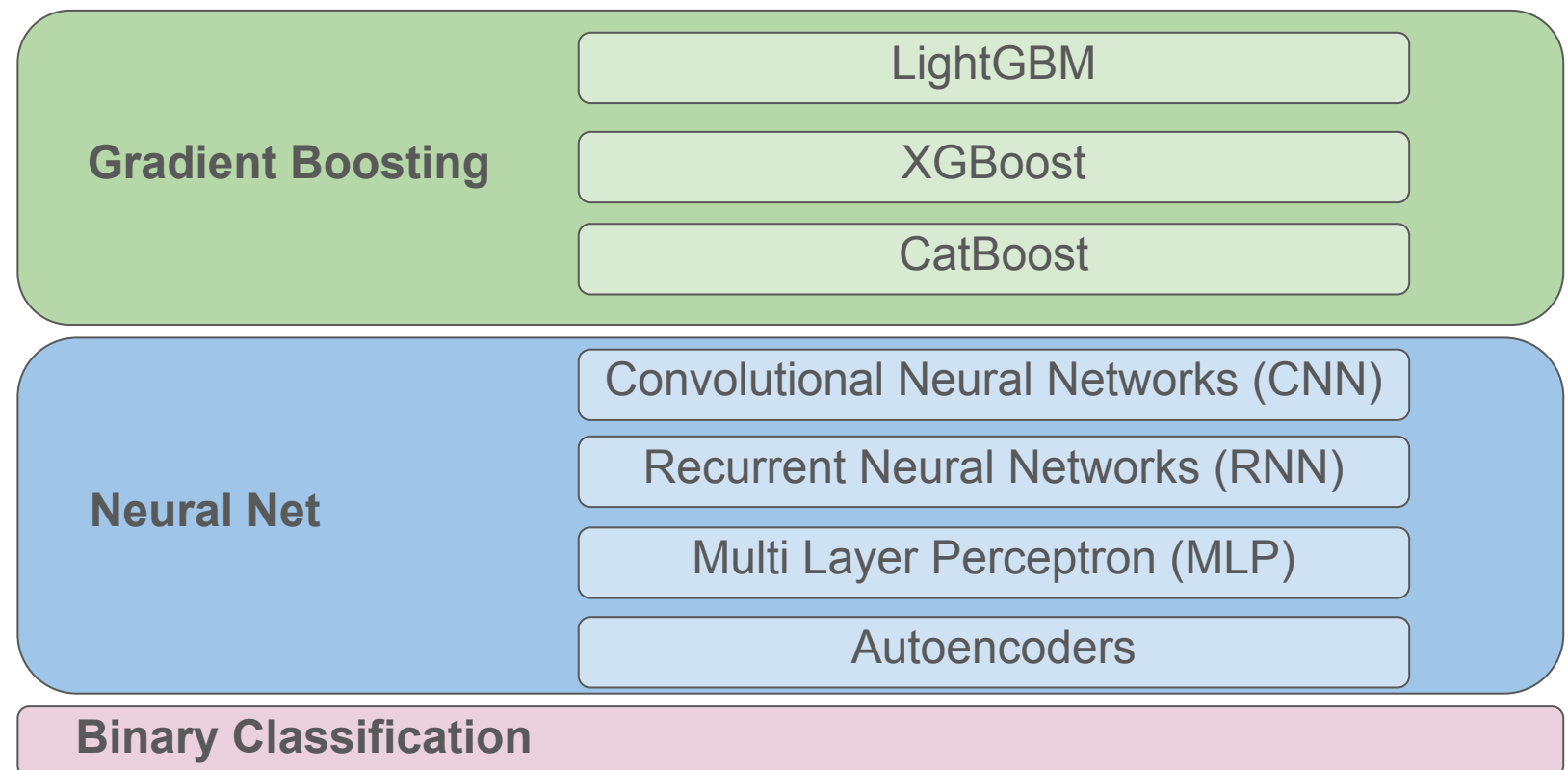
LSST Project · 1,094 teams · 17 days ago

Solutions posted on Kaggle

Useful Features

- Light curve fitting -- Bazin, GP, template fitting (SALT2, SN templates)
- Flux ratio (color)
- Flux difference
- Host galaxy photo-z
- $\text{flux} * \text{distance} ** 2$

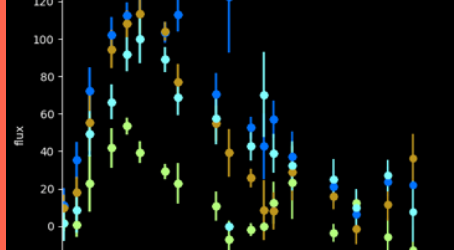
Popular Models among Kagglers



Credit M. Dai

Part II: typing with photometry

Datasets: PLAsTiCC



Featured Prediction Competition

PLAsTiCC Astronomical Classification
Can you help make sense of the Universe?

\$25,000
Prize Money

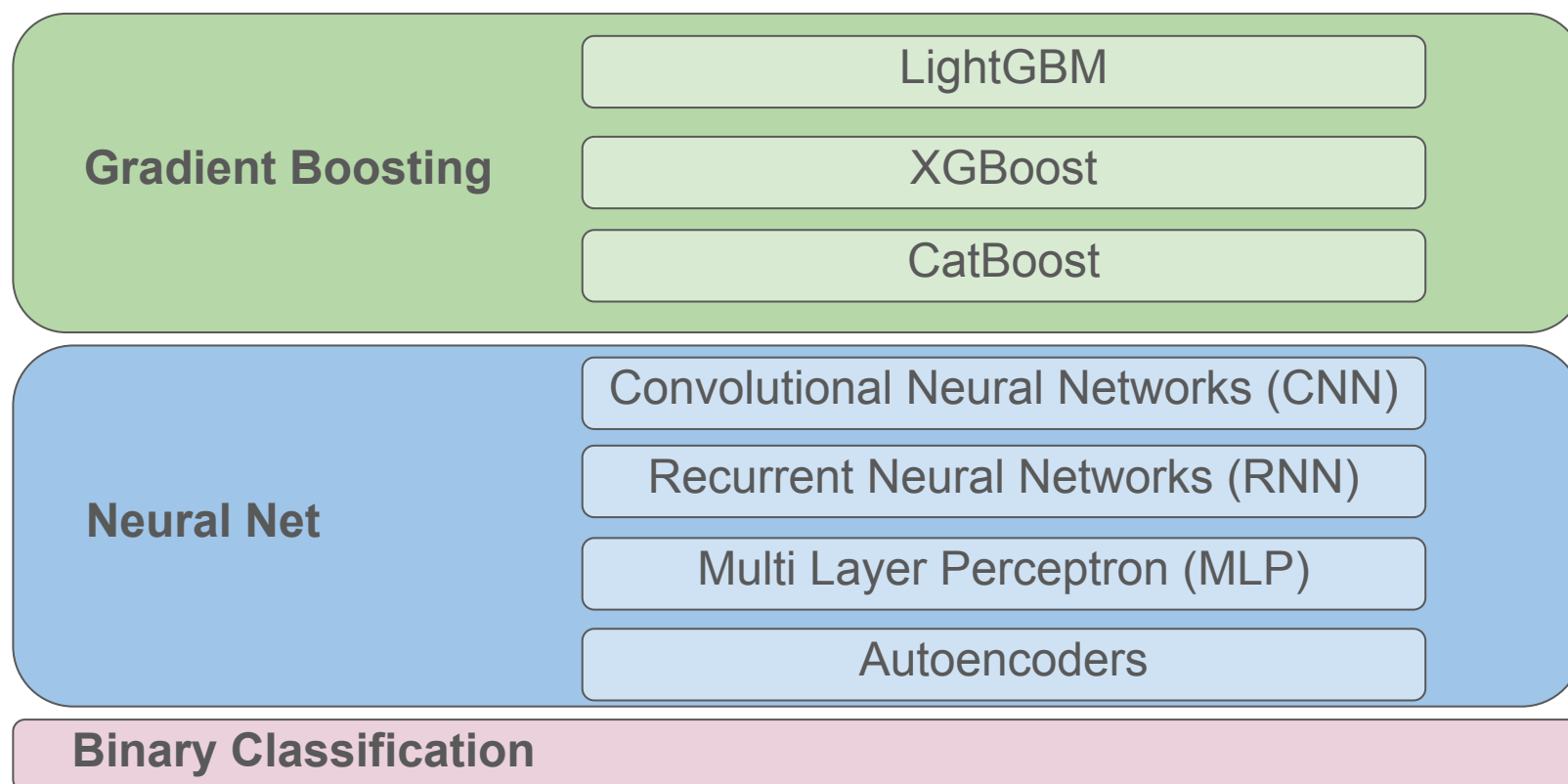
LSST Project · 1,094 teams · 17 days ago

Solutions posted on Kaggle

Useful Features

- Light curve fitting -- Bazin, GP, template fitting (SALT2, SN templates)
- Flux ratio (color)
- Flux difference
- Host galaxy photo-z
- $\text{flux} * \text{distance} ** 2$

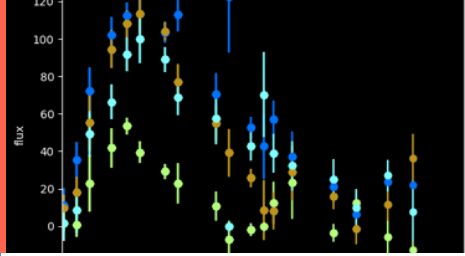
Popular Models among Kagglers



Credit M. Dai

- common algorithms perform great, e.g. BDTs
- feature extraction is key (domain knowledge + irregular time series)
- labeled set (for training) was crucial, not large enough, not representative of the test set

Part II: typing with photometry



Datasets: PLAsTiCC

Featured Prediction Competition

PLAsTiCC Astronomical Classification
Can you help make sense of the Universe?

\$25,000
Prize Money

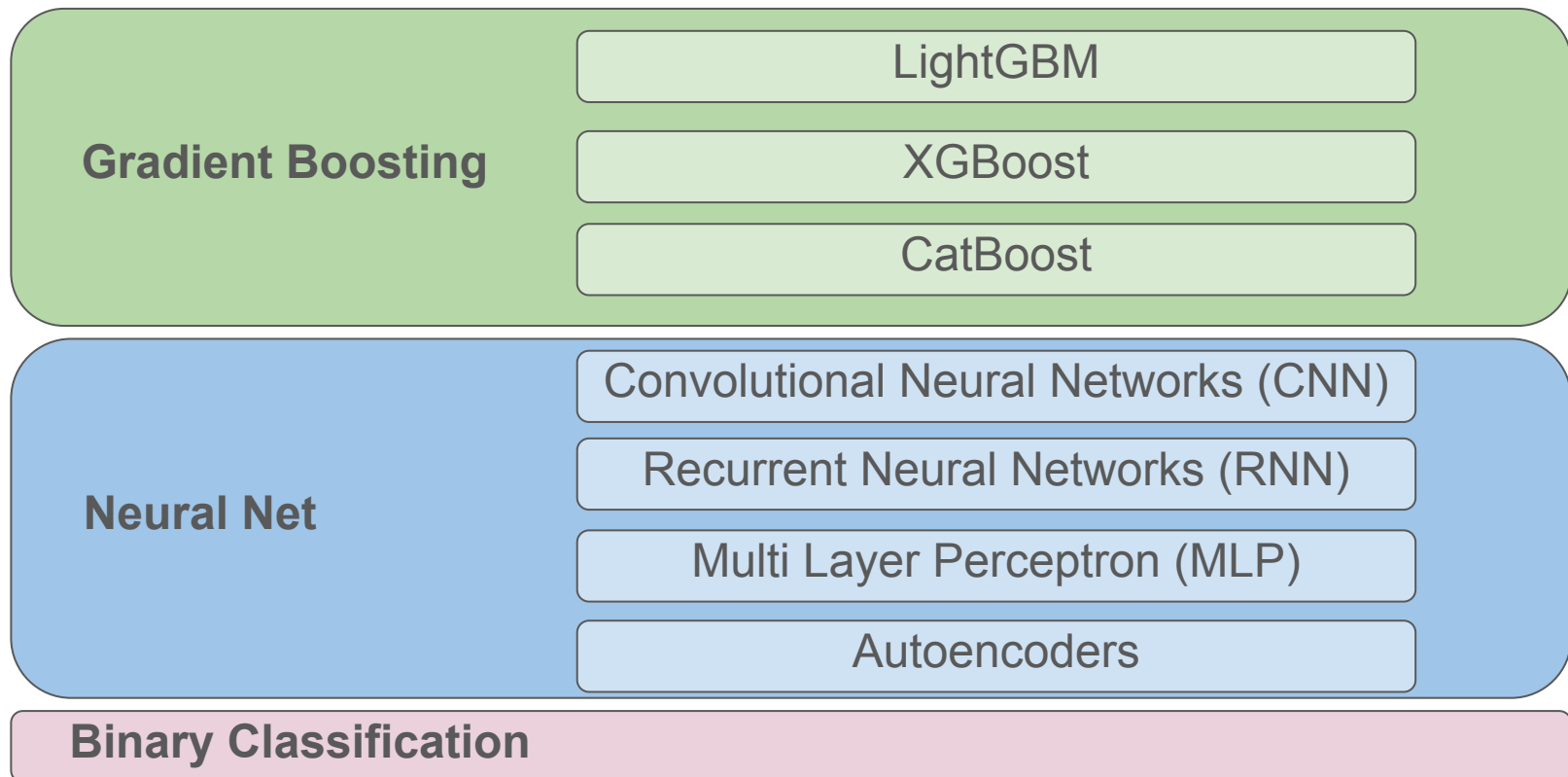
LSST Project · 1,094 teams · 17 days ago

Solutions posted on Kaggle

Useful Features

- Light curve fitting -- Bazin, GP, template fitting (SALT2, SN templates)
- Flux ratio (color)
- Flux difference
- Host galaxy photo-z
- $\text{flux} * \text{distance} ** 2$

Popular Models among Kagglers

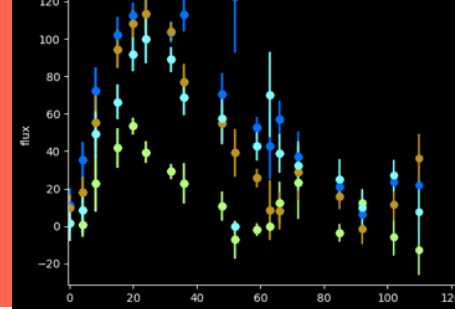


Credit M. Dai

But feature extraction biases samples!

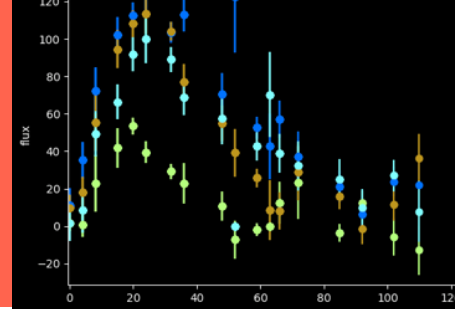
- common algorithms perform great, e.g. BDTs
- feature extraction is key (domain knowledge + **irregular time series**)
- labeled set (for **training**) was crucial, not large nough, **not representative of the test set**

Part II: typing with photometry



Möller & de Boissière 2019

Part II: typing with photometry

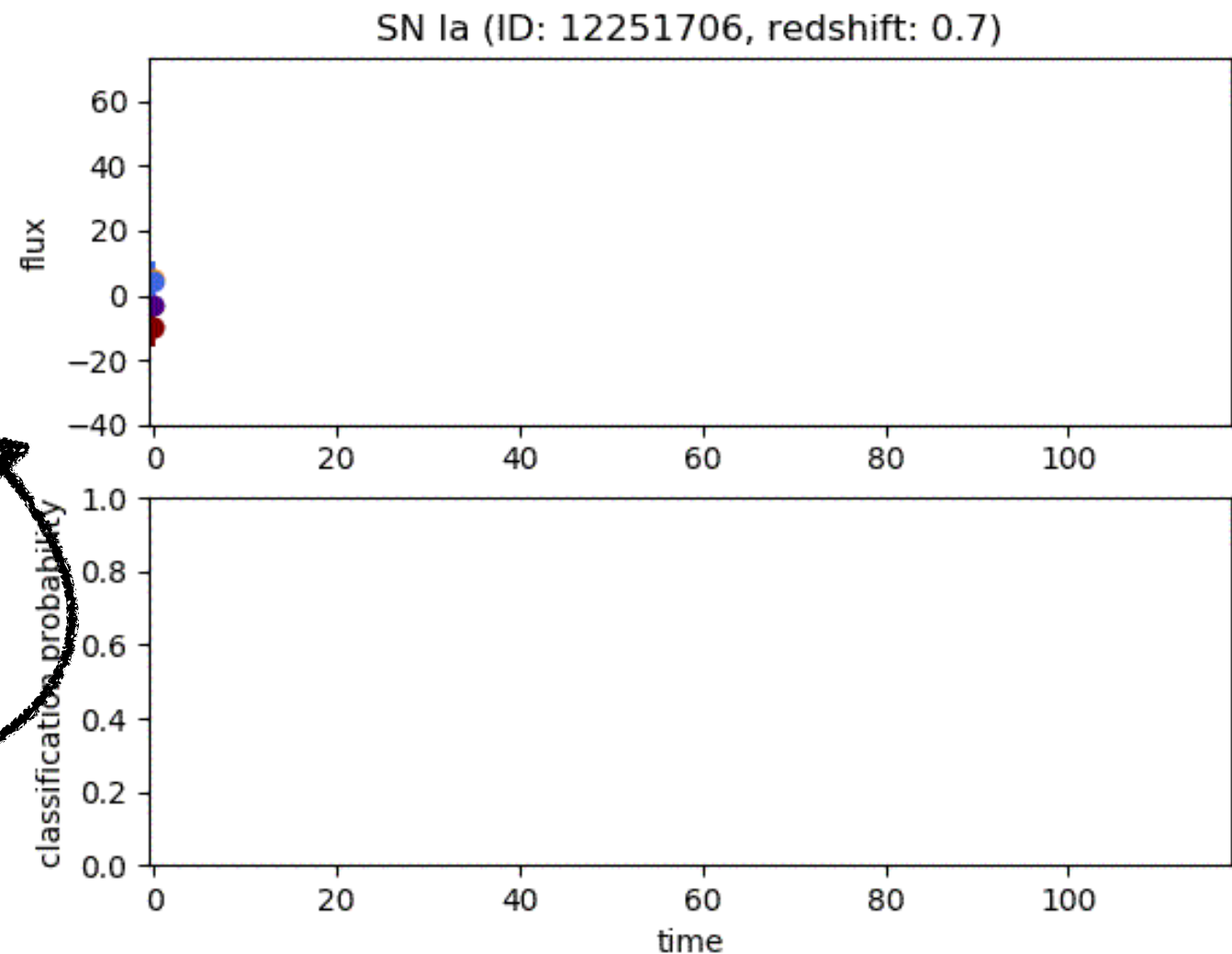


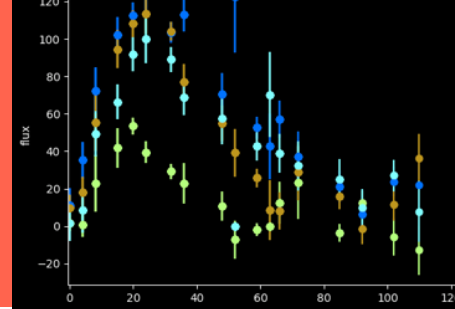
Möller & de Boissière 2019

Inputs:

- Flux in different band-passes
- Flux measurement errors
- Time-step between measurements
- Optional: other features (e.g. host-redshift)

Allows using irregular time series without extrapolation!

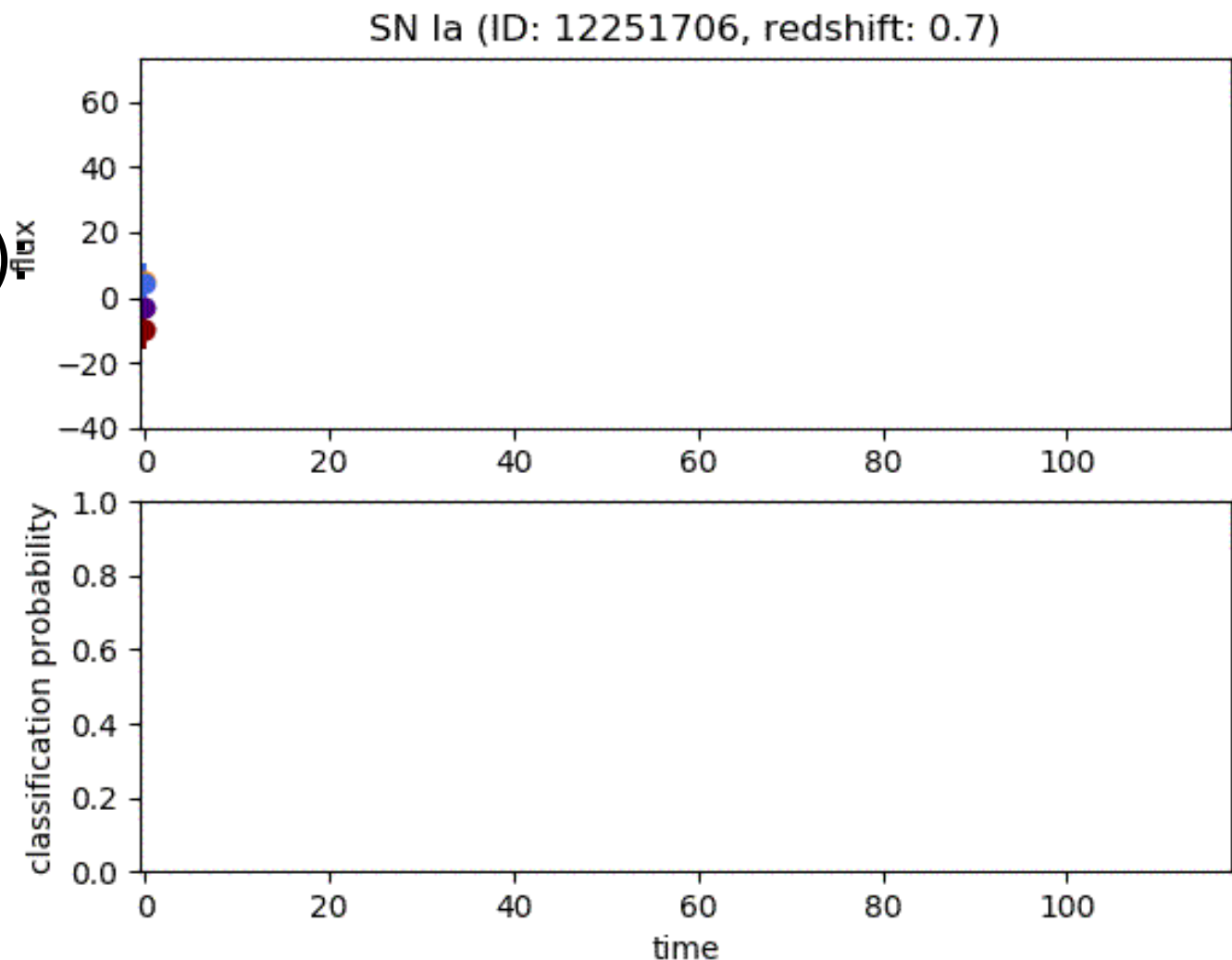




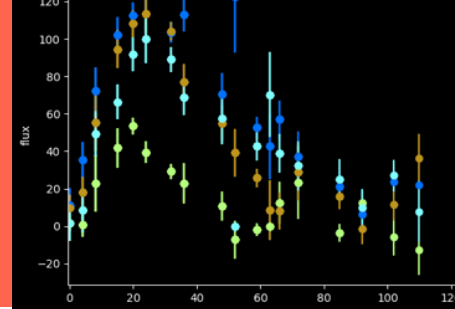
Möller & de Boissière 2019

Deep learning for cosmology

- Recurrent Neural Networks (RNNs)
 - LSTM
 - GRU
- Bayesian RNNs
 - Variational (Gal+2016)
 - Bayes by Backprop (Fortunato+2017)
- Convolutional NN



*Other time-series applications with:
RNNs: Charnock & Moss 2016, Muthukrishna+2019
CNN: Kimura +2017*



Möller & de Boissière 2019

Accuracy

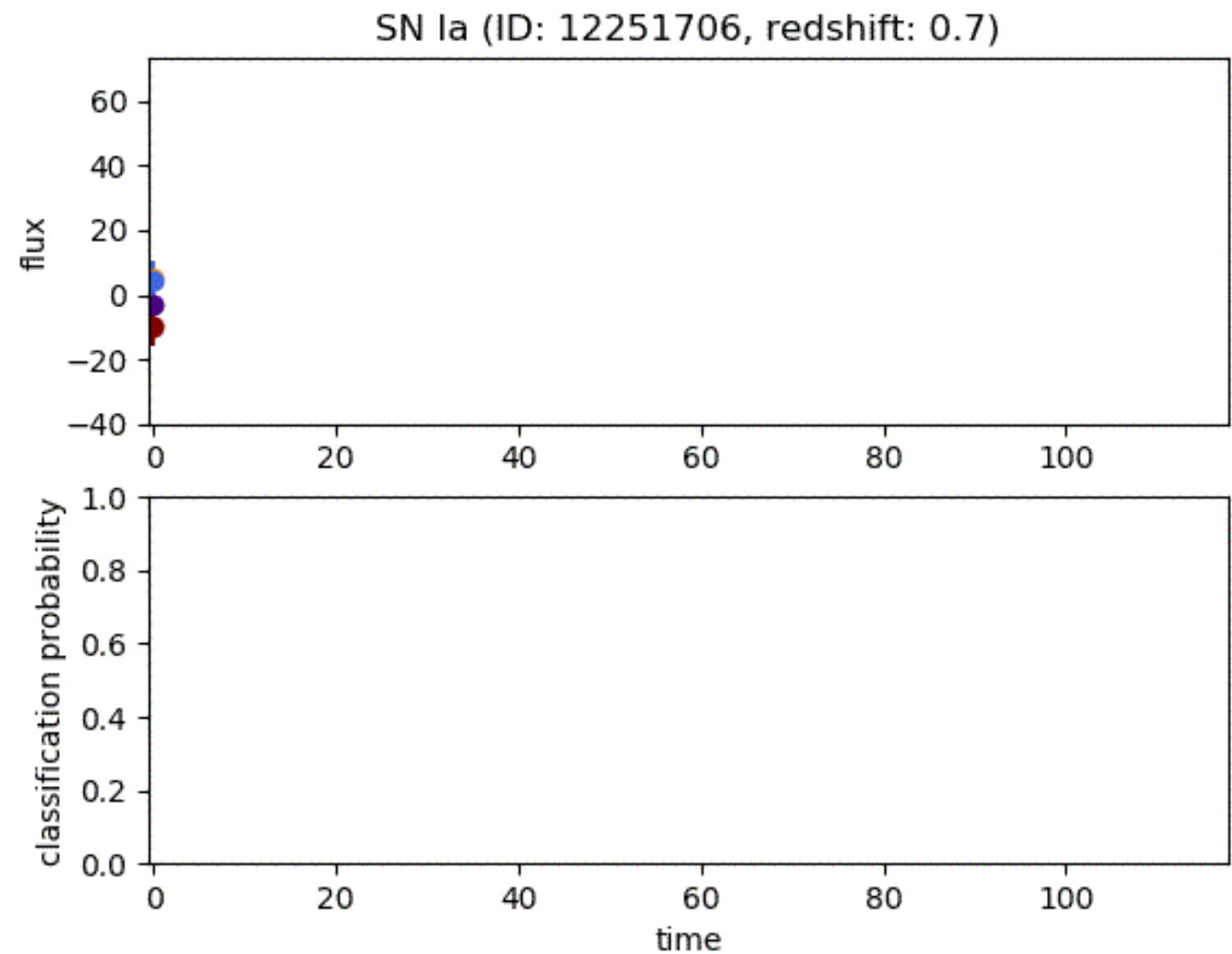
early classification 87.59 ± 0.13

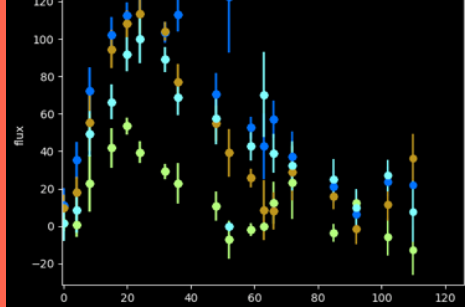
for:

- brokers
- follow-up for promising candidates

reproducible selection functions

Improving trying samples e.g. Ishida + 2018





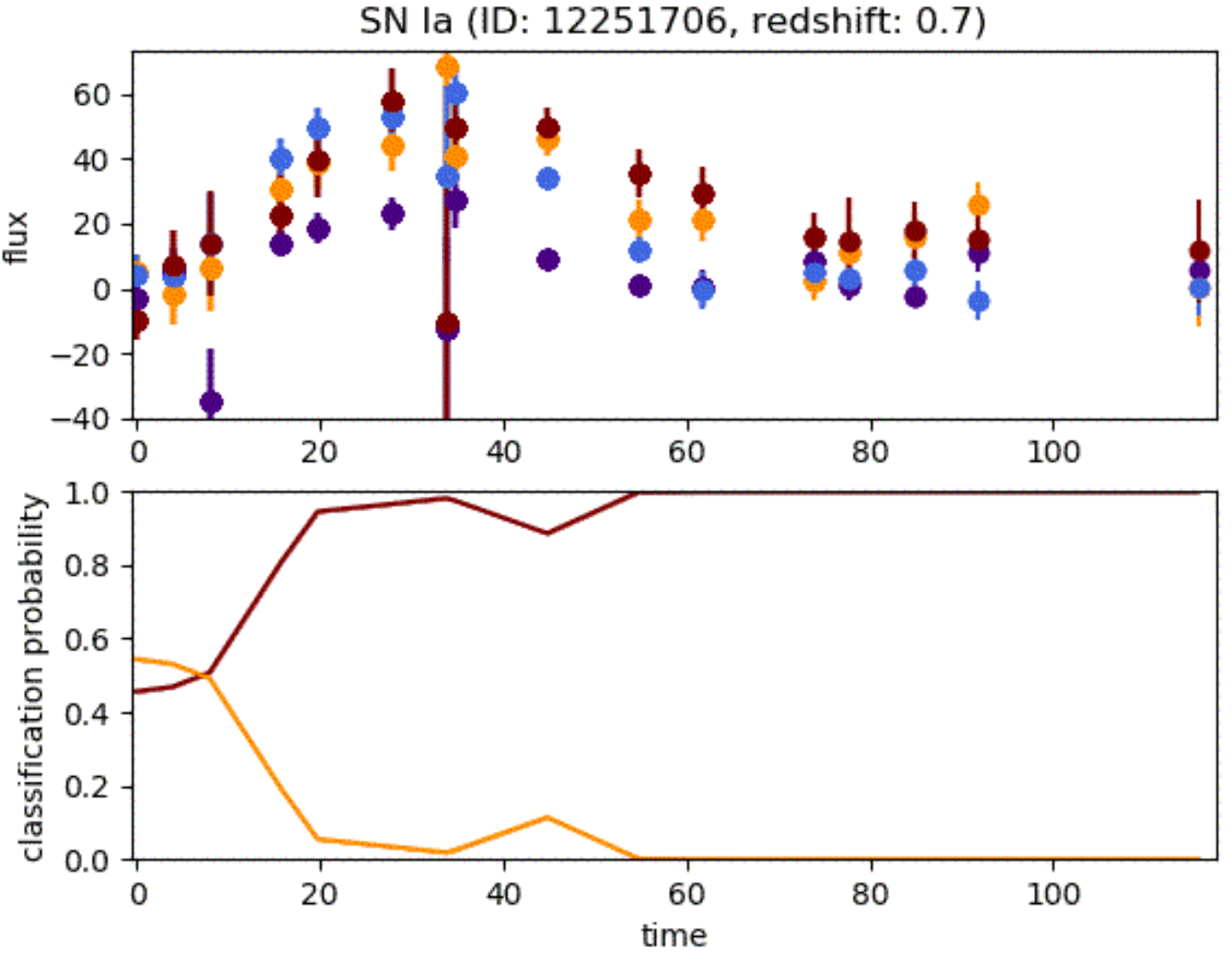
Möller & de Boissière 2019

Accuracy

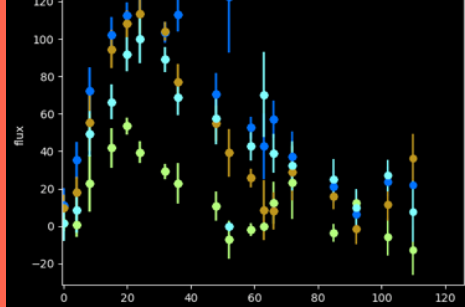
early classification 87.59 ± 0.13

complete 96.97 ± 0.06

- for:
- larger & more reliable samples
 - probing new parameter space



Part II: typing with photometry



Möller & de Boissière 2019

Accuracy

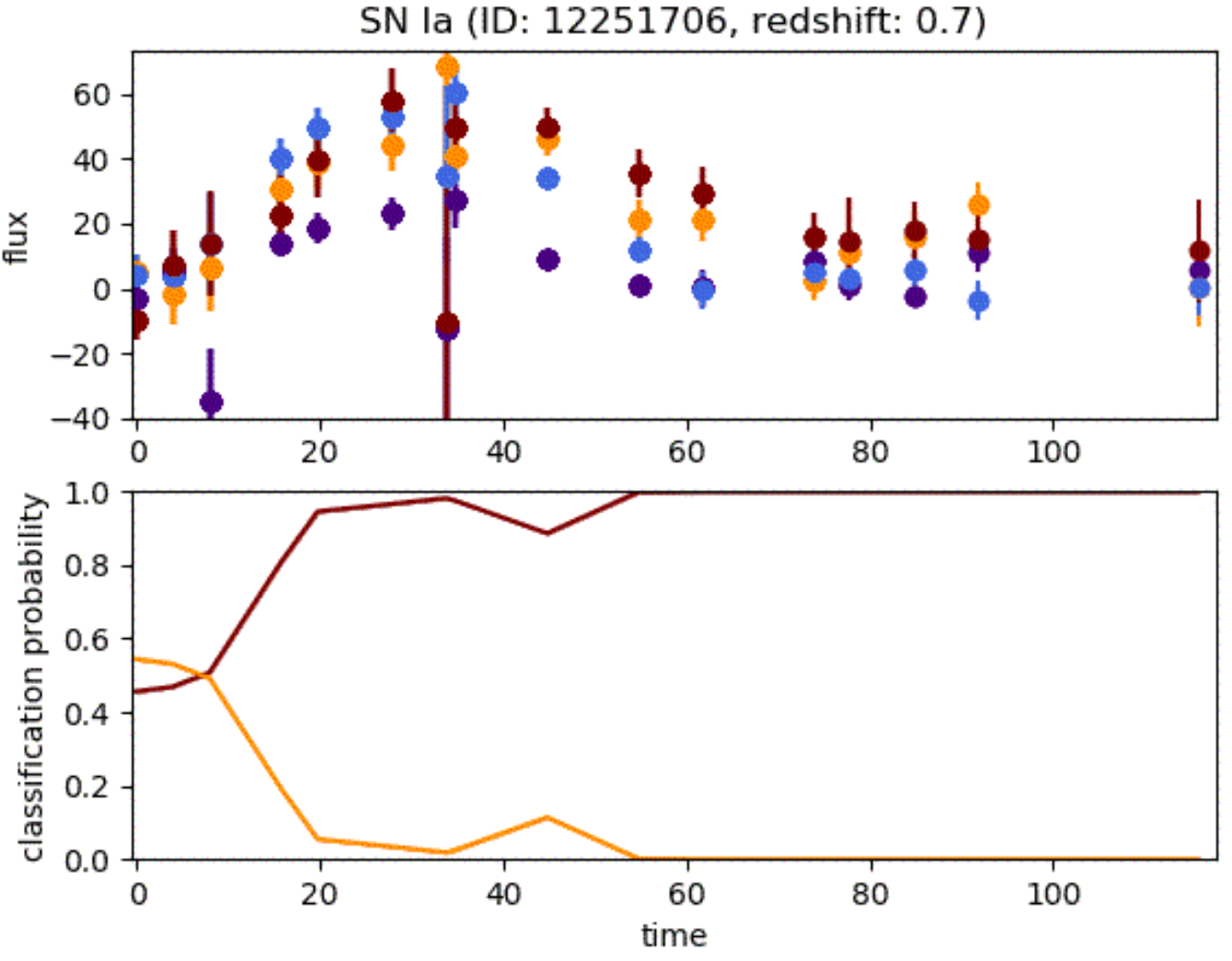
early classification 87.59 ± 0.13

complete 96.97 ± 0.06

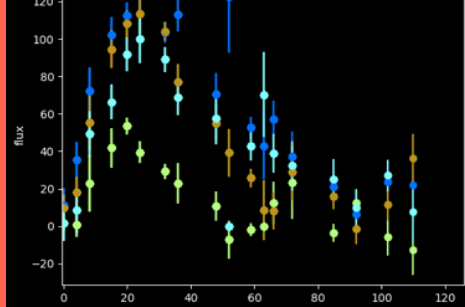
- for:
- larger & more reliable samples
 - probing new parameter space

cosmology, systematic studies
e.g. Hlozek + 2012, Jones+2016

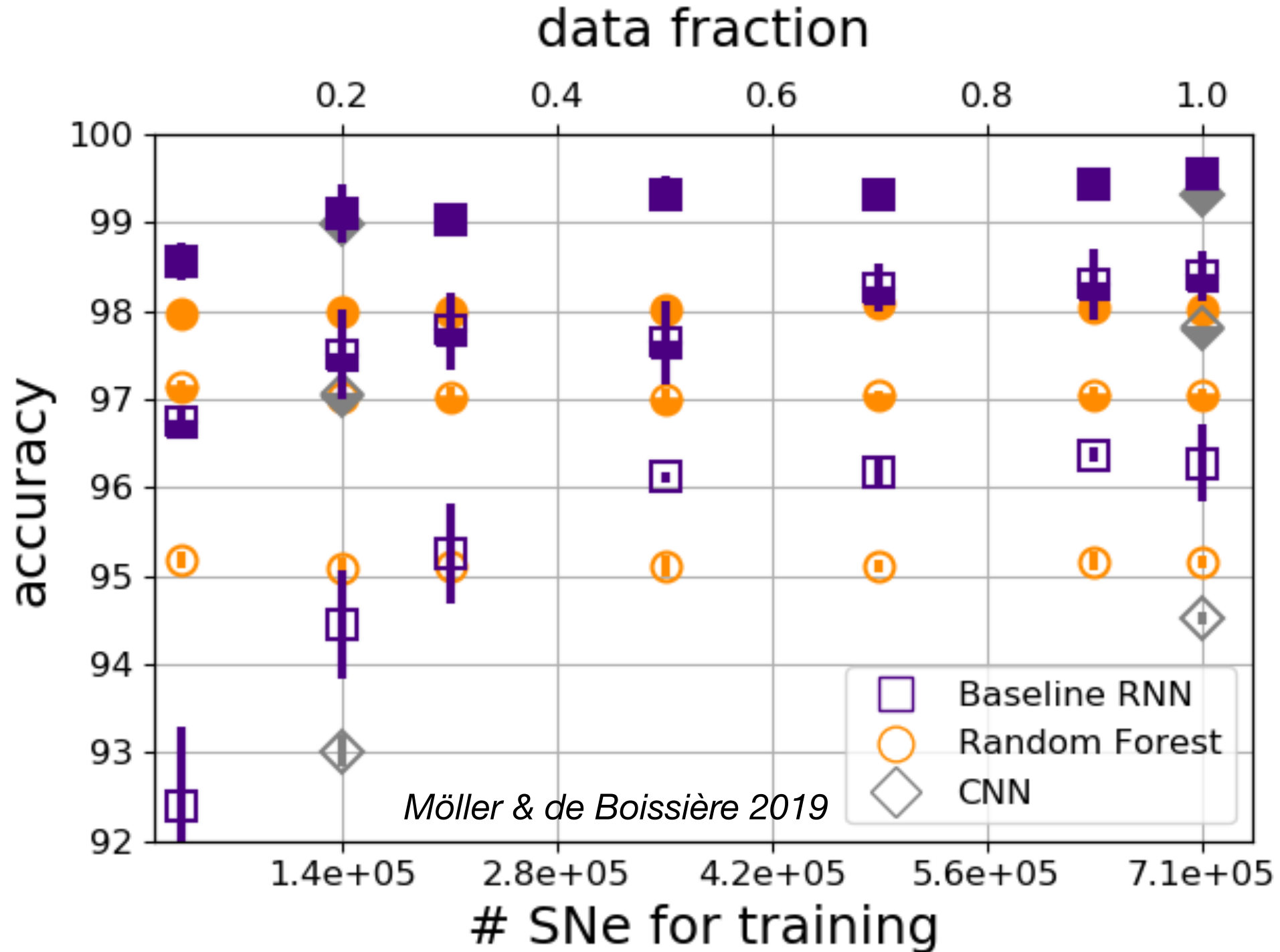
**Cosmology limitation:
Modelling core-collapse contamination**



Current efforts include Hinton + 2018, Vincenzi + 2019

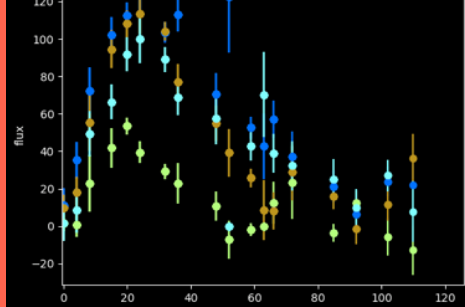


Accuracy

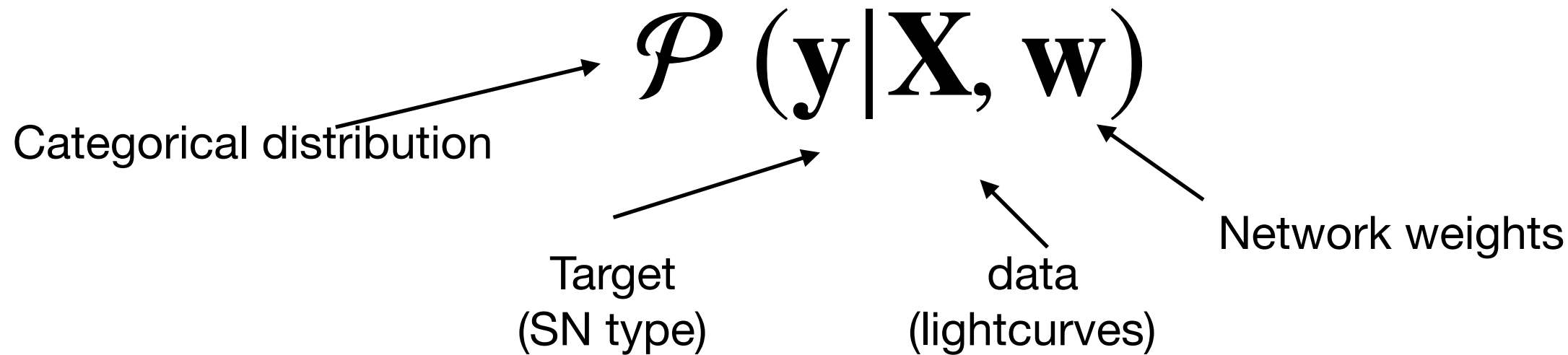


Möller & de Boissière 2019

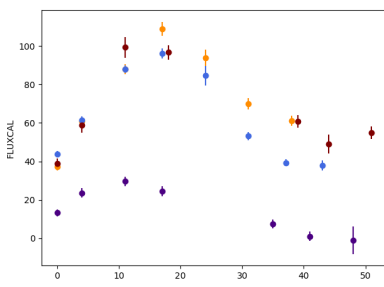
Beware, features in RF are very tuned for Its!

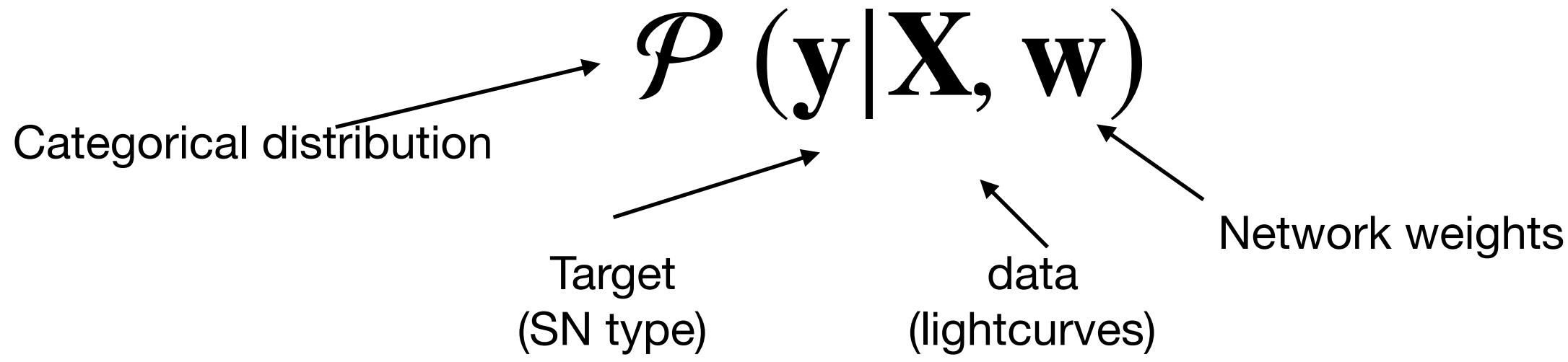
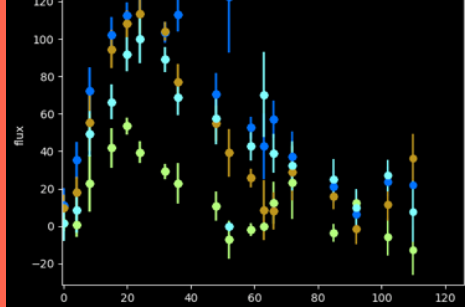


Bayesian NNs

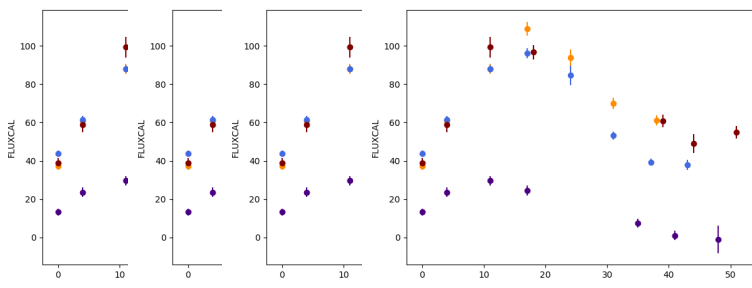


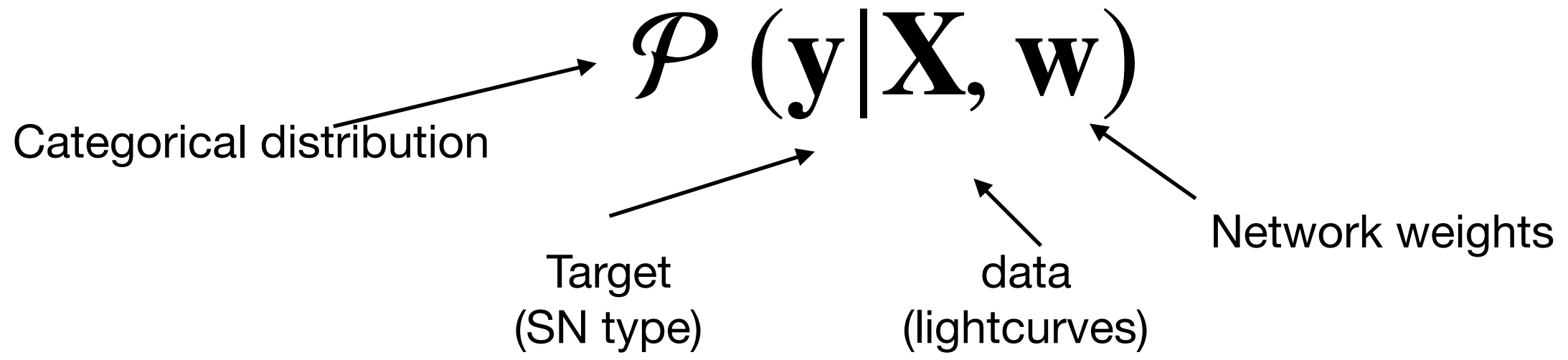
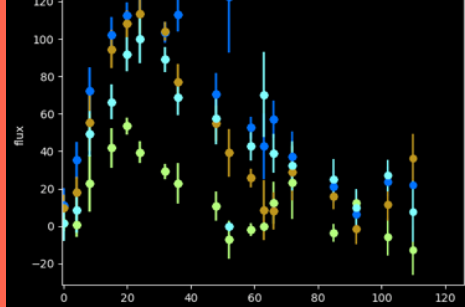
Ia vs. Non Ia



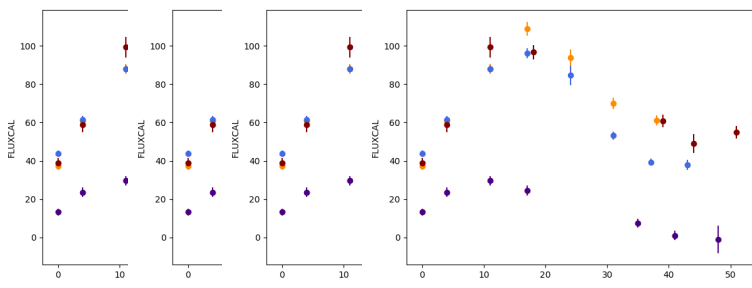


data $\mathcal{D} = (\mathbf{X}_k, \mathbf{y}_k)_{k=1 \dots K}$



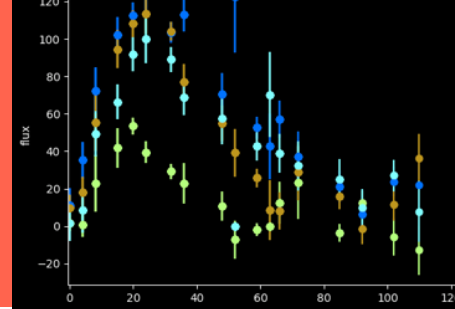


data $\mathcal{D} = (\mathbf{X}_k, \mathbf{y}_k)_{k=1 \dots K}$



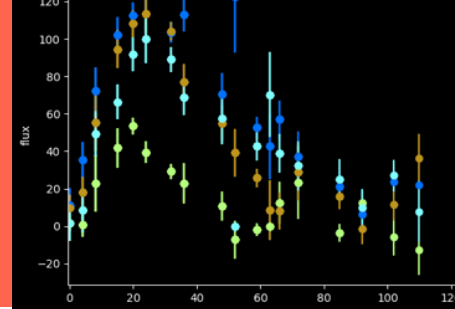
Training minimisation

$$NLL = \min_{\mathbf{w}} \sum_{k=1}^K -\log \mathcal{P}(\mathbf{y}_k | \mathbf{X}_k, \mathbf{w})$$



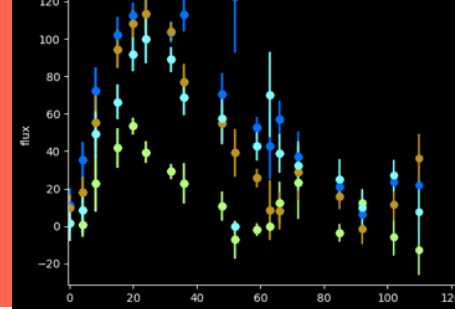
$$\mathcal{P}(\hat{\mathbf{y}} | \mathbf{x}) = \int \mathcal{P}(\hat{\mathbf{y}} | \mathbf{x}, \mathbf{w}) \mathcal{P}(\mathbf{w} | \mathcal{D}) d\mathbf{w}$$

Bayesian: distribution of weights



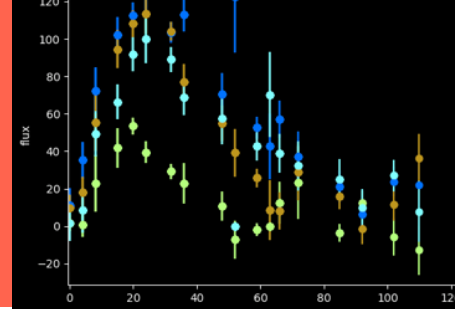
$$\mathcal{P}(\hat{\mathbf{y}} | \mathbf{x}) = \int \mathcal{P}(\hat{\mathbf{y}} | \mathbf{x}, \mathbf{w}) \mathcal{P}(\mathbf{w} | \mathcal{D}) d\mathbf{w}$$

posterior is intractable for deep neural networks



$$\mathcal{P}(\hat{\mathbf{y}} | \mathbf{x}) = \int \mathcal{P}(\hat{\mathbf{y}} | \mathbf{x}, \mathbf{w}) \mathcal{P}(\mathbf{w} | \mathcal{D}) d\mathbf{w}$$

$$\mathcal{P}(\mathbf{w} | \mathcal{D}) \approx q(\mathbf{w} | \theta) \quad \text{variational distribution}$$

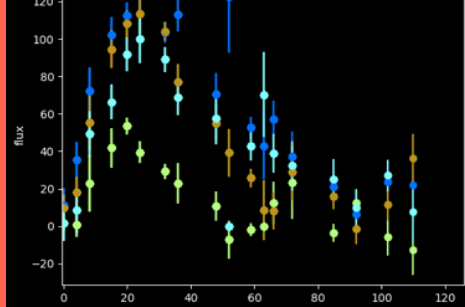


$$\mathcal{P}(\hat{\mathbf{y}} | \mathbf{x}) = \int \mathcal{P}(\hat{\mathbf{y}} | \mathbf{x}, \mathbf{w}) \mathcal{P}(\mathbf{w} | \mathcal{D}) d\mathbf{w}$$

$$\mathcal{P}(\mathbf{w} | \mathcal{D}) \approx q(\mathbf{w} | \theta) \quad \text{variational distribution}$$

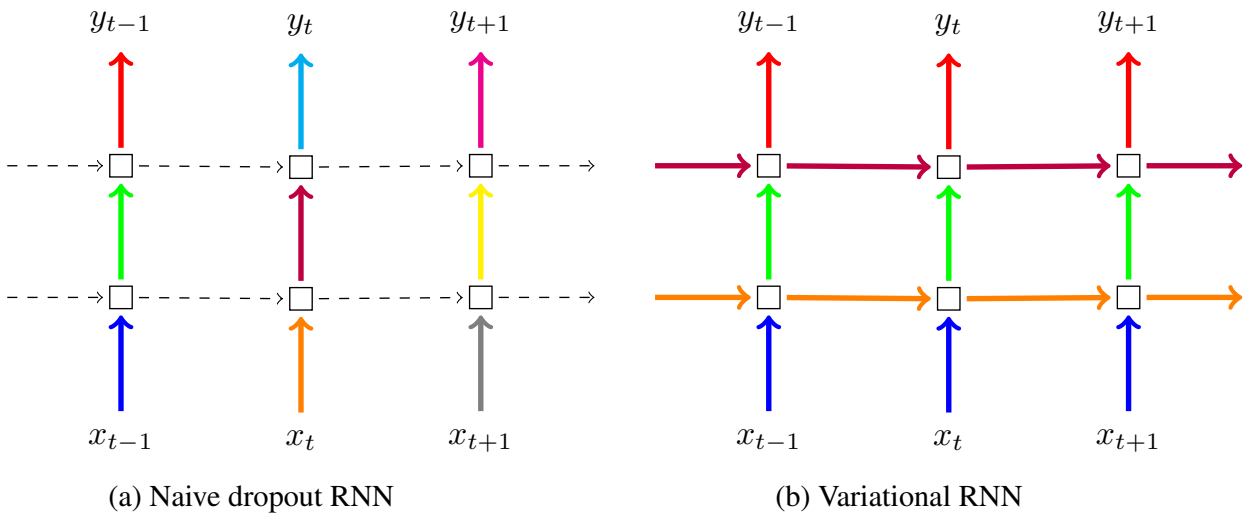
Training minimisation

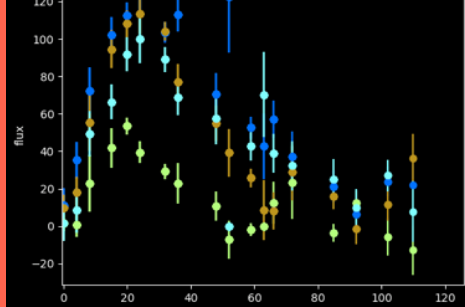
$$\hat{\theta} = \min_{\theta} \mathbf{KL} (q(\mathbf{w} | \theta) || \mathcal{P}(\mathbf{w} | \mathcal{D}))$$



Approximating the variational distribution $q(\mathbf{w}|\theta)$

1. MC dropout *Gal & Ghahramani 2016*

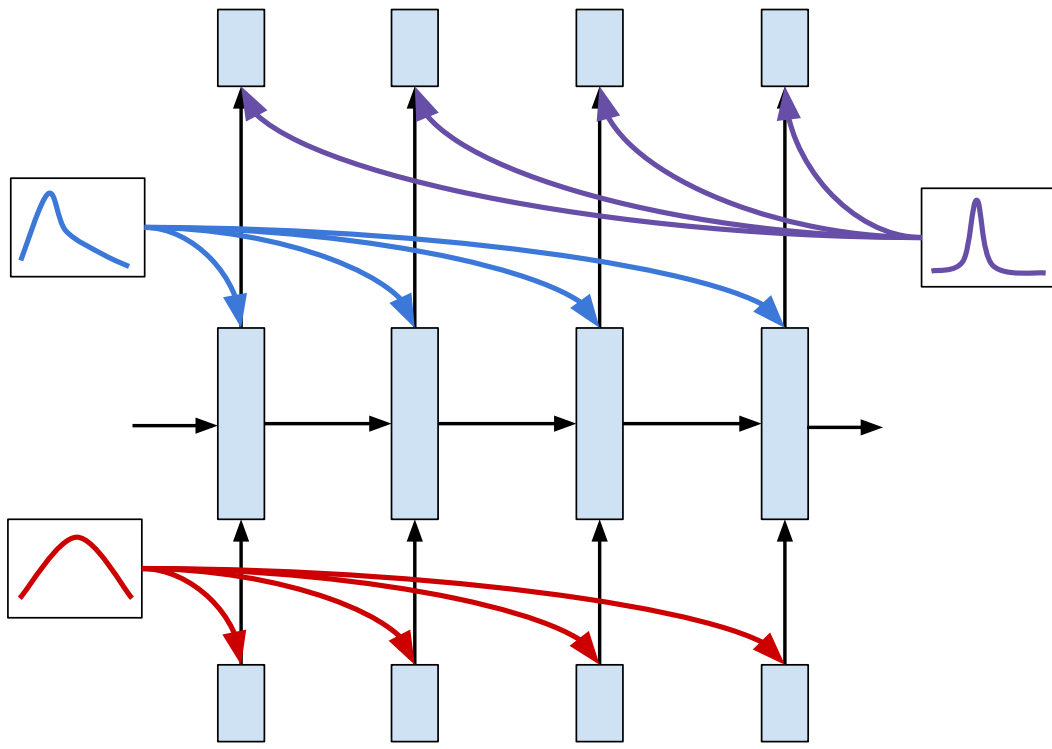
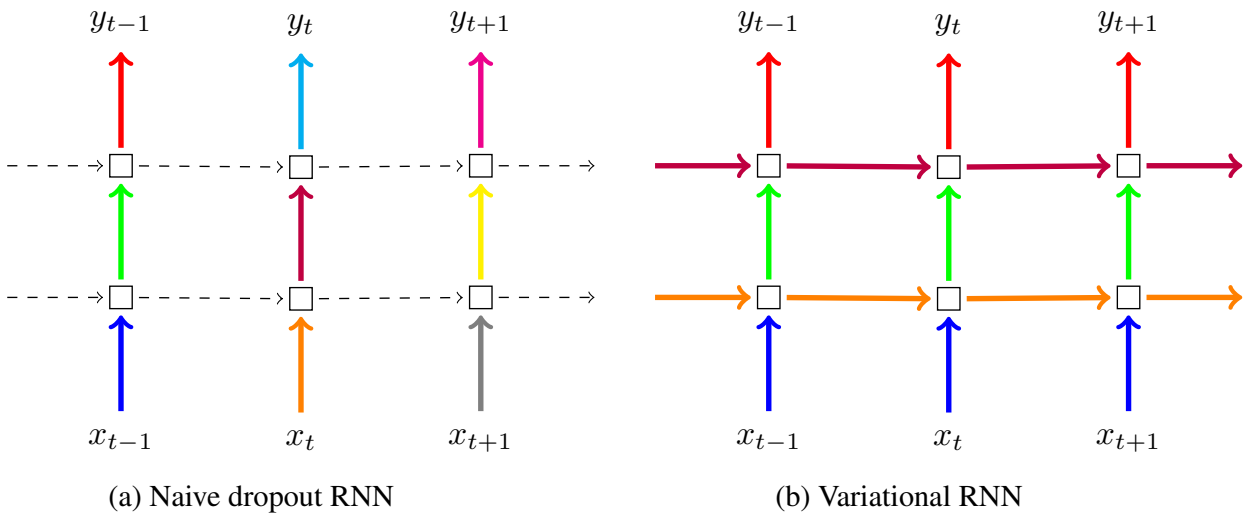


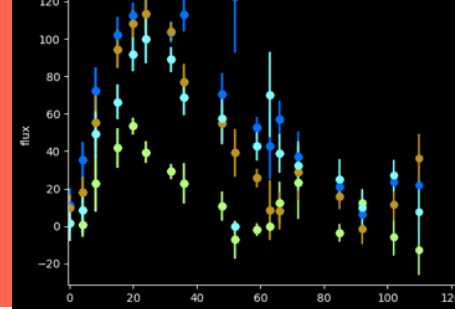


Approximating the variational distribution $q(\mathbf{w}|\theta)$

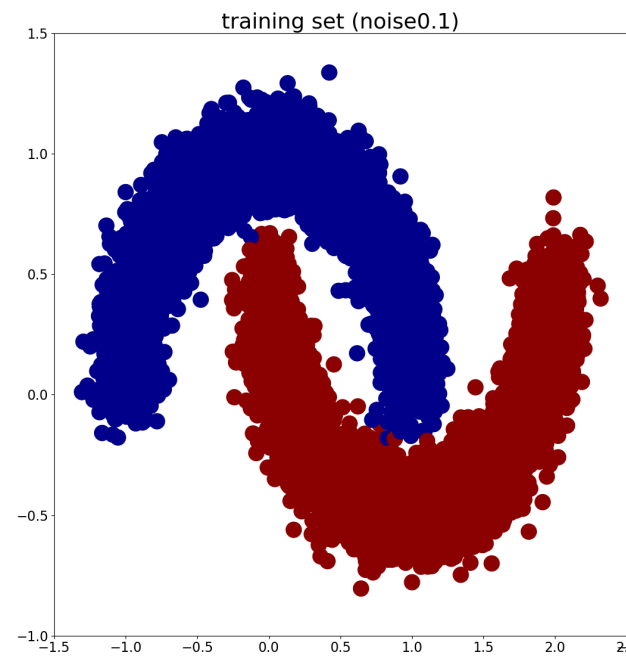
1. **MC dropout**
Gal & Ghahramani 2016

2. **Bayes by Backprop**
Fortunato+ 2017

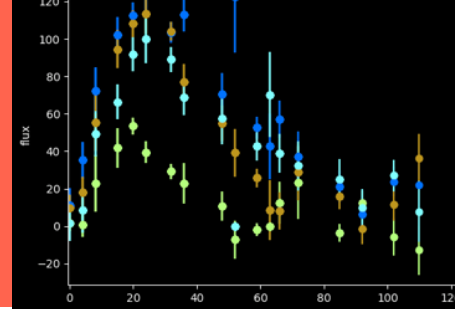




training

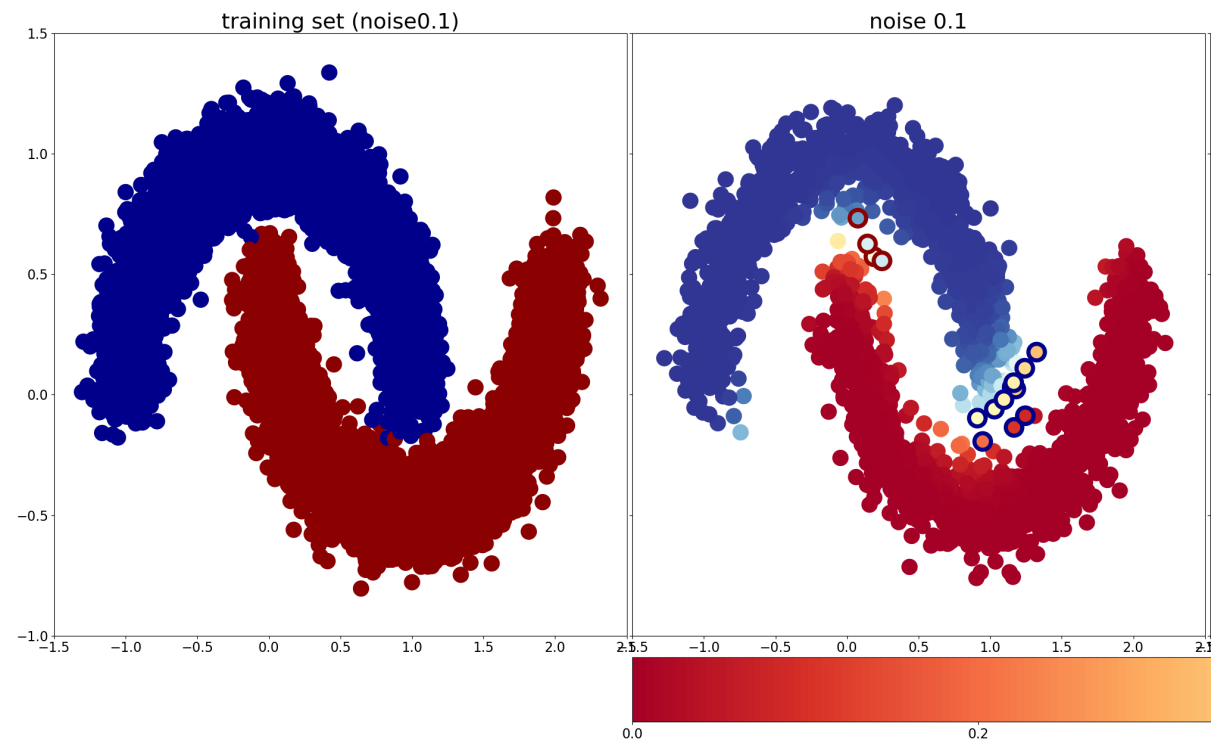


classification probability

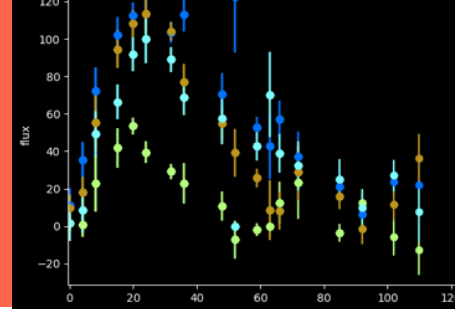


training

classification probability



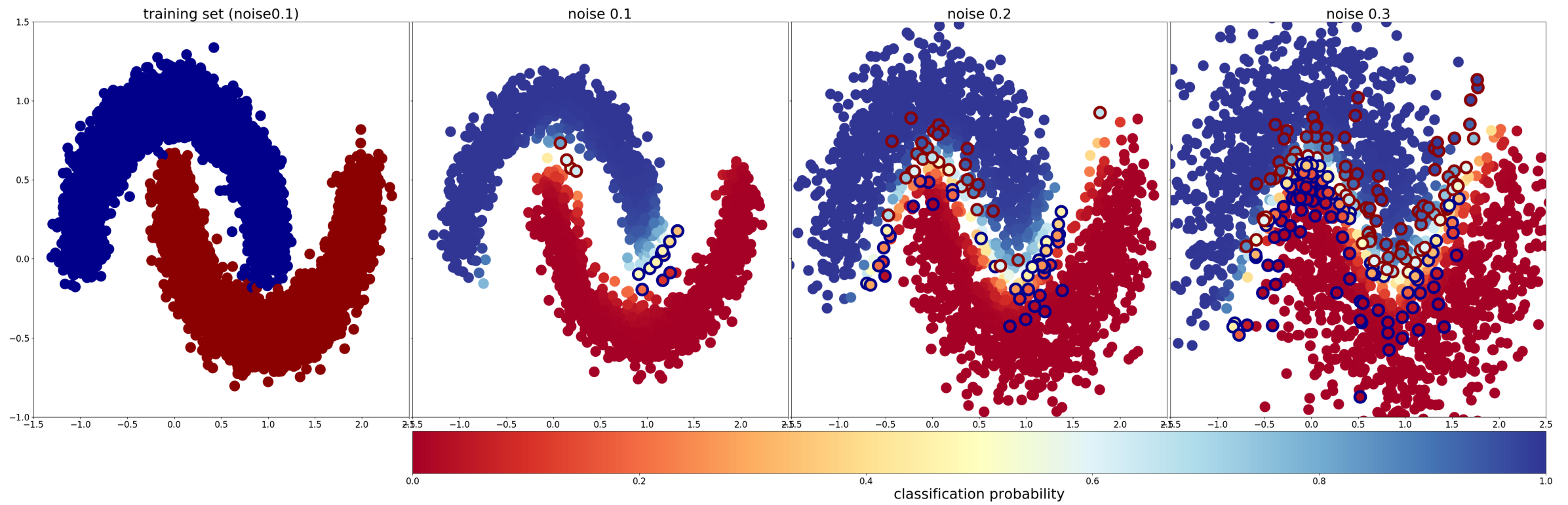
Part II: typing with photometry



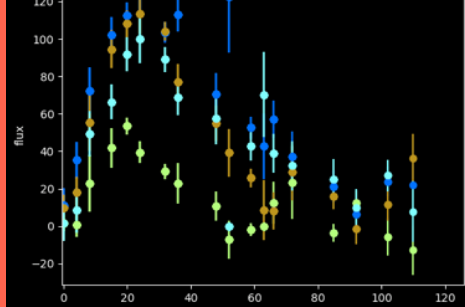
Bayesian NNs

training

classification probability

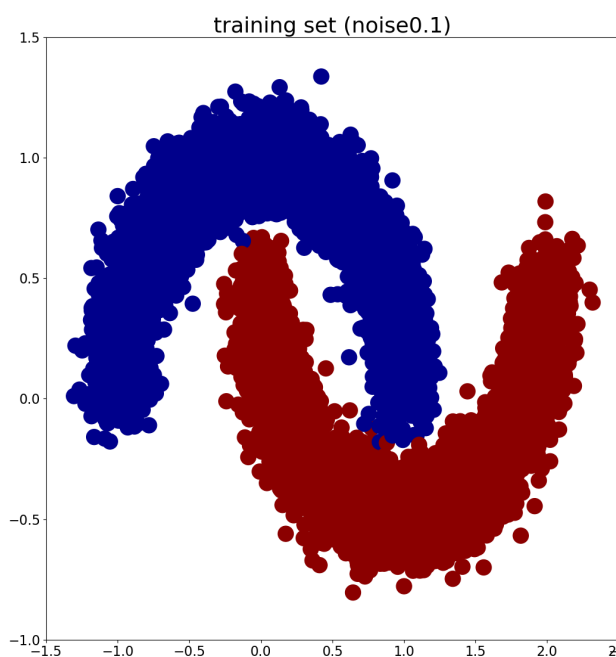


Part II: typing with photometry

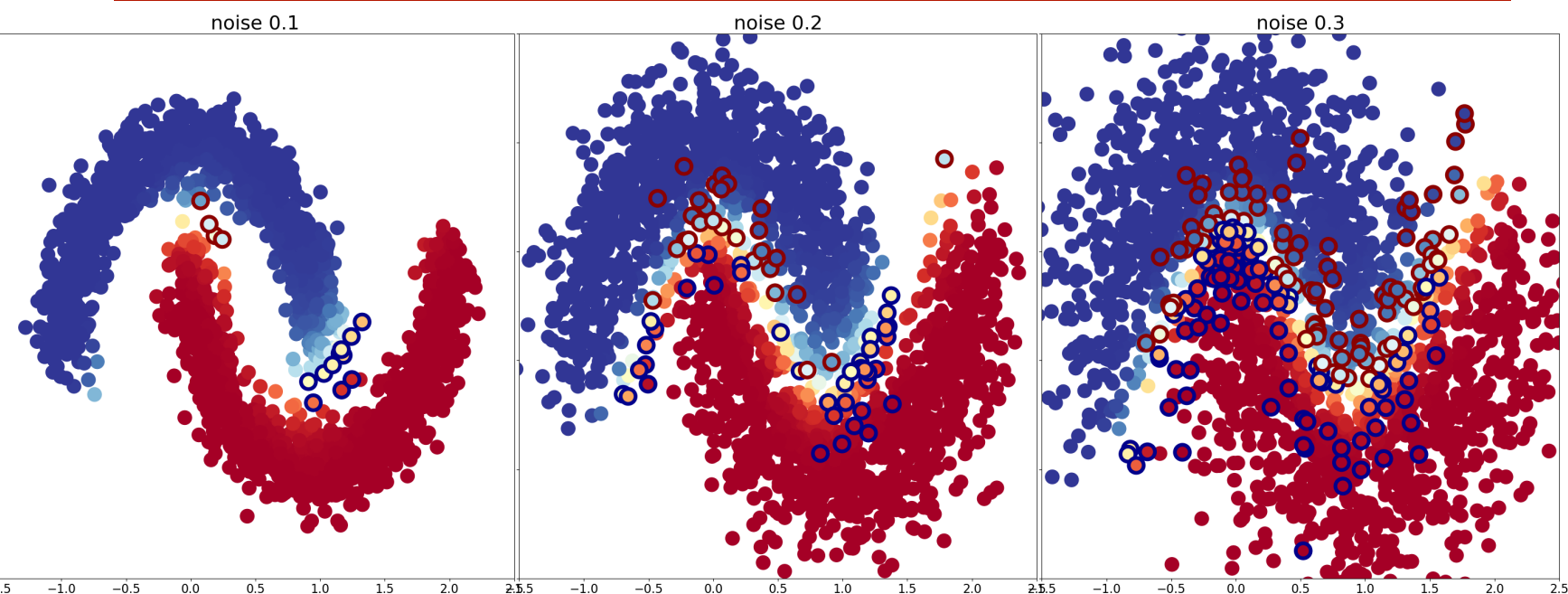


Bayesian NNs

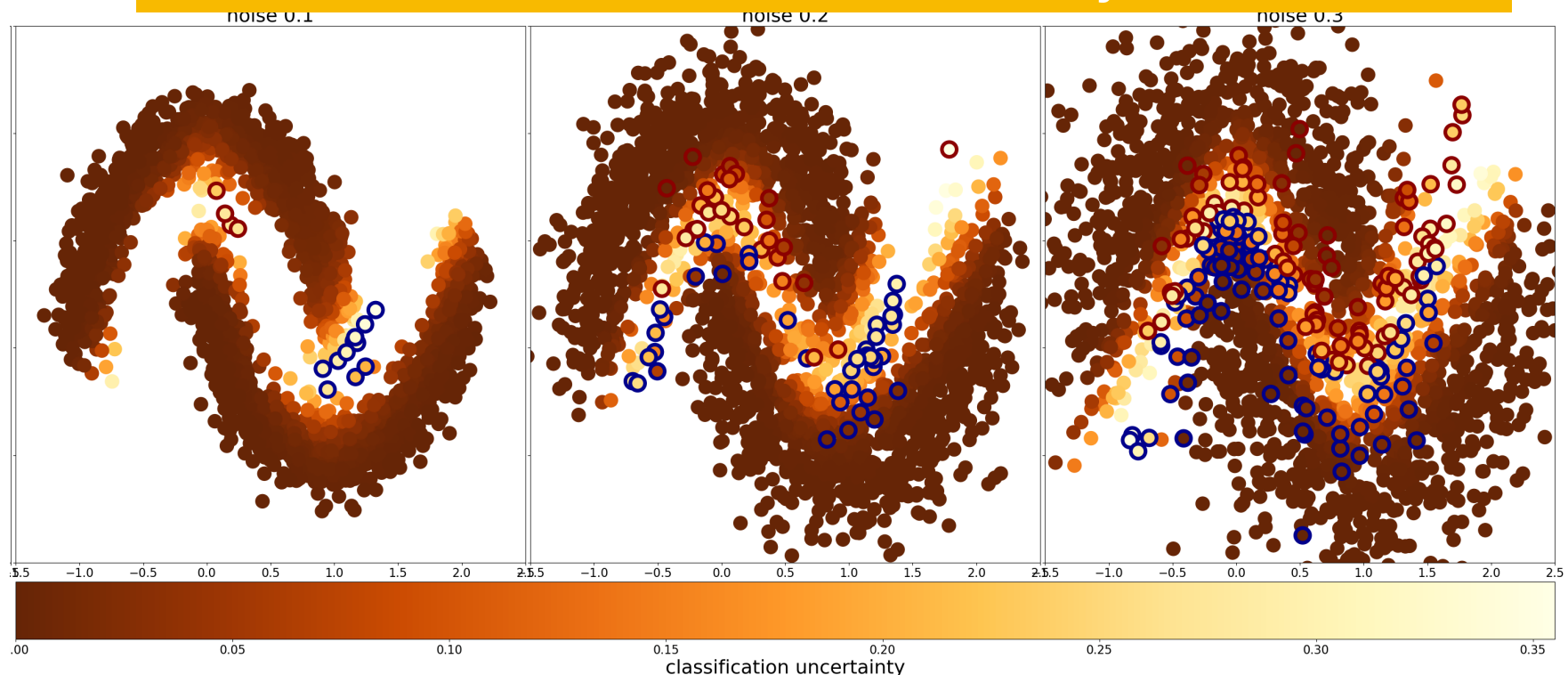
training

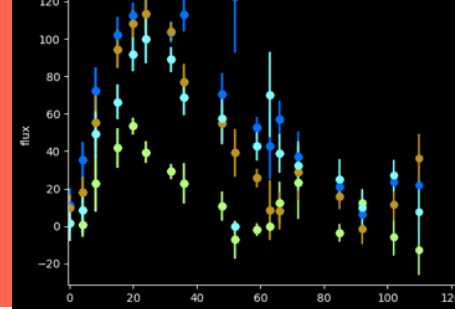


classification probability



classification uncertainty

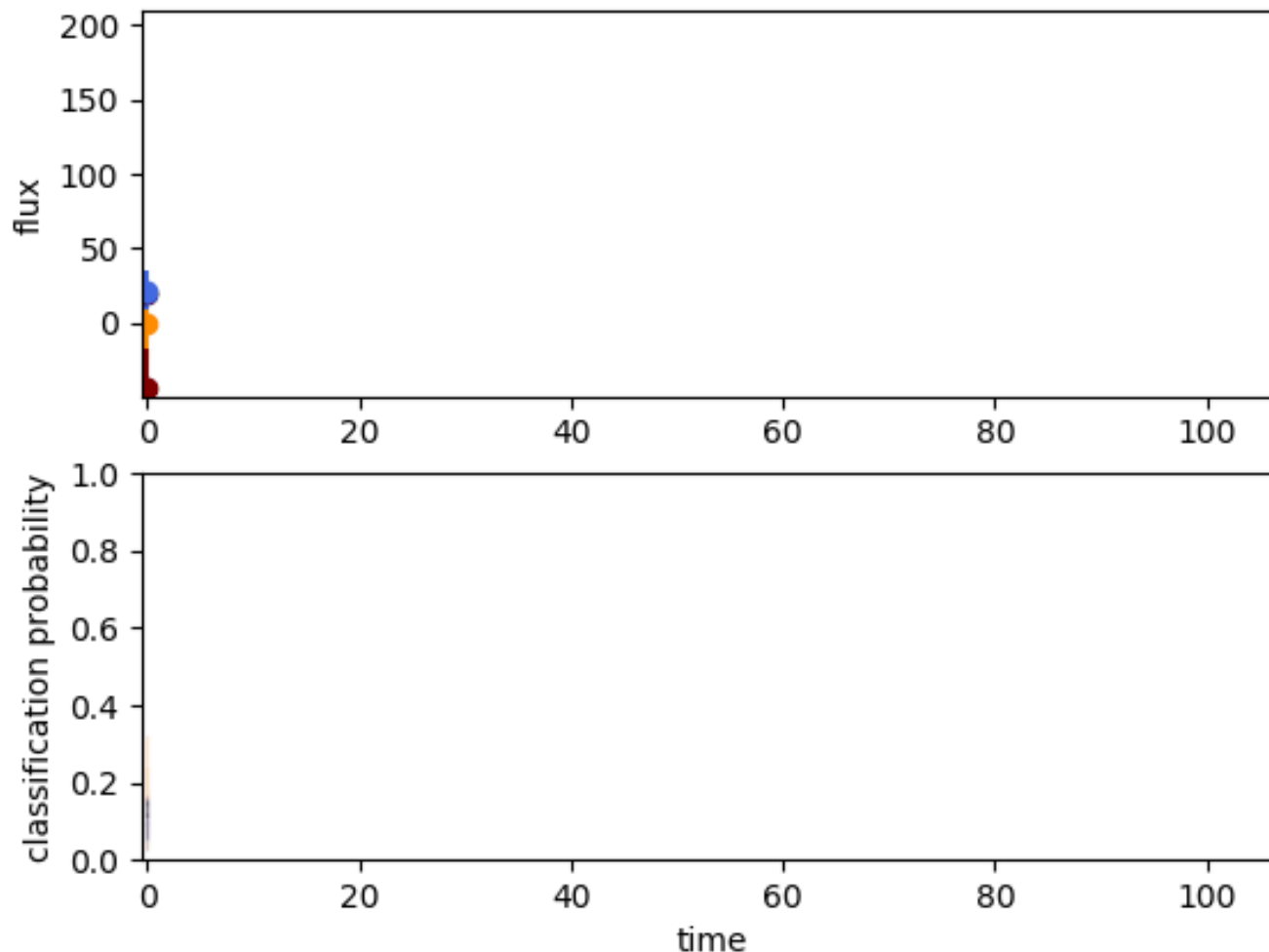




1. MC dropout

Gal & Ghahramani 2016

Ia (ID: 31577798, redshift: 0.374)



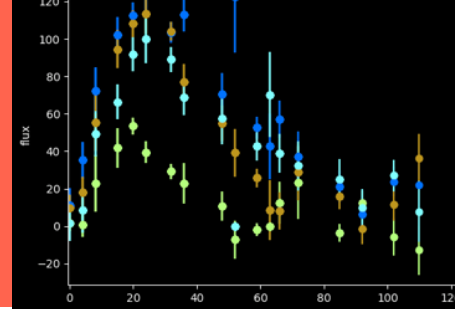
2. Bayes by Backprop

Fortunato+ 2017

Posterior that provides epistemic uncertainties

Epistemic uncertainties:

express our ignorance about the model that generated the data.



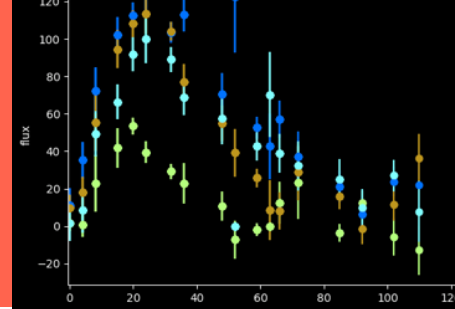
Training sets are:

! not representative

! incomplete (we don't know/can't simulate)

? *Can we use output from ML classifiers for cosmology or any statistical analyses?*

Part II: typing with photometry



ML limitations
representativity

simulation



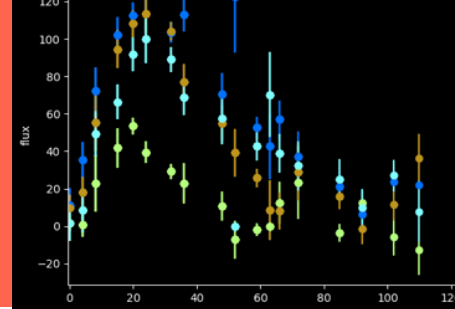
peak brightness i

data



peak brightness i

Distribution of properties of SNe



ML limitations representativity

simulation



peak brightness i



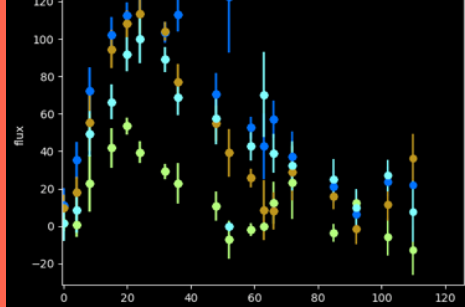
train classification algorithm

data



peak brightness i

Part II: typing with photometry



ML limitations representativity

simulation



peak brightness i



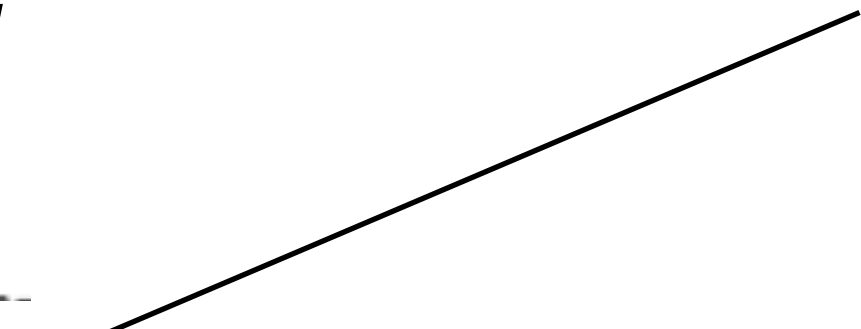
train classification algorithm

data

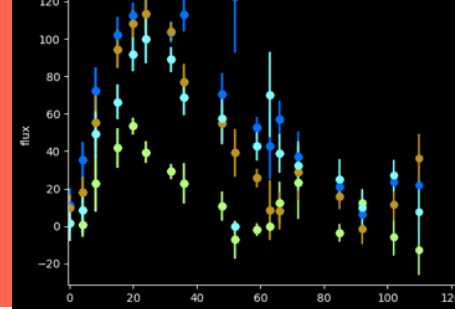


peak brightness i

classify the data



accuracy decreases (*Lochner+ 2015, Charnock+2017*)



ML limitations representativity

simulation



peak brightness i

train classification algorithm

data



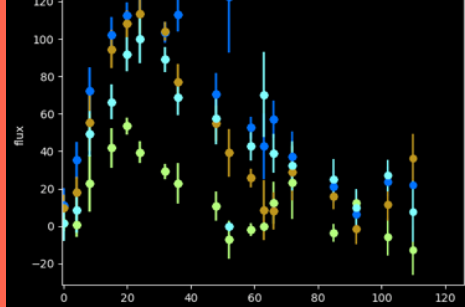
peak brightness i

classify the data

Either we improve our training sets or search for robust methods!

Pasquet+ 2019, Möller+ 2019

Part II: typing with photometry



ML limitations representativity

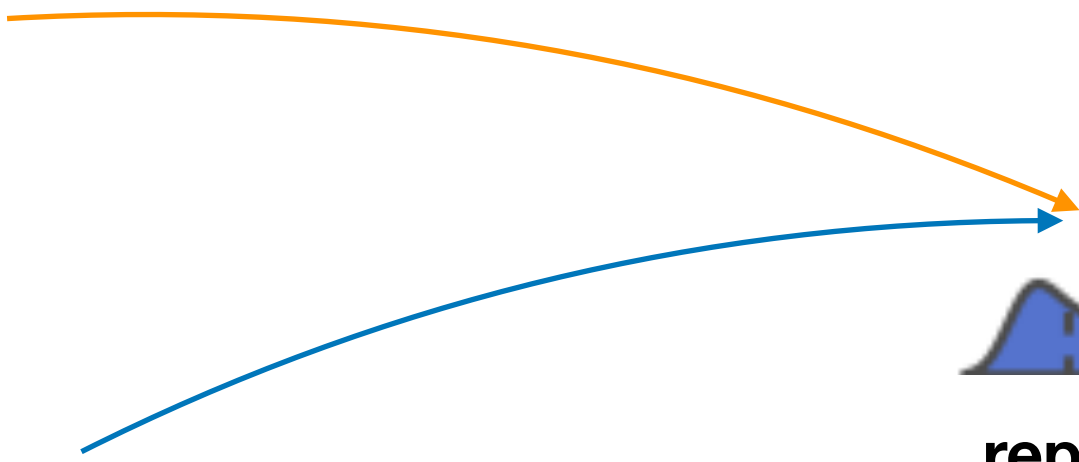
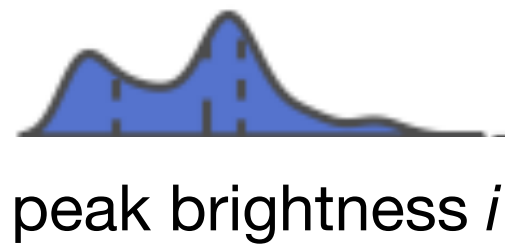


SuperNNova *BNNs*
open source photometric
classification

**Simplistic
simulation**



**representative
simulation**

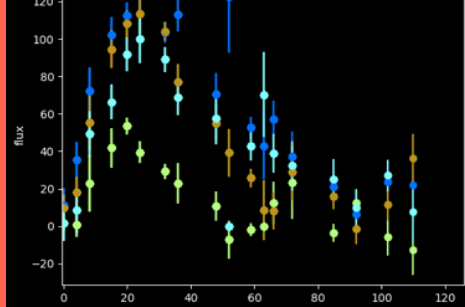


classify



**representative
simulation**

Part II: typing with photometry



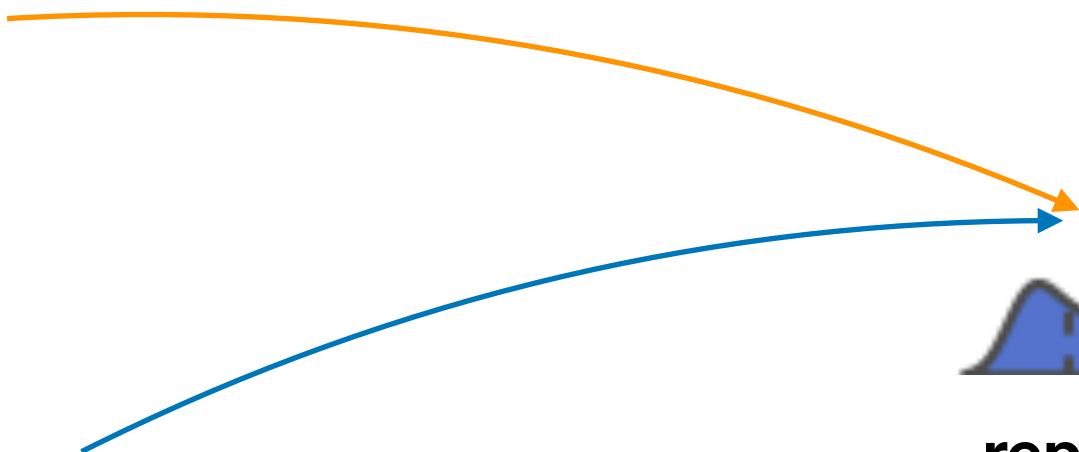
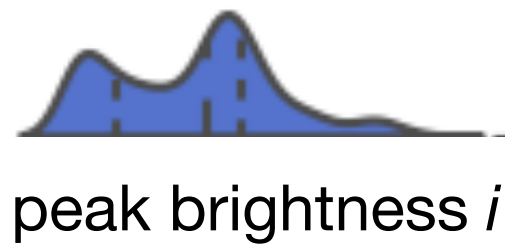
ML limitations representativity

SuperNNova *BNNs*
open source photometric classification

Simplistic simulation



representative simulation



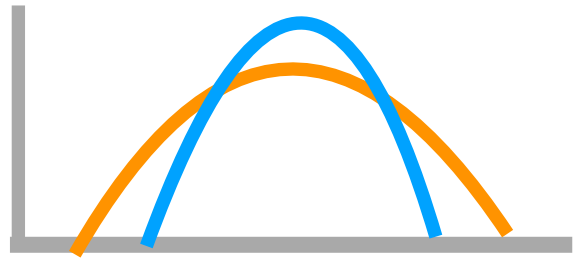
classify



representative simulation

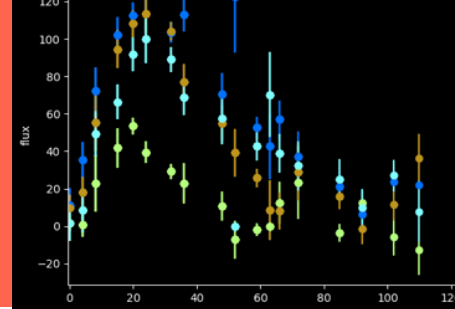
accuracy changes slightly ($\langle \text{prob} \rangle$ are not the most indicative)

non-representative models give larger uncertainties!



Probability

Part II: typing with photometry



ML limitations incompleteness

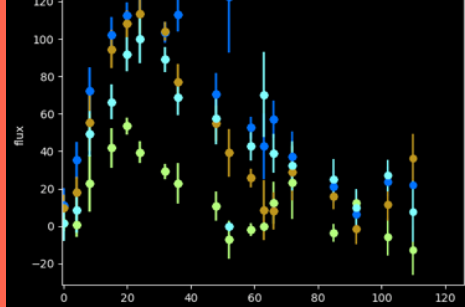


training set



to classify

Part II: typing with photometry



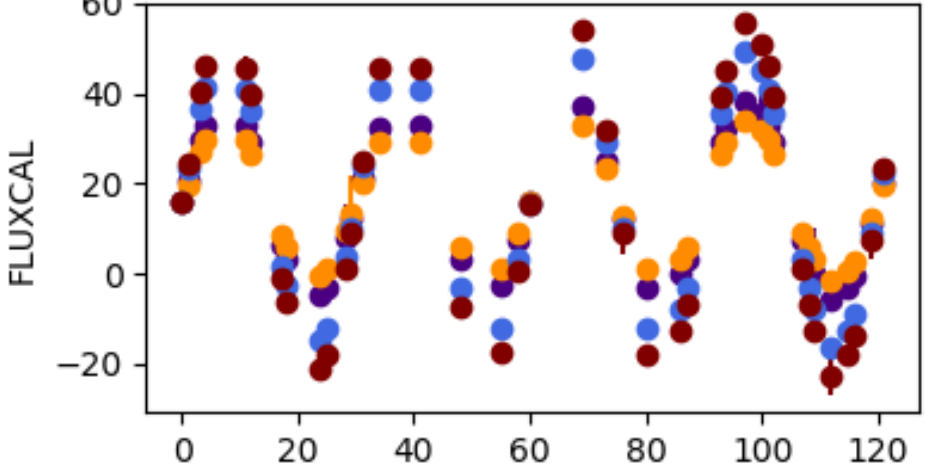
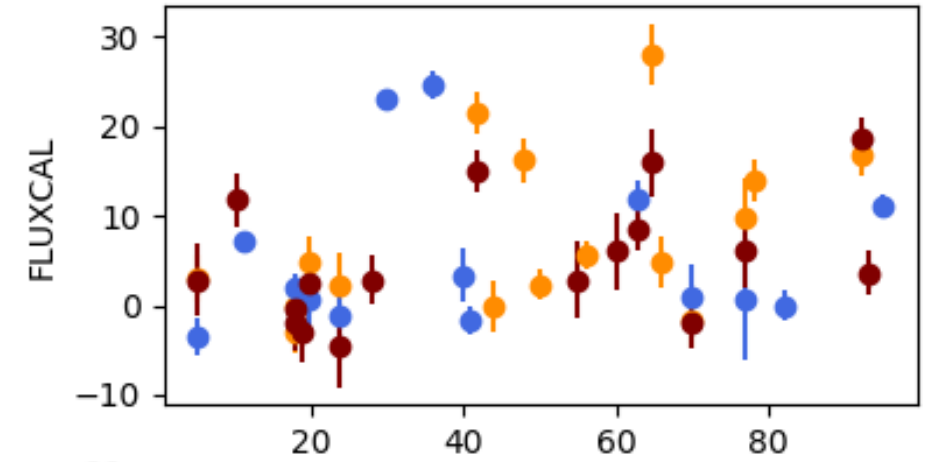
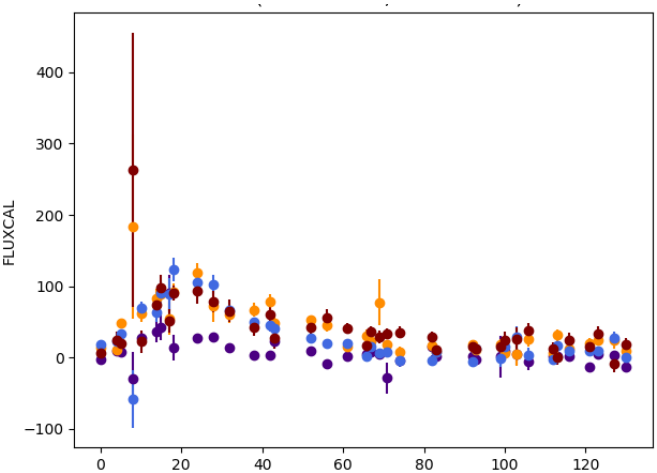
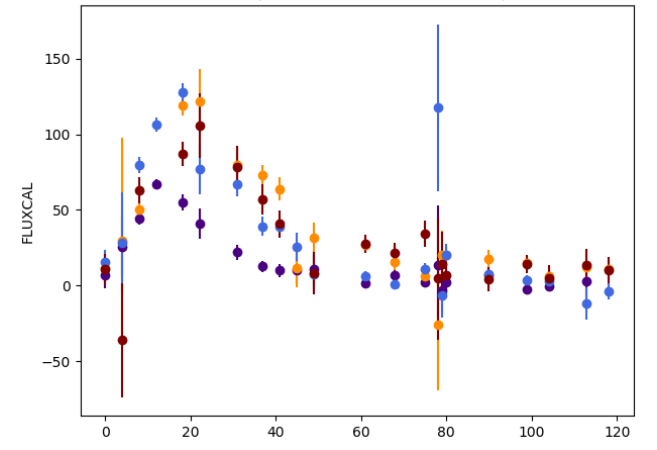
ML limitations incompleteness



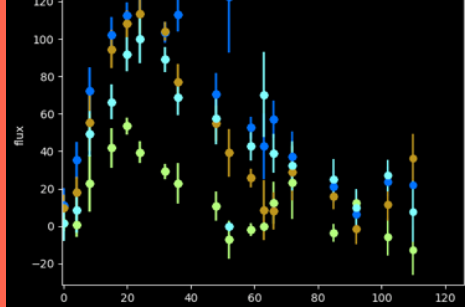
training set



to classify



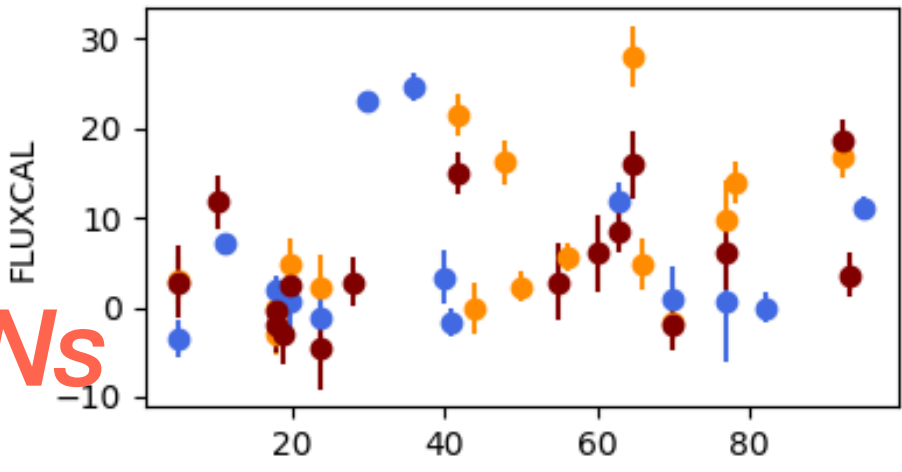
Part II: typing with photometry



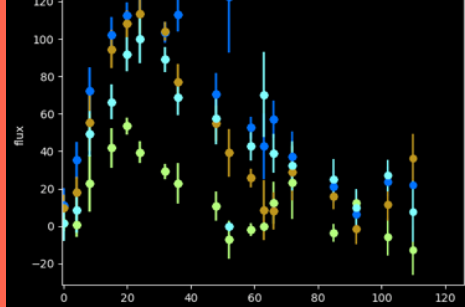
ML limitations incompleteness

SuperNNova
open source photometric
classification

with BNNs

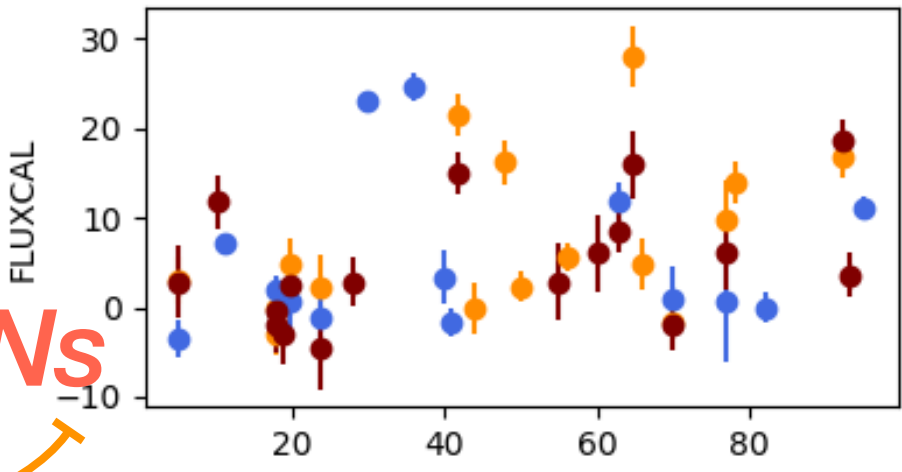


Part II: typing with photometry

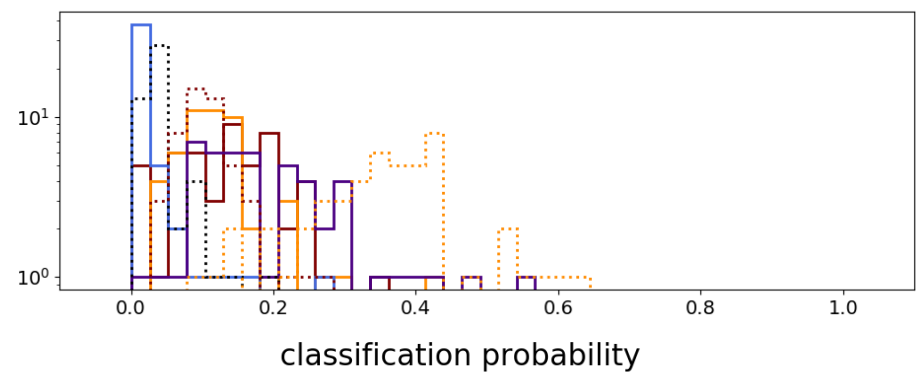


ML limitations incompleteness

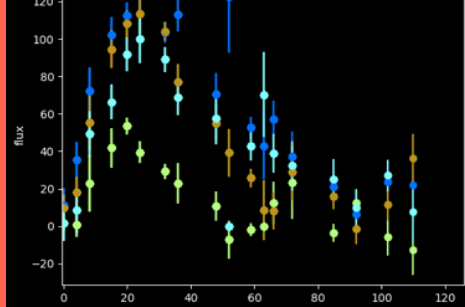
SuperNNova *with BNNs*
open source photometric
classification



low probability for any class

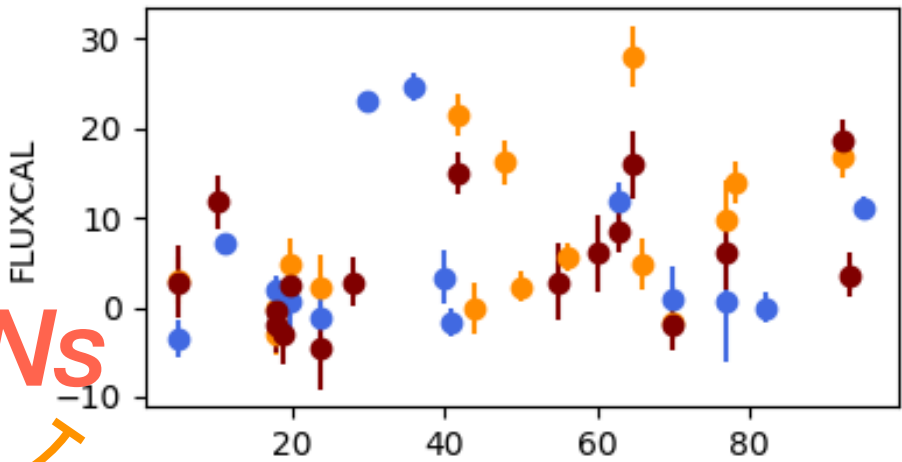


Part II: typing with photometry

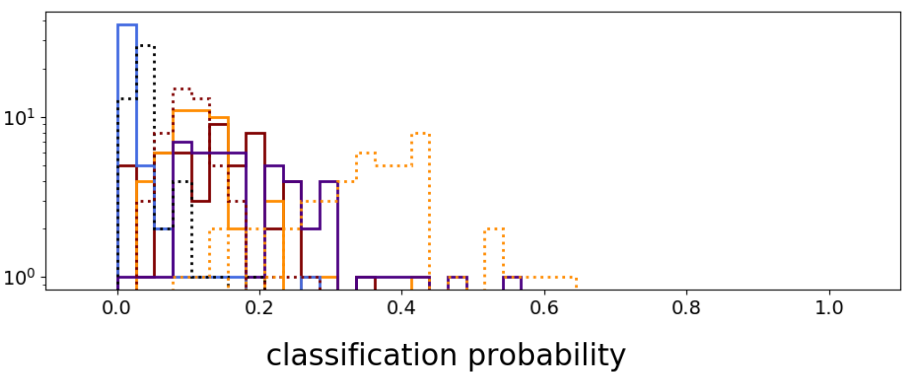


ML limitations incompleteness

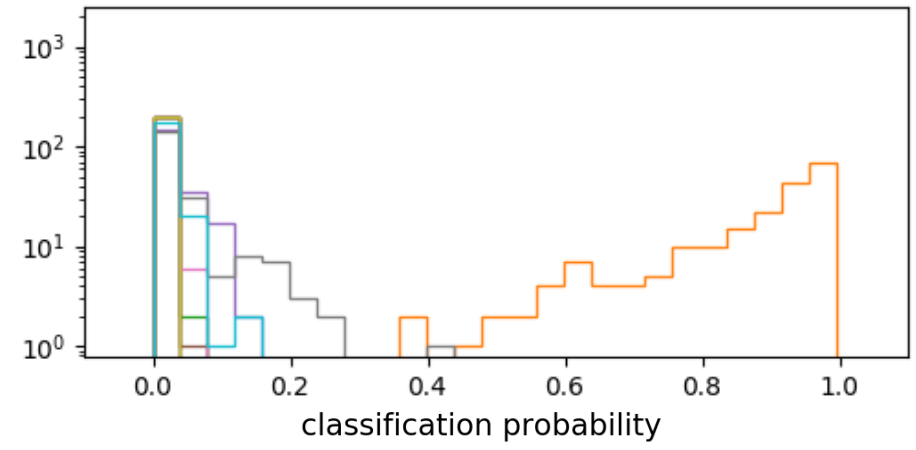
SuperNNova *with BNNs*
open source photometric classification



low probability for any class

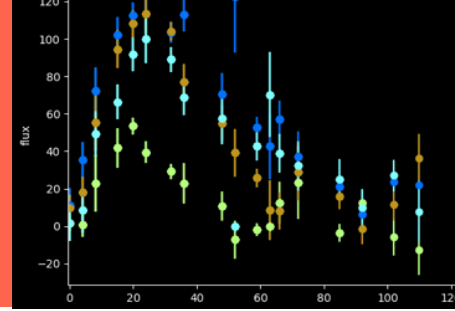


high probability for "less-known" class



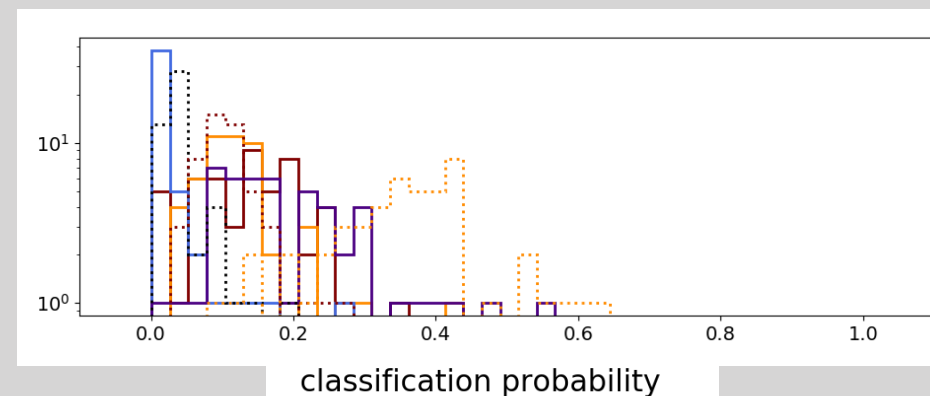
but... BNNs can give us high-probability but large uncertainty

"an increase in average classification uncertainties for these anomalies" Möller + 2019

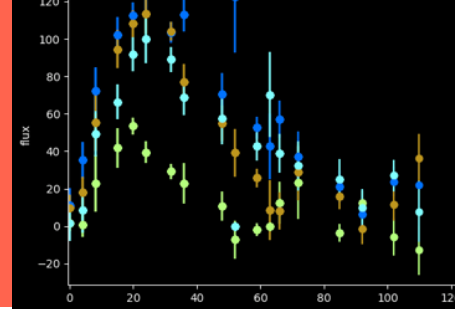


ML probabilities for statistical analyses?

Selecting a SN Ia sample:
cutting on “classification probabilities” for selection

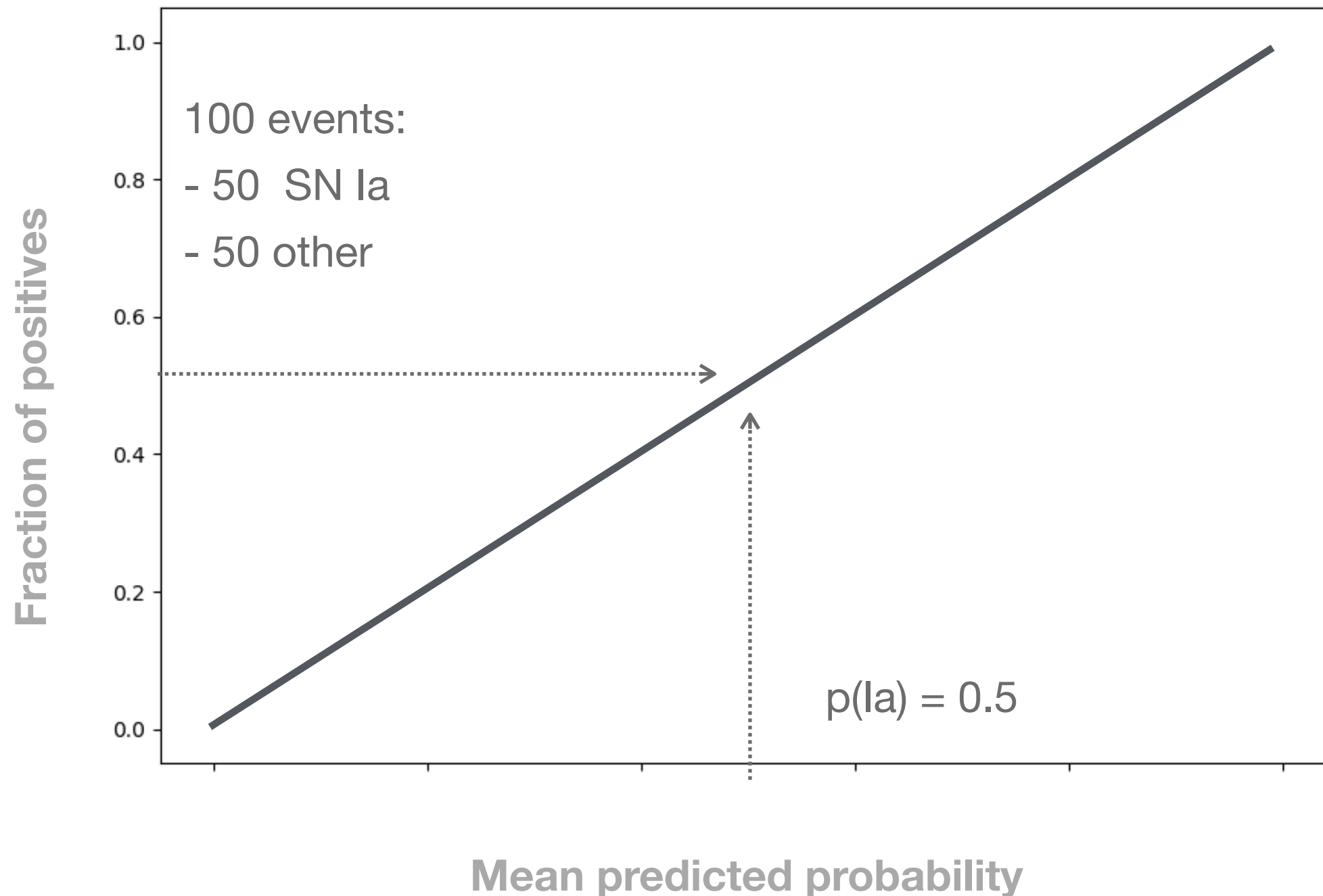


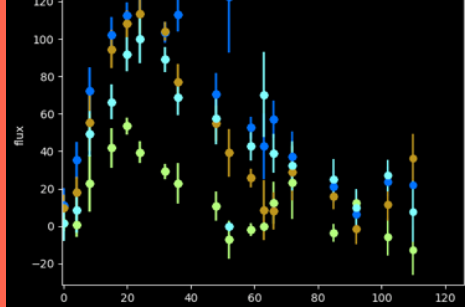
Can use a “weight” in the analysis using these “classification probabilities” *Jones+2018, Hinton+2018*



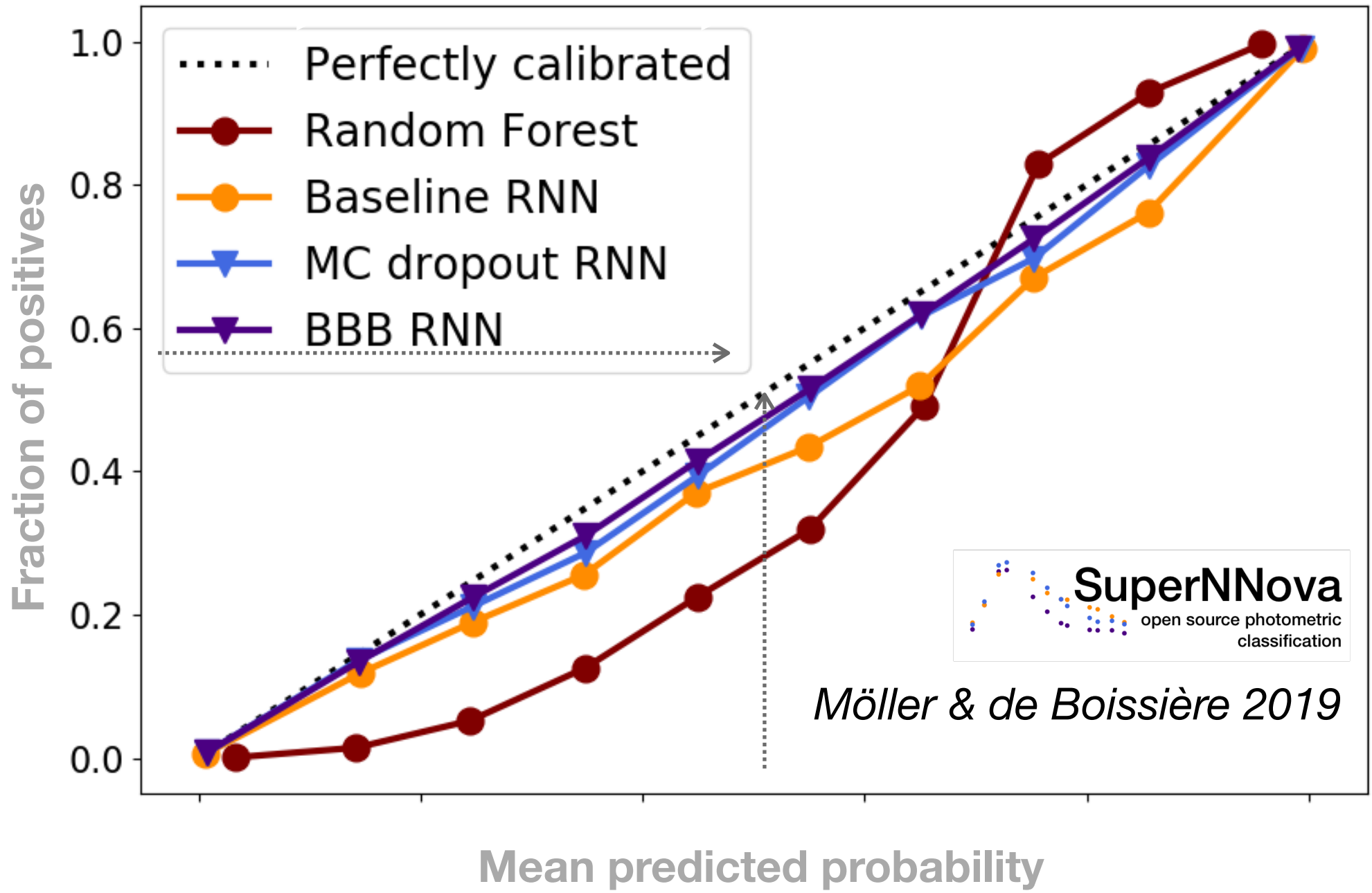
ML probabilities for statistical analyses?

De Groot+ 1983, Niculezcu-Mizil+ 2005, Guo+ 2017

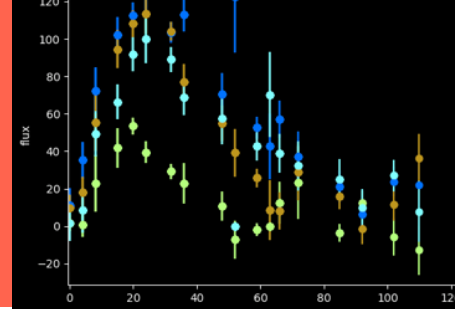




ML probabilities for statistical analyses?



Part II: typing with photometry

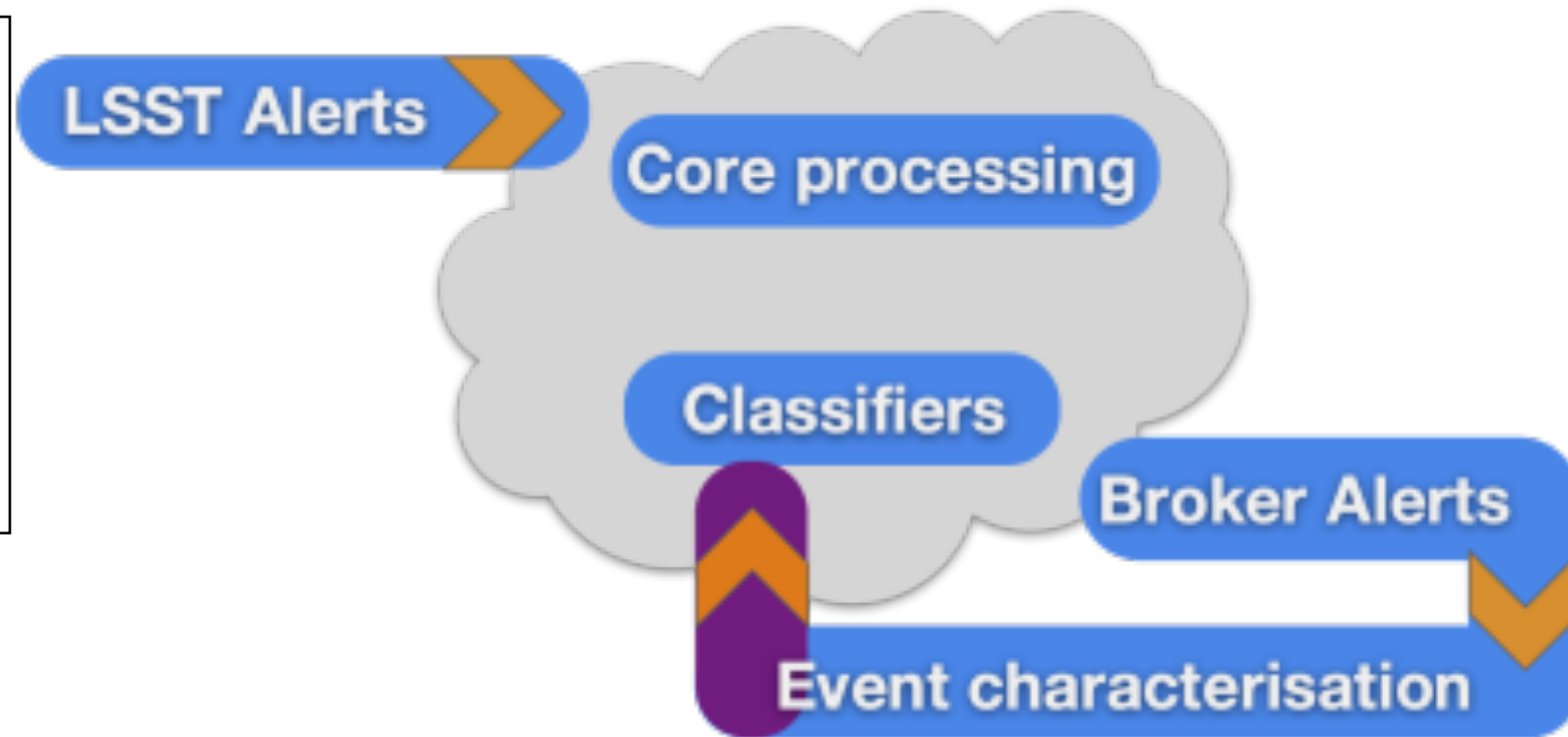


FINK

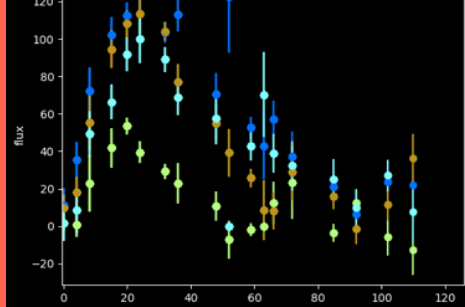
find, classify, improve

- 10 IN2P3/CNRS labs and centers
- Big data (15 TB/night)
- **ML on data stream**
- Automated decision making
- Minimum duration of 10 years
- Connection across experiments

J. Peloton, E. Ishida, A. Möller
+28 signatories



Part II: typing with photometry

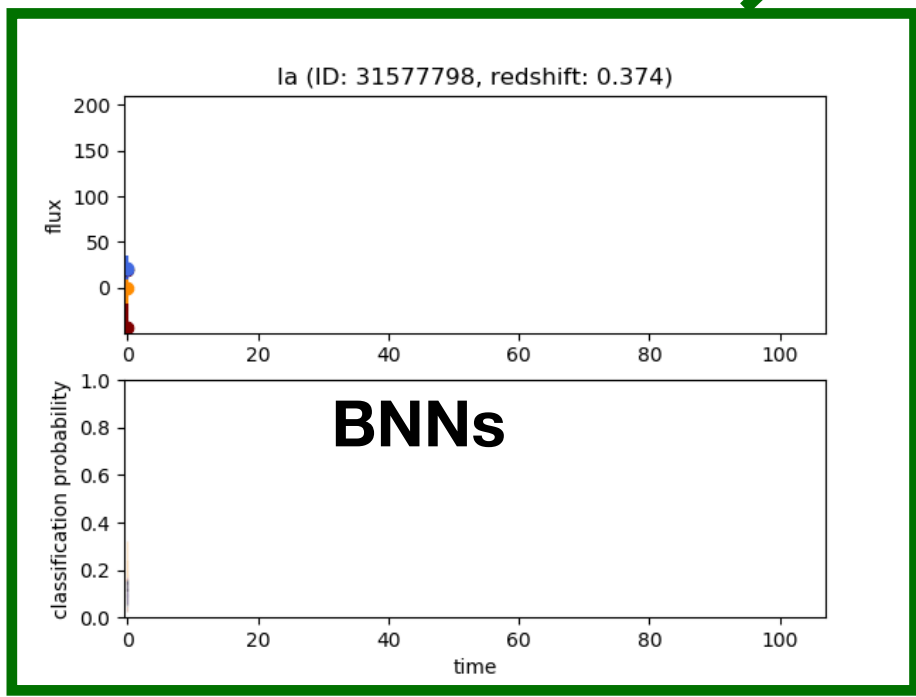
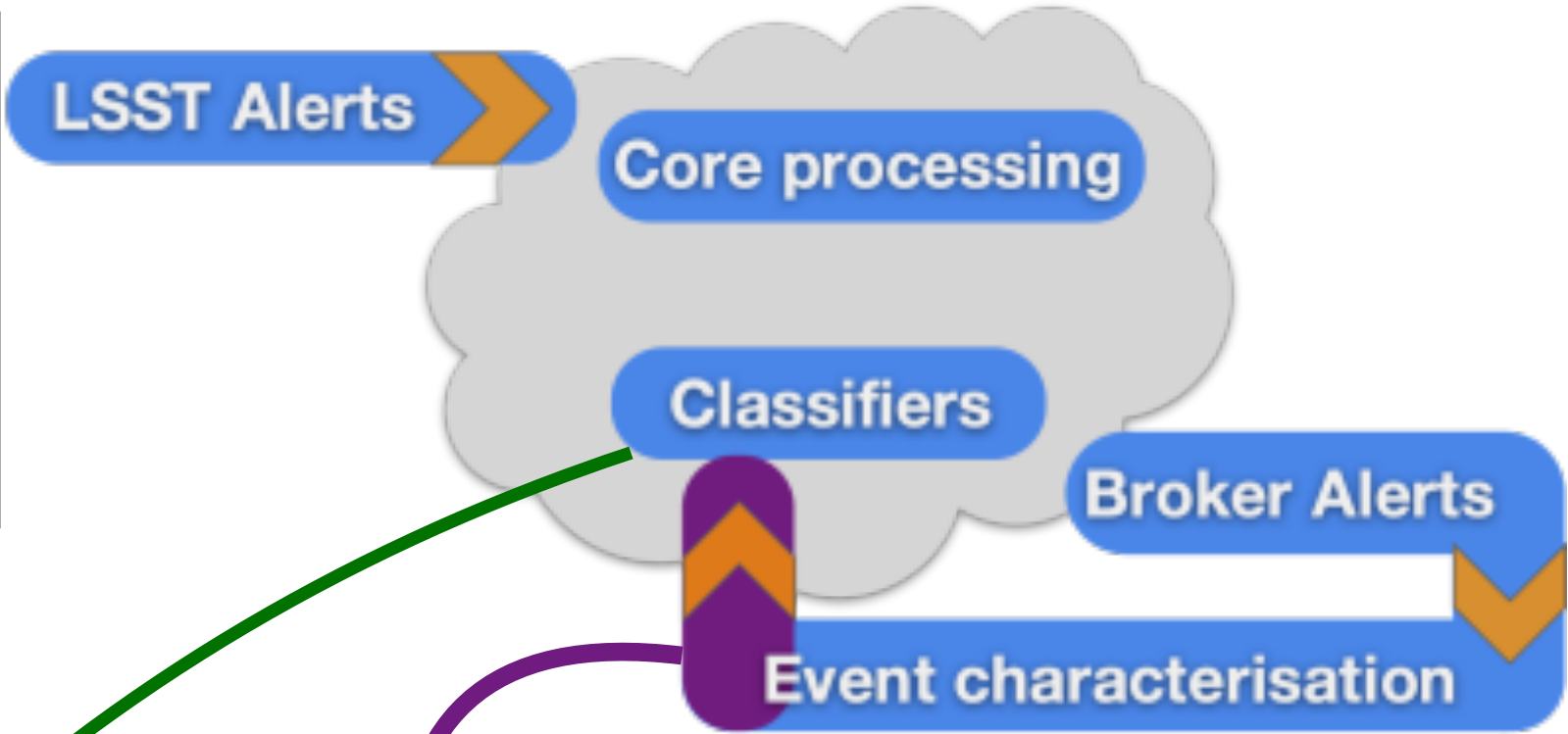


FINK

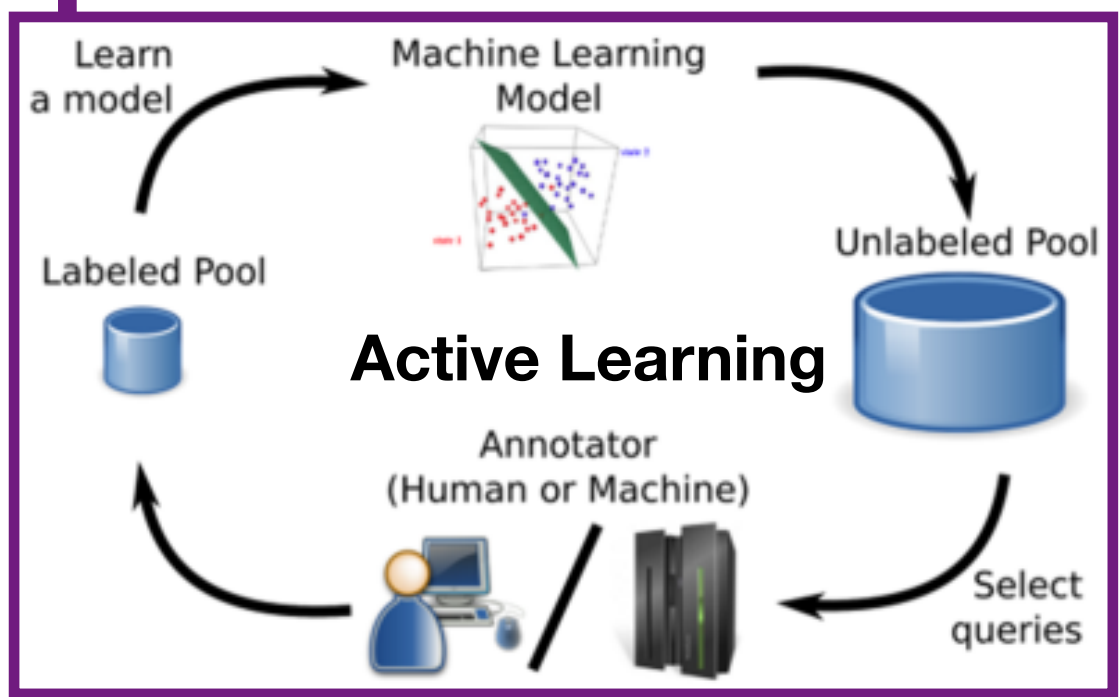
find, classify, improve

- 10 IN2P3/CNRS labs and centers
- Big data (15 TB/night)
- **ML on data stream**
- Automated decision making
- Minimum duration of 10 years
- Connection across experiments

J. Peloton, E. Ishida, A. Möller
+28 signatories



Möller + 2019



Ishida + 2018

take away

- The big *astronomical* data era requires the use of ML methods
- Machine learning is key for supernova cosmology.
- The last decade we have seen large advancements on problems like real vs. bogus & photometric classification.
- To exploit our large SN samples we need to start evaluating the robustness as one of the key components of our classifiers.
- Bayesian Neural Networks are promising for statistical analyses.

github: [supernnova/SupernNova](https://github.com/supernnova/SupernNova)

