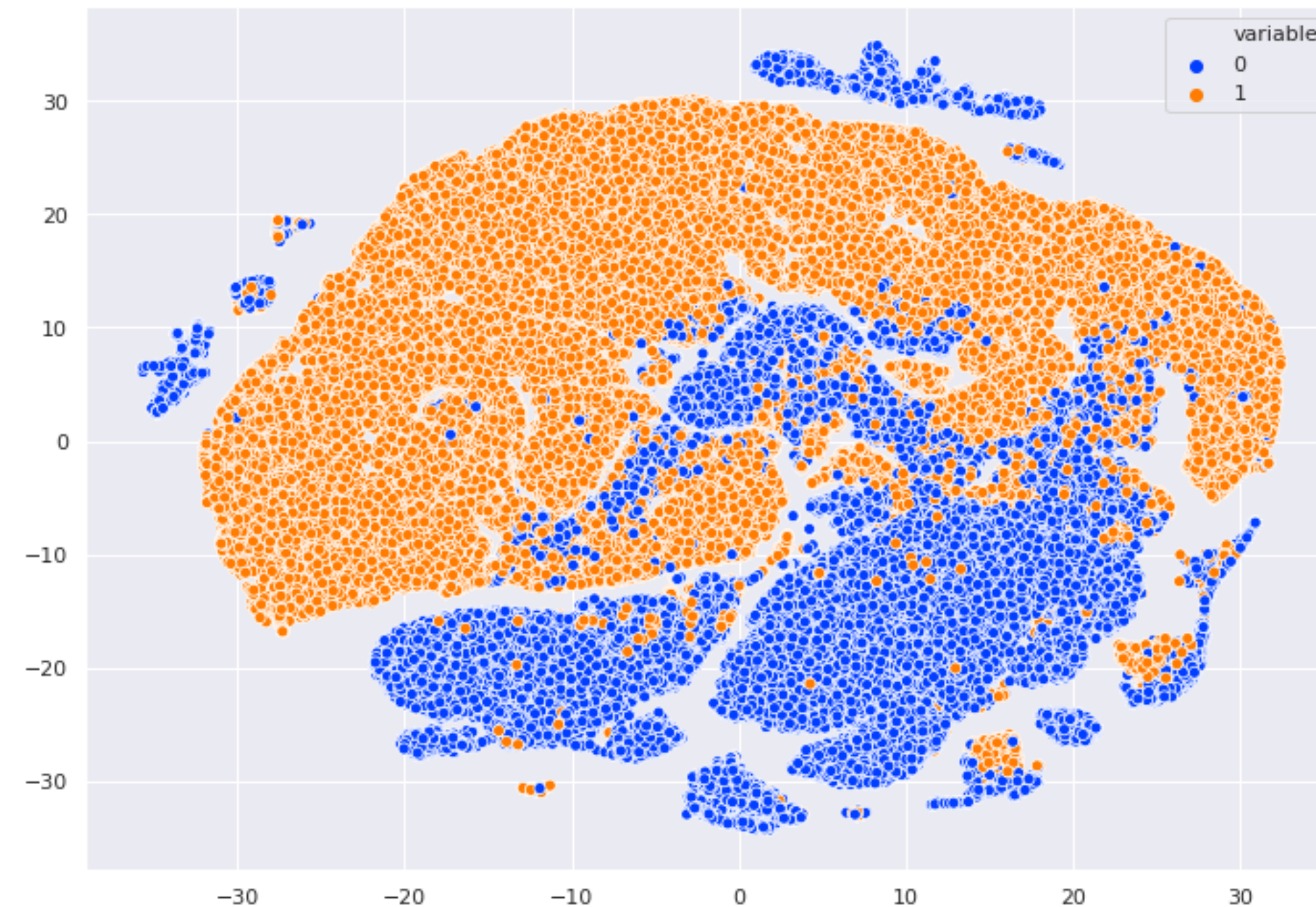


ZTF ML - Status and Plans



Ashish Mahabal

ZTF ML Lead

Caltech, 2020-10-19

On behalf of the ML group and ZTF

- Current/Recent work

- SNe
- Transients
- TDEs
- NEOs
- Comets
- Variables

- Future/plans

- meta-classification
- more classes
- Outliers
- Focus improvements
- Statistical studies

+ lot of science (e.g. microlensing, YSOs, redbacks/blackwidows, ...)

SNe

~3000 SN Ia, ~2000 non SN Ia

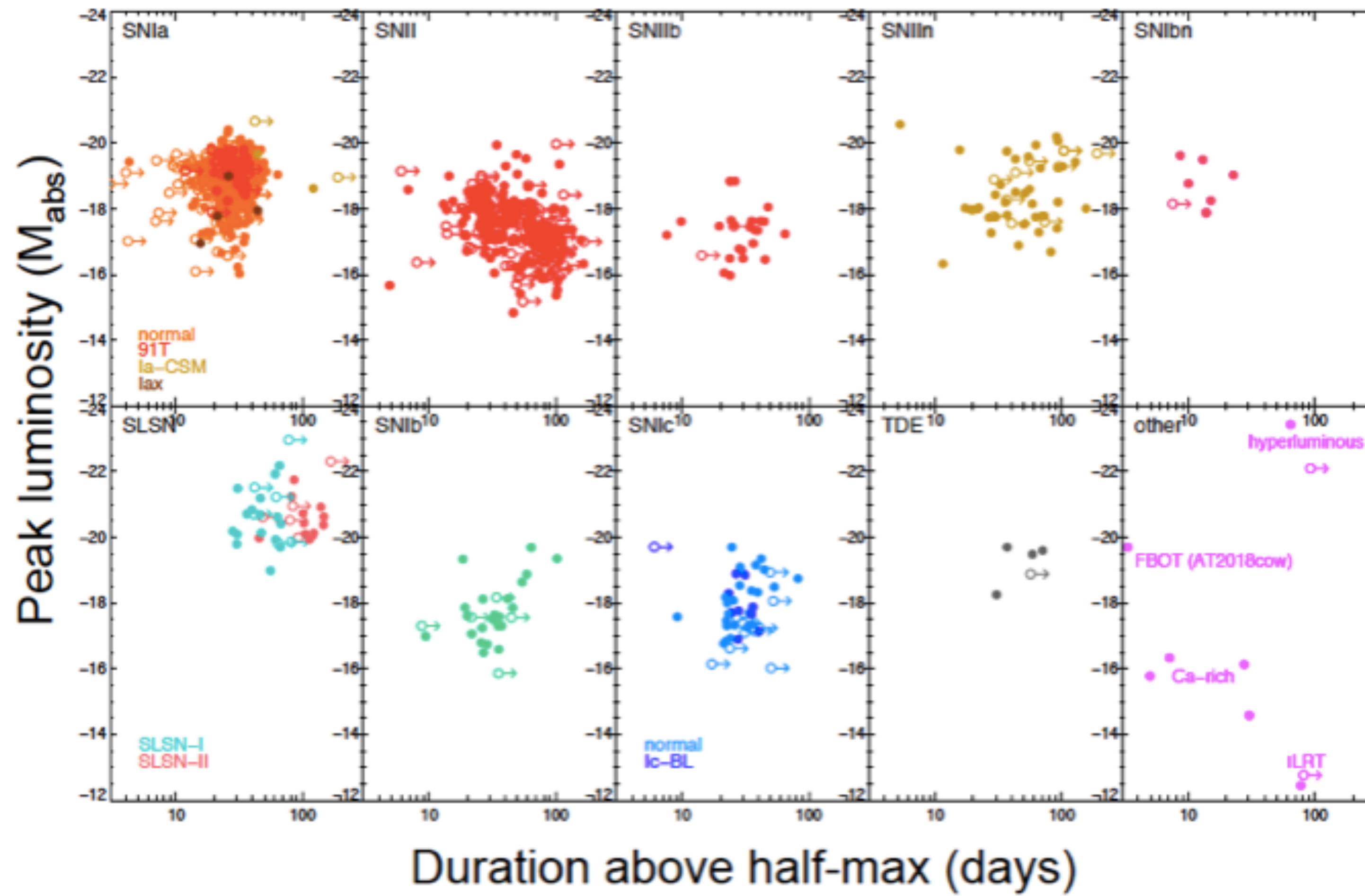
- SN Ia/non-SN Ia classification using SEDM spectra
 - (Christoffer Fremling's talk - Wed 7 AM PT)
- Subclassifying non SN Ia spectra using SEDM spectra (mostly SNe)
 - (Yashvi Sharma, Mahabal, Duev, Fremling, ...)
 - Needs augmentation
 - Multi-class CNN
 - multiple binary classifiers

Transients

- Based on light curve features and metadata
 - Rise-time, fade-time, max-brightness etc.
- XGBoost

with

- Dan Perley
- Christoffer Fremling
- Jakob Nordin
- Nicolas Miranda
- Rahul Biswas
- Anna Ho
- ++



Finding TDEs

with

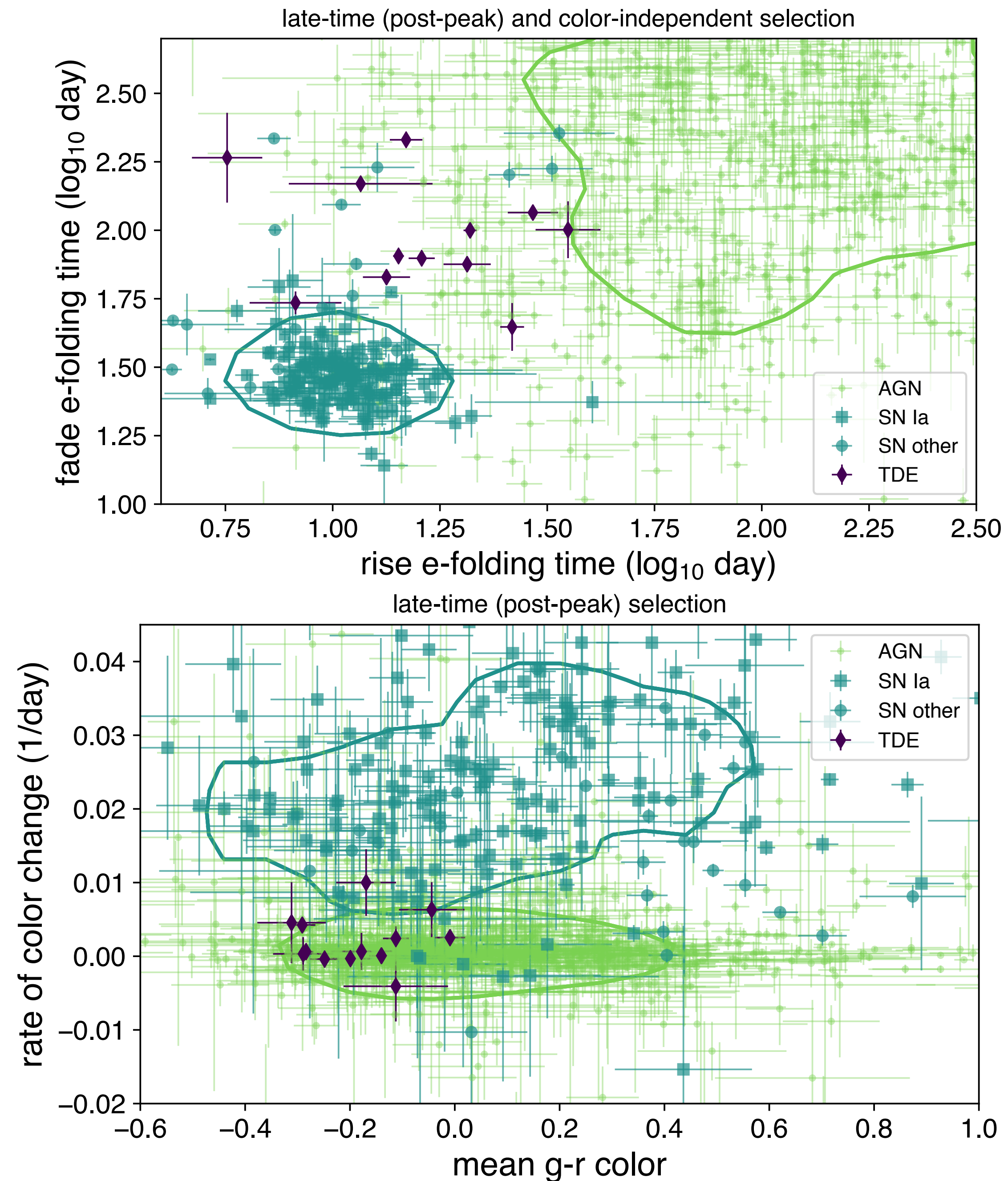
Sjoert Van Velzen

Matthew Graham

Suvi Gezari

++

Van Velzen et al.



NEO Discoveries

Bolin, Helou, Kramer, Masci, Prince

>180 NEOs

ZMODE: Point sources

detection of Main Belt Asteroids, Comets, distant
NEAs

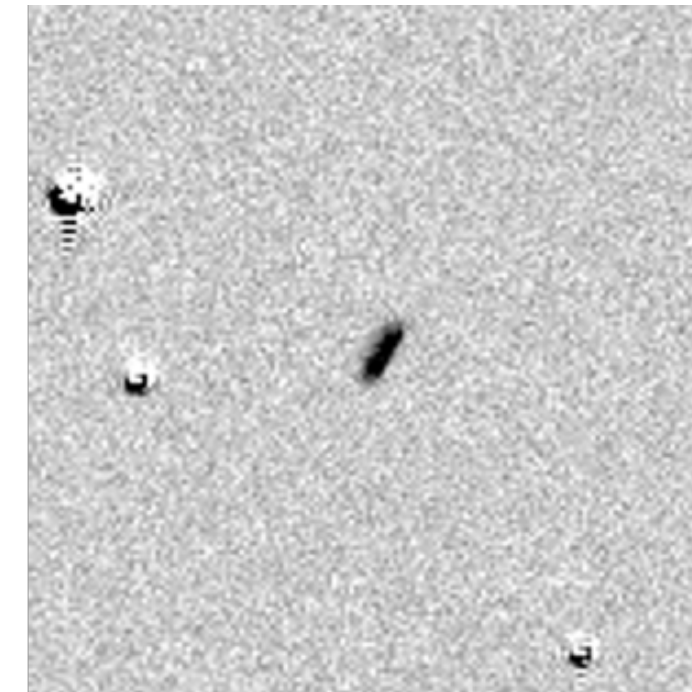
Masci et al 2019

ZSTREAKS: Fast-moving objects

Asteroids < 0.01 AU (5 Lunar distances)

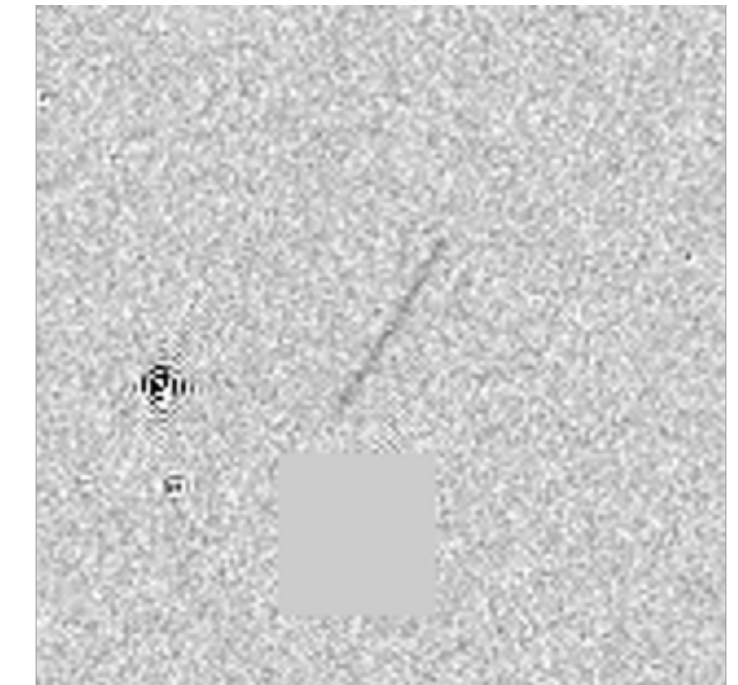
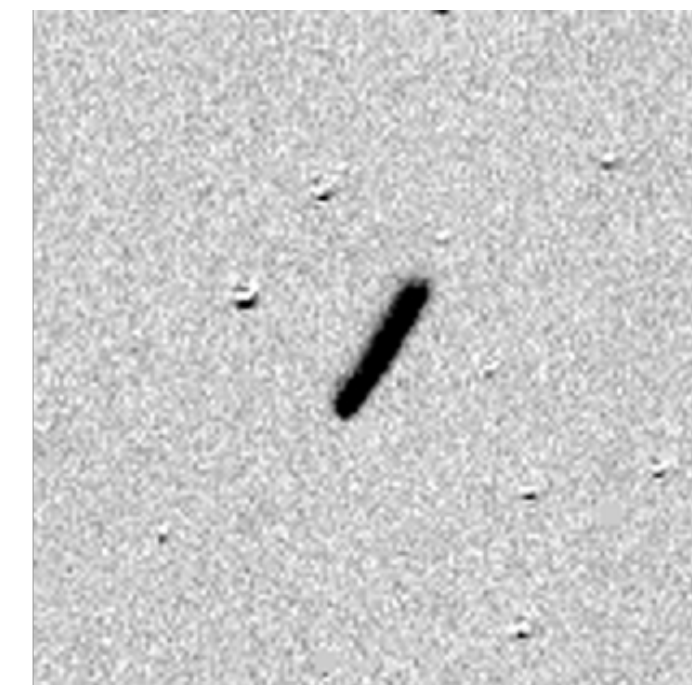
DeepStreaks reduced False Positives by 100x

Duev, Mahabal, ...



Asteroid 2018 KM1

2 Unknowns



Waszczak et al. 2013, Ye et al. 2019, Duev et al. 2019

Tails: DL-assisted discovery of comets (in ZTF)

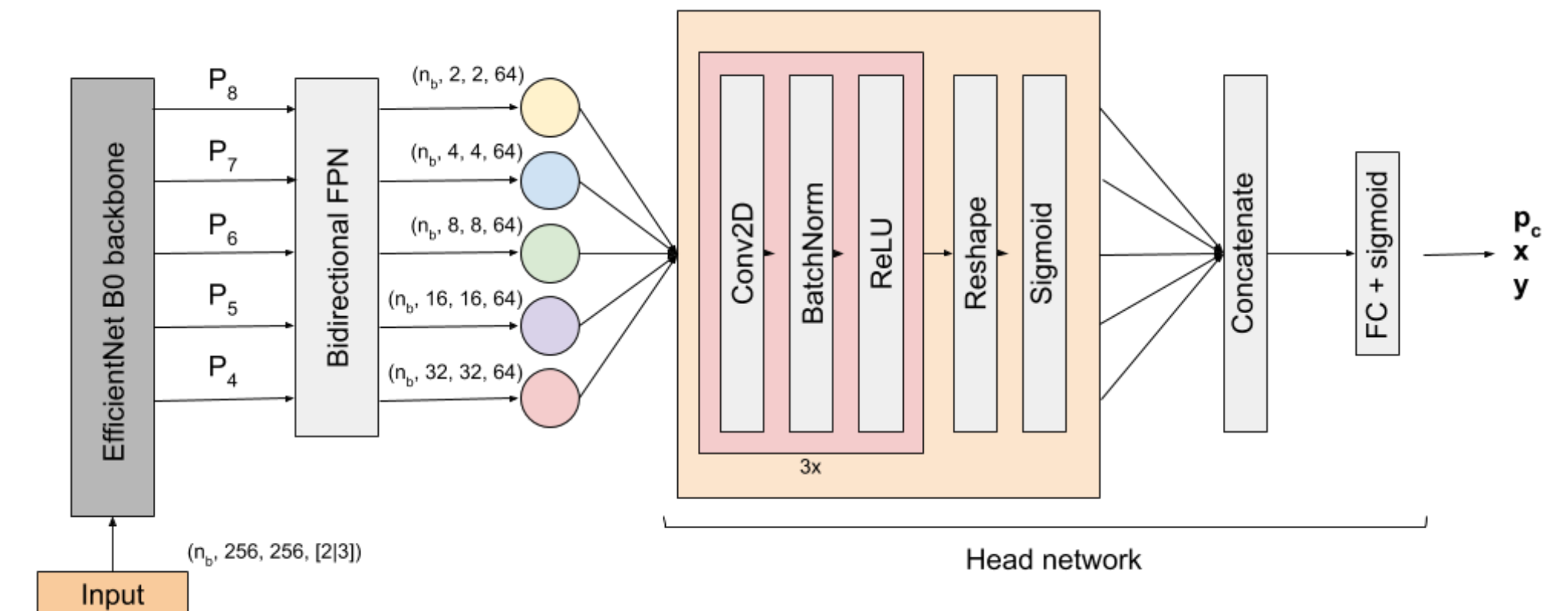
Custom state-of-the-art EfficientDet-based architecture

active learning, no difference image required

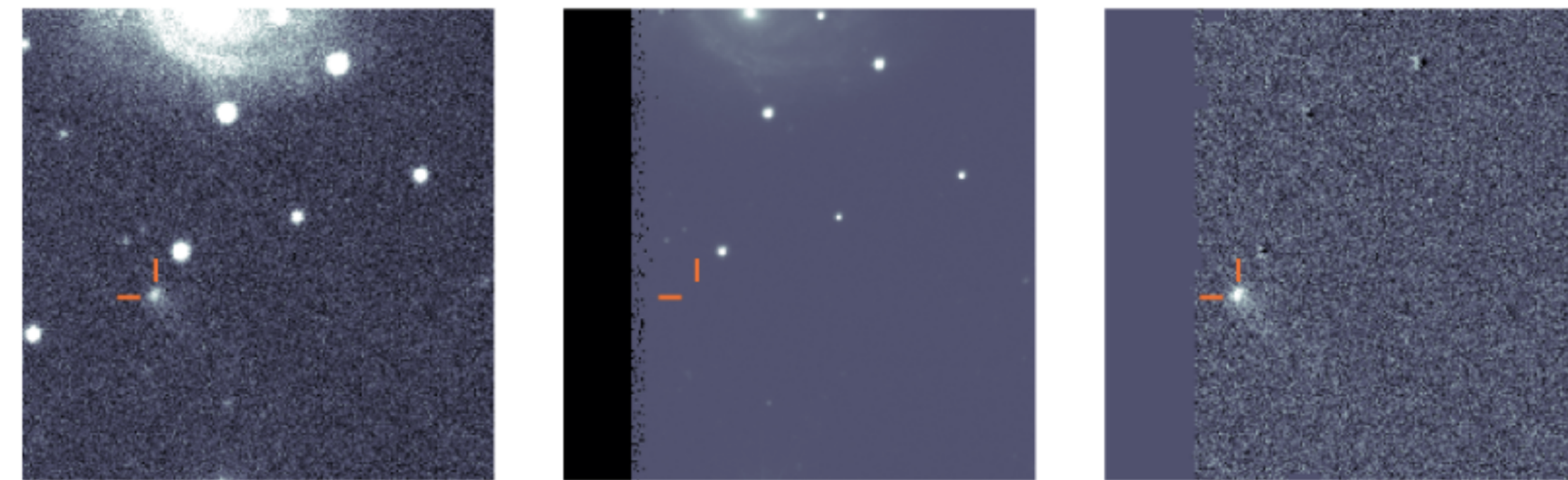
>99% accuracy, 1-2 pix RMSE (Horizons)

Production service running on Twilight data for 2 months

~10-20 nightly candidates



2I/Borisov from October 15, 2019



Duev+ (in prep)

Bryce to provide the first exciting results (next talk)

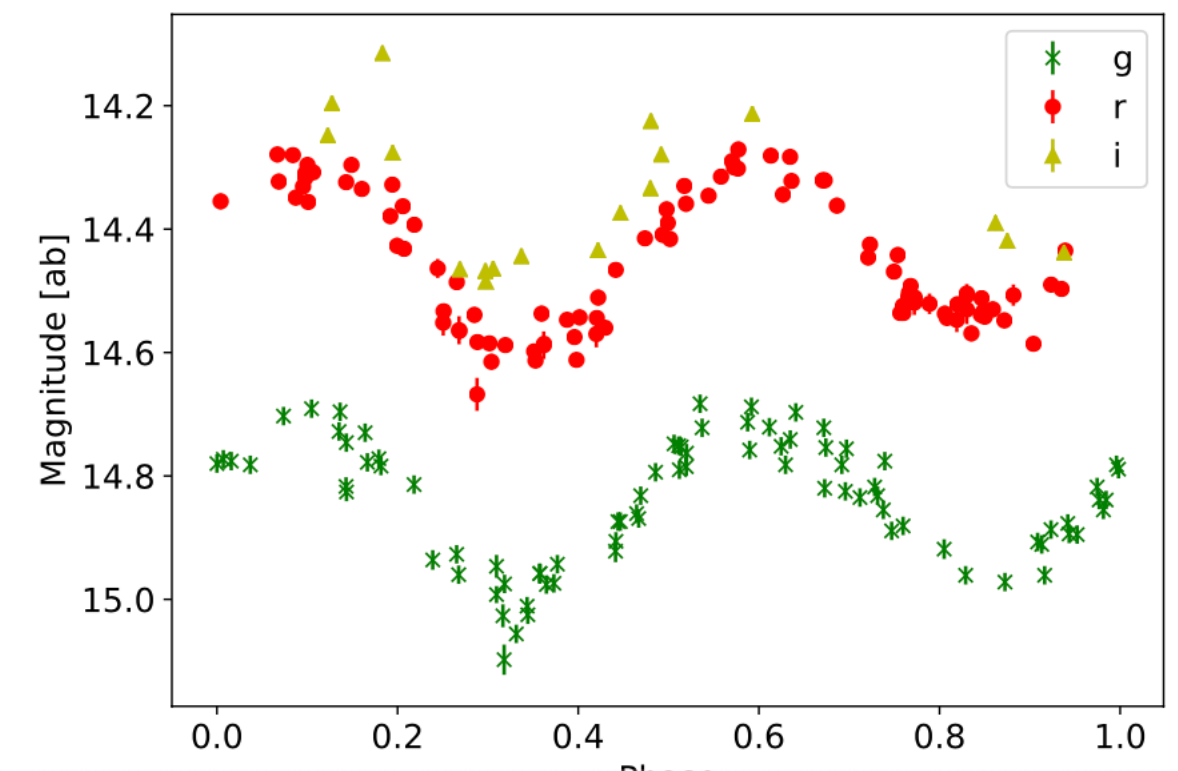
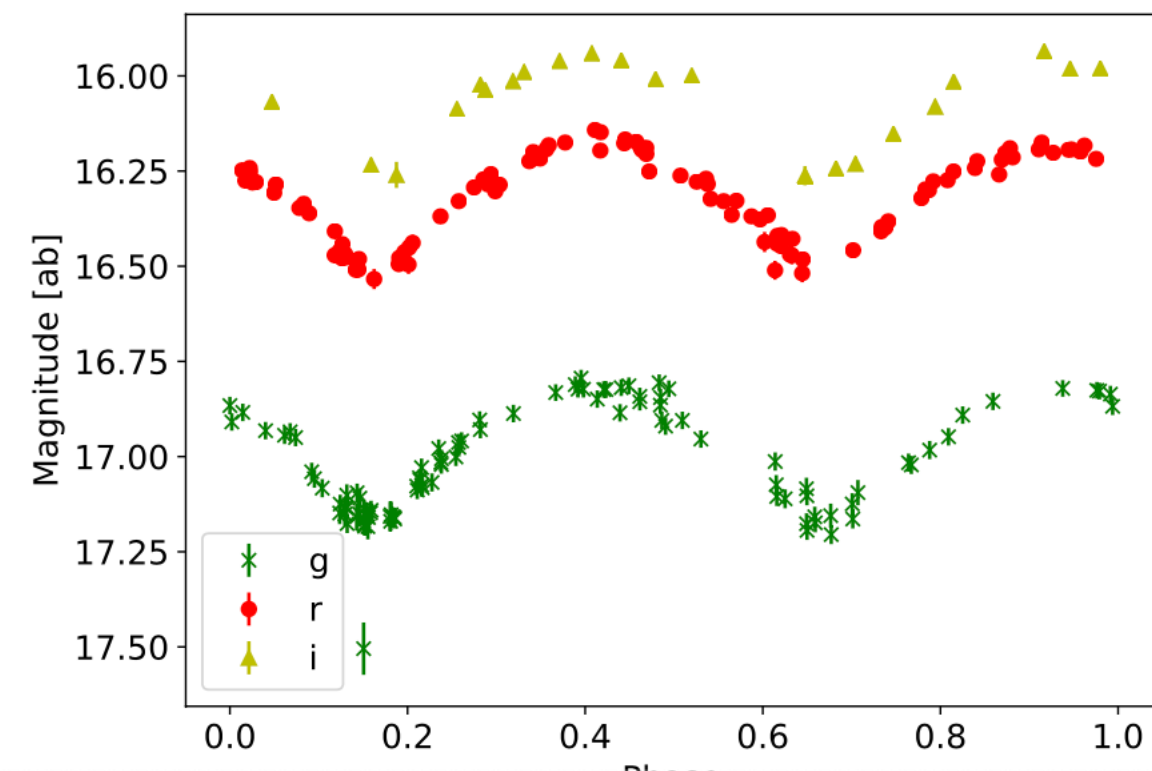
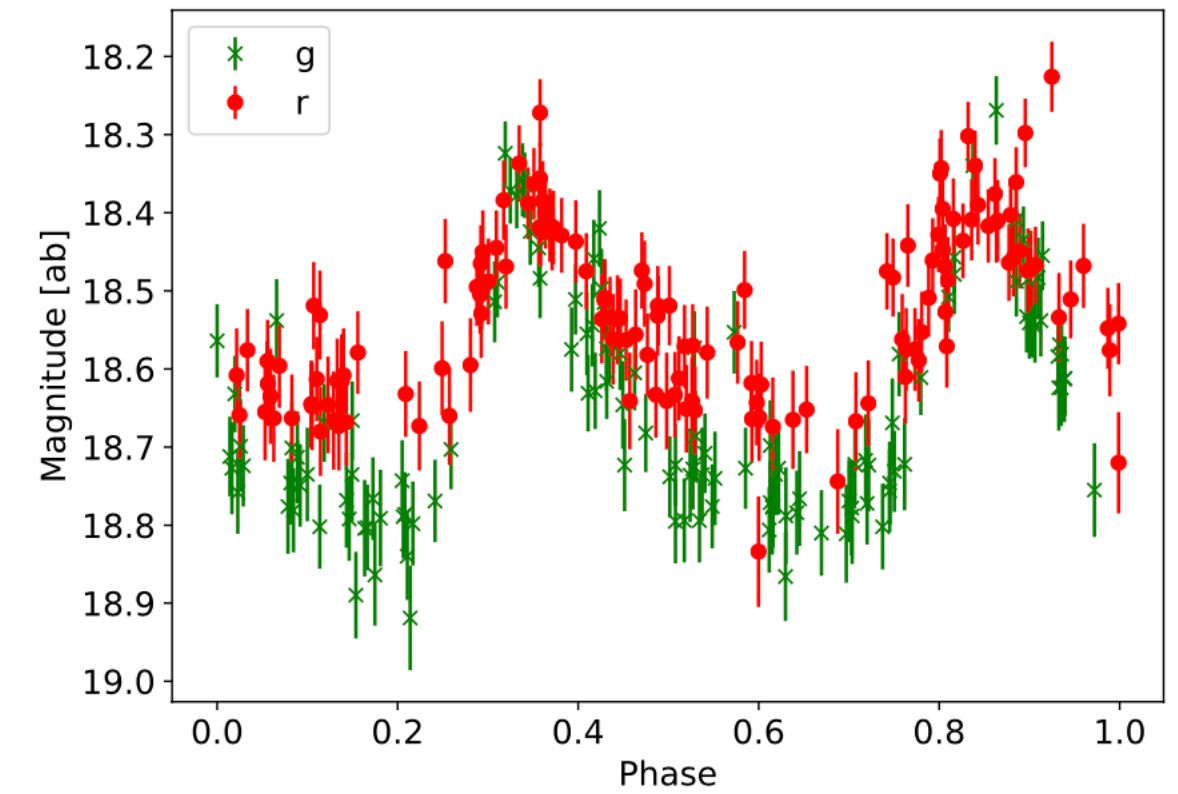
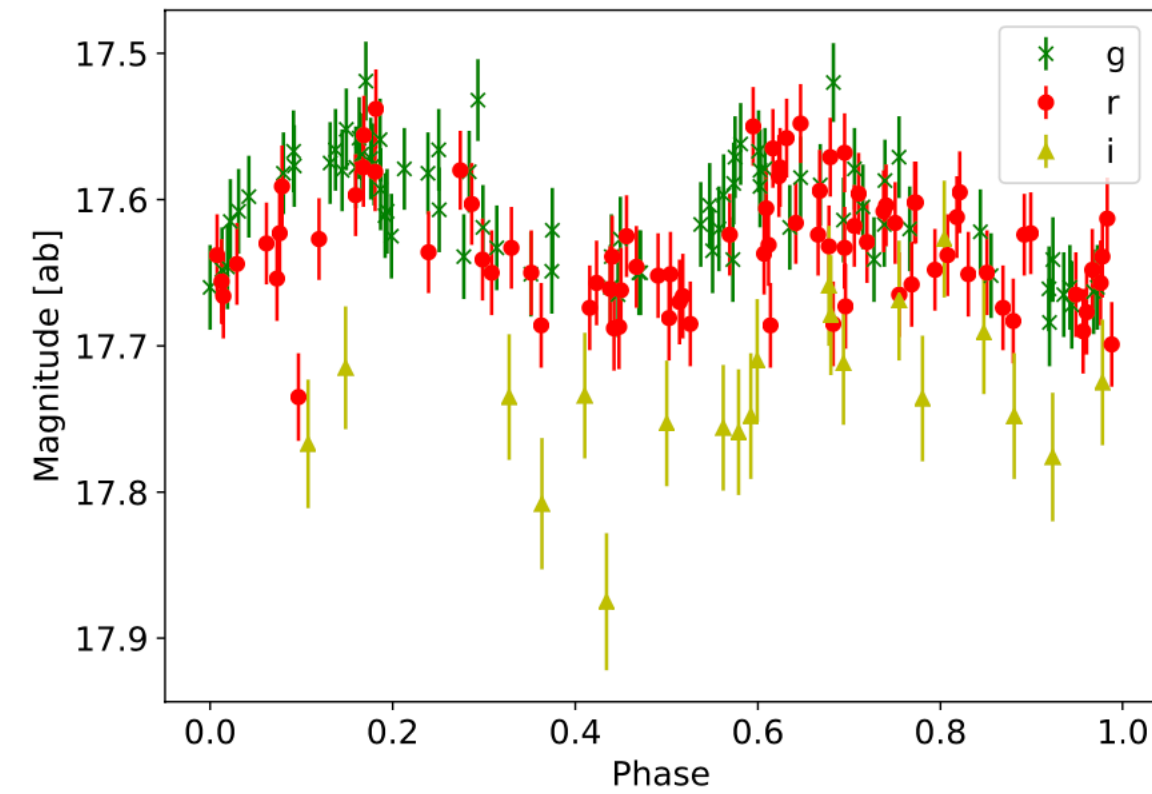
Variability

Period searching

- GPU-based methods
 - Lomb-Scargle
 - Analysis of Variance
 - Conditional Entropy

Micahel Coughlin
Przemek Mroz
Ethan Jaszewski
Matthew Graham

++



Coughlin et al. 2020

with

Jan Van Roestel

Dmitry Duev

Michael Coughlin

Przemek Mroz

Andrew Drake

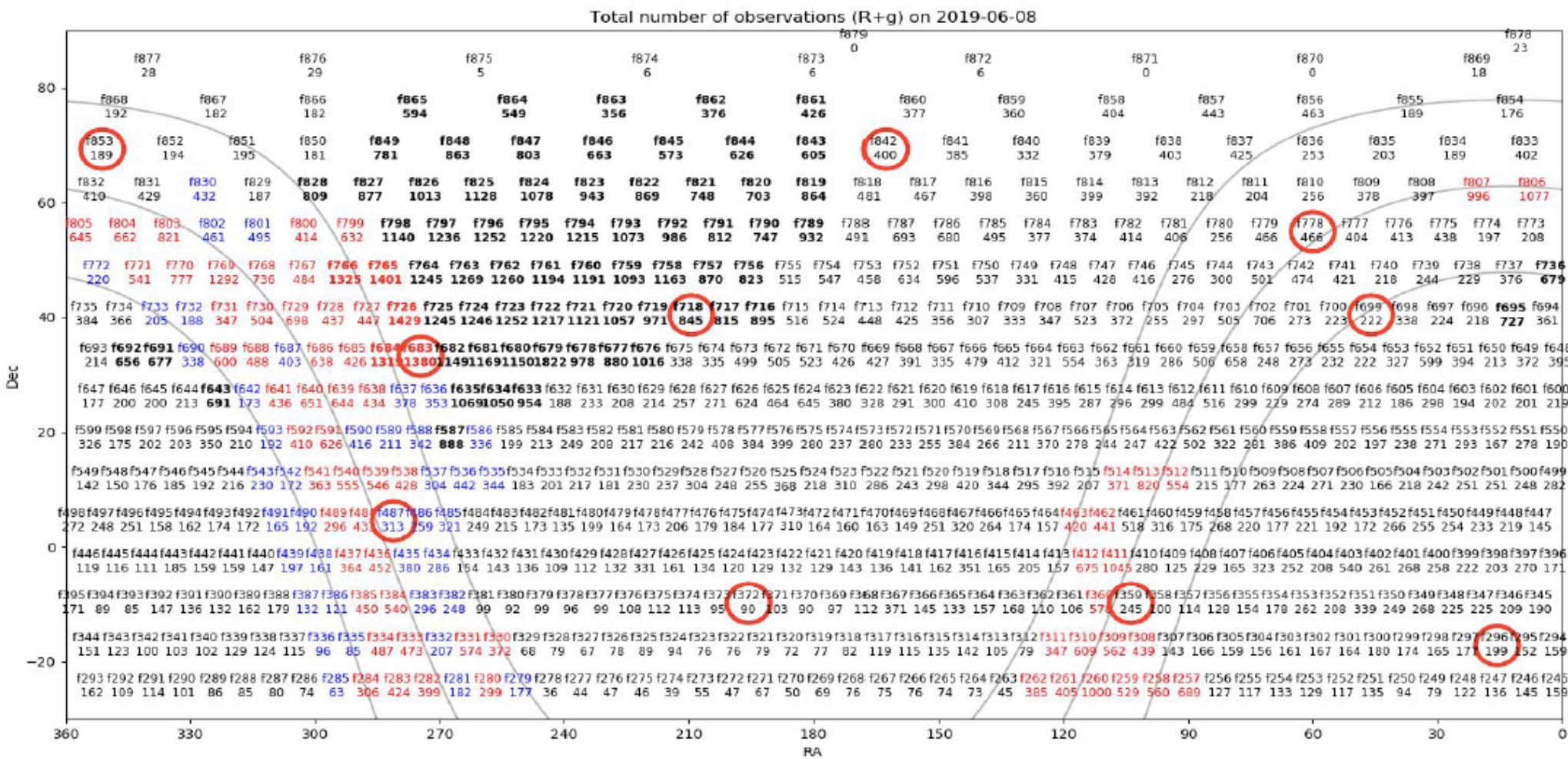
Matthew Graham

Lynne Hillenbrand

Variability

20 Fields

ZTF variable star catalog



• 34 M sources

• DNN (Light curve features + dmdt)

• XGBoost

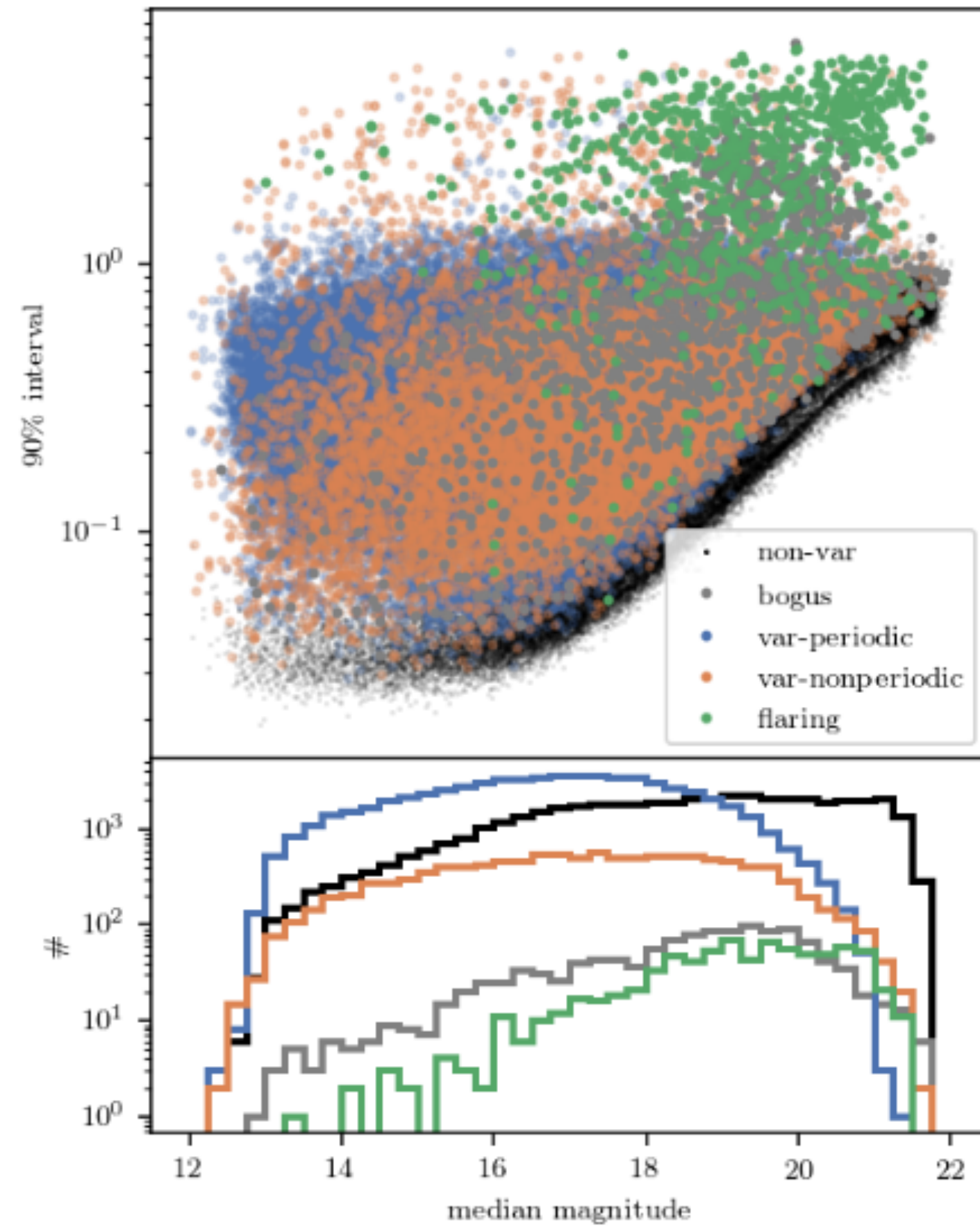
• A dozen binary classifiers

• Extensible to more classes

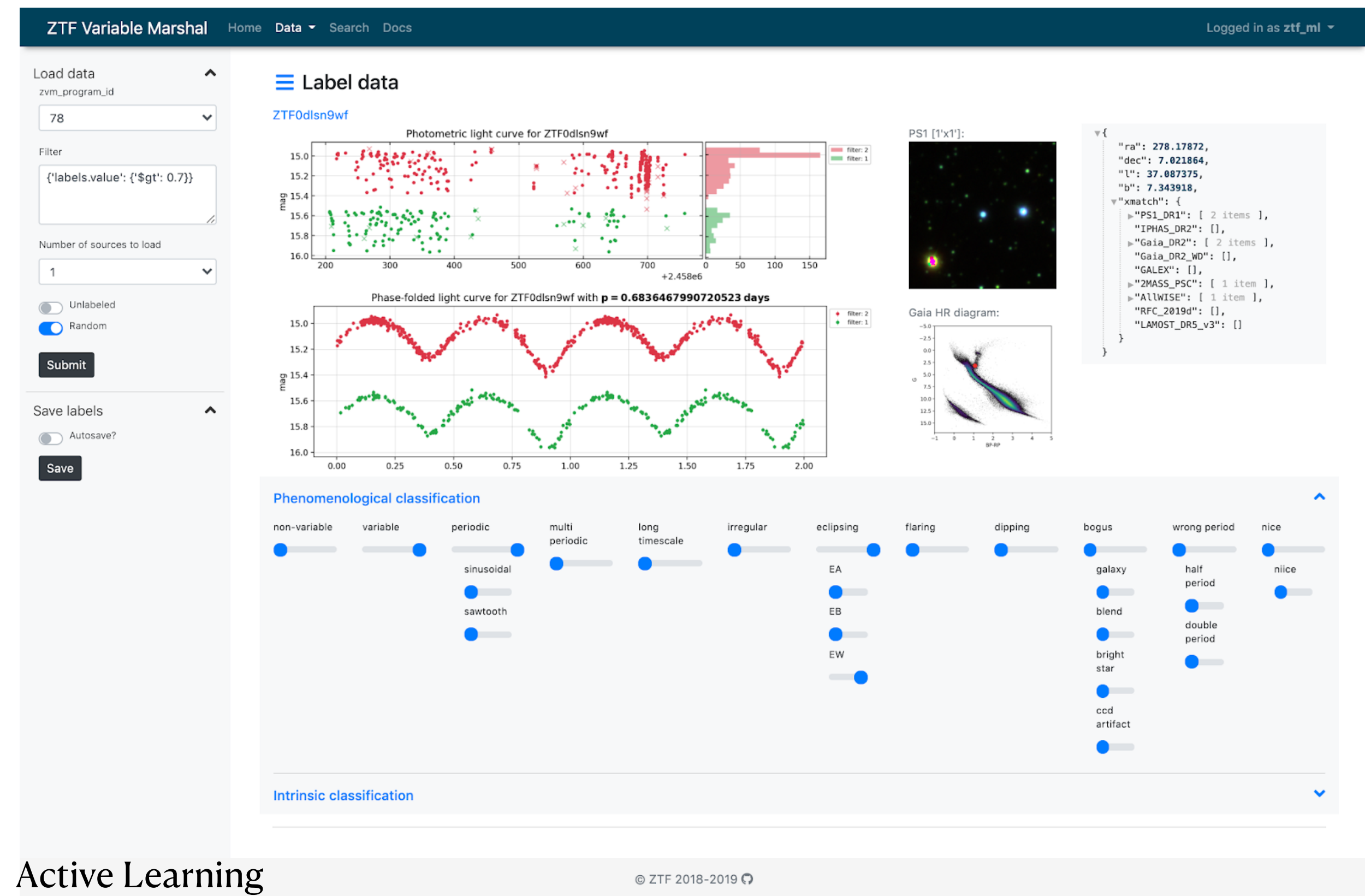
• Being extended to the entire sky (DR2)

Van Roestel et al. (being revised)

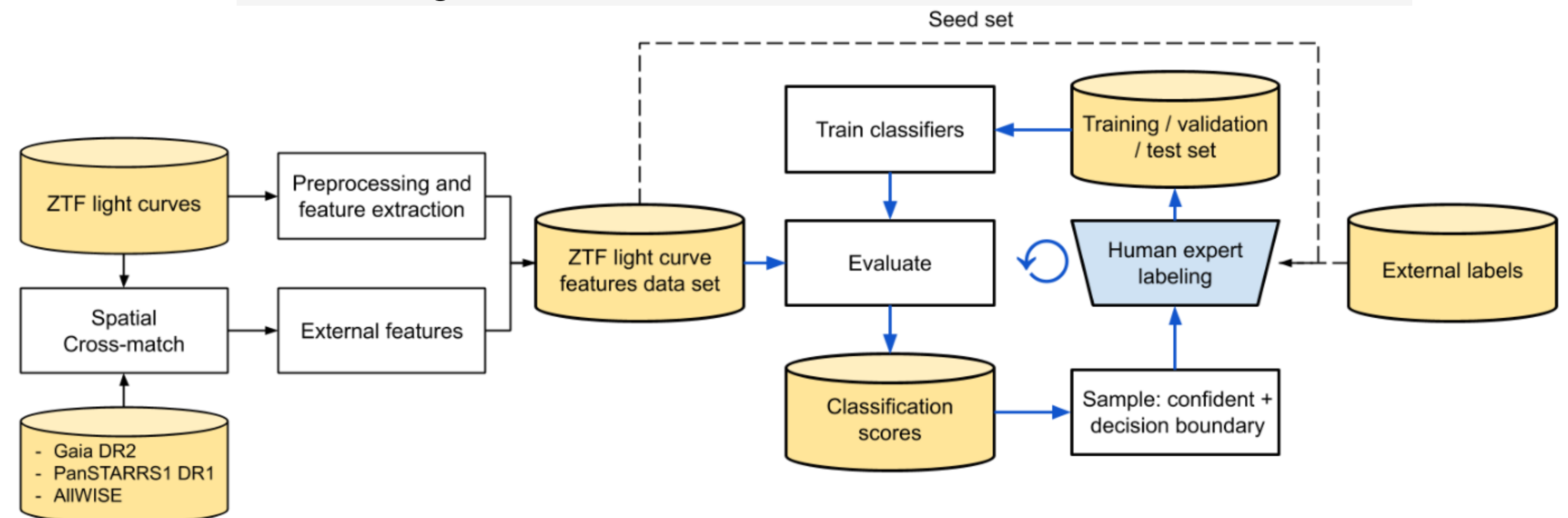
Variability



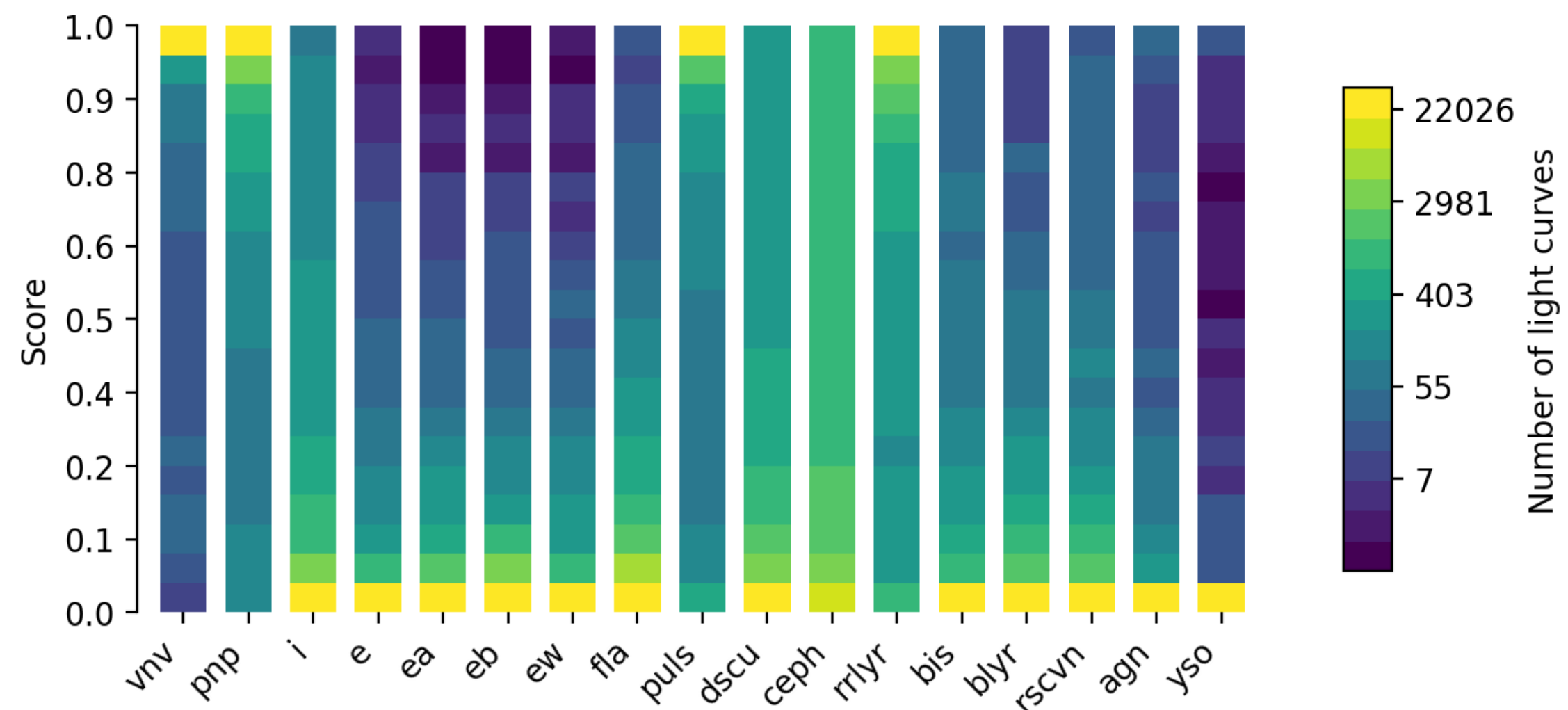
Includes external data



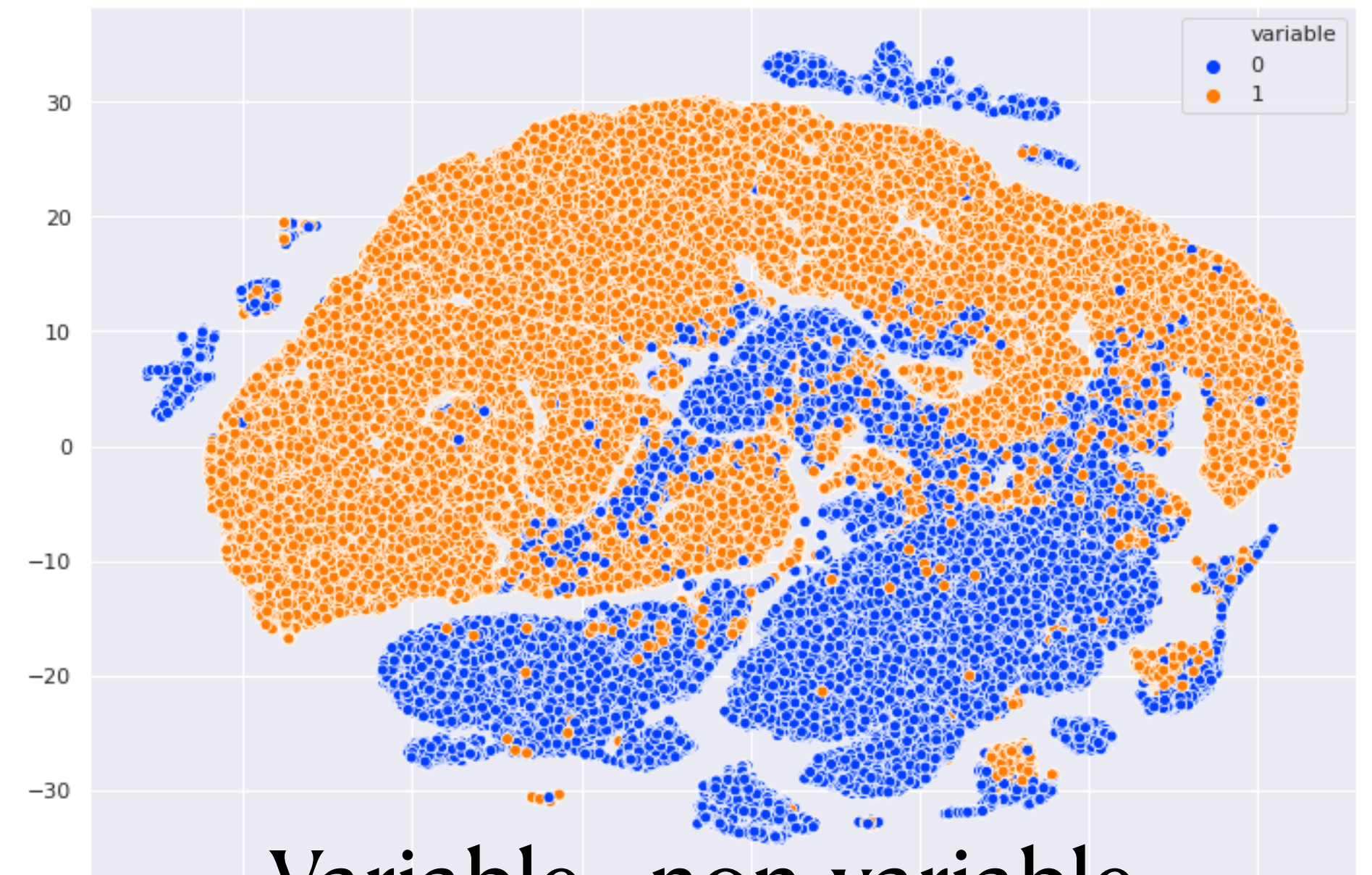
Active Learning



Variability

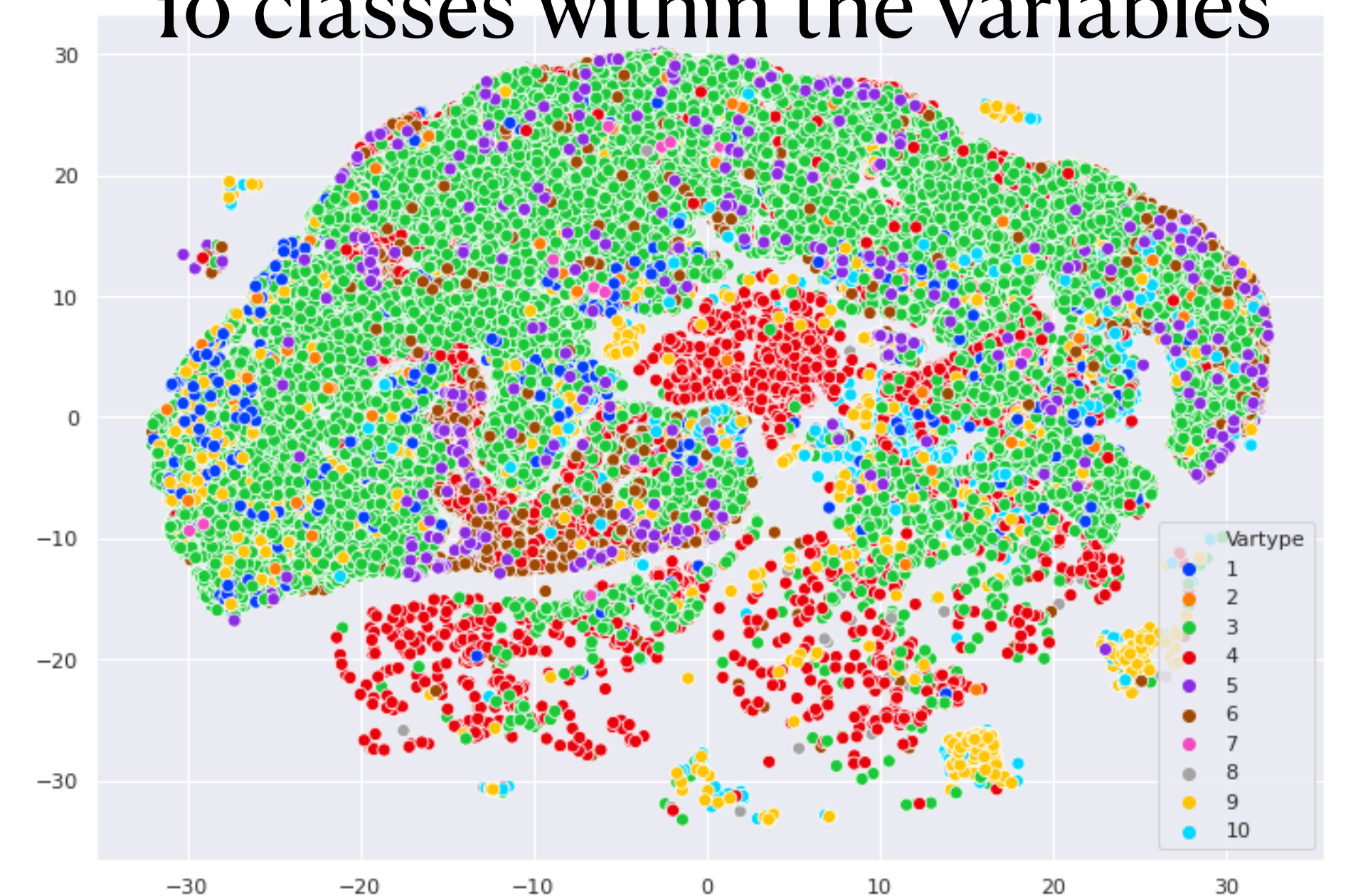


Set of binary classifiers



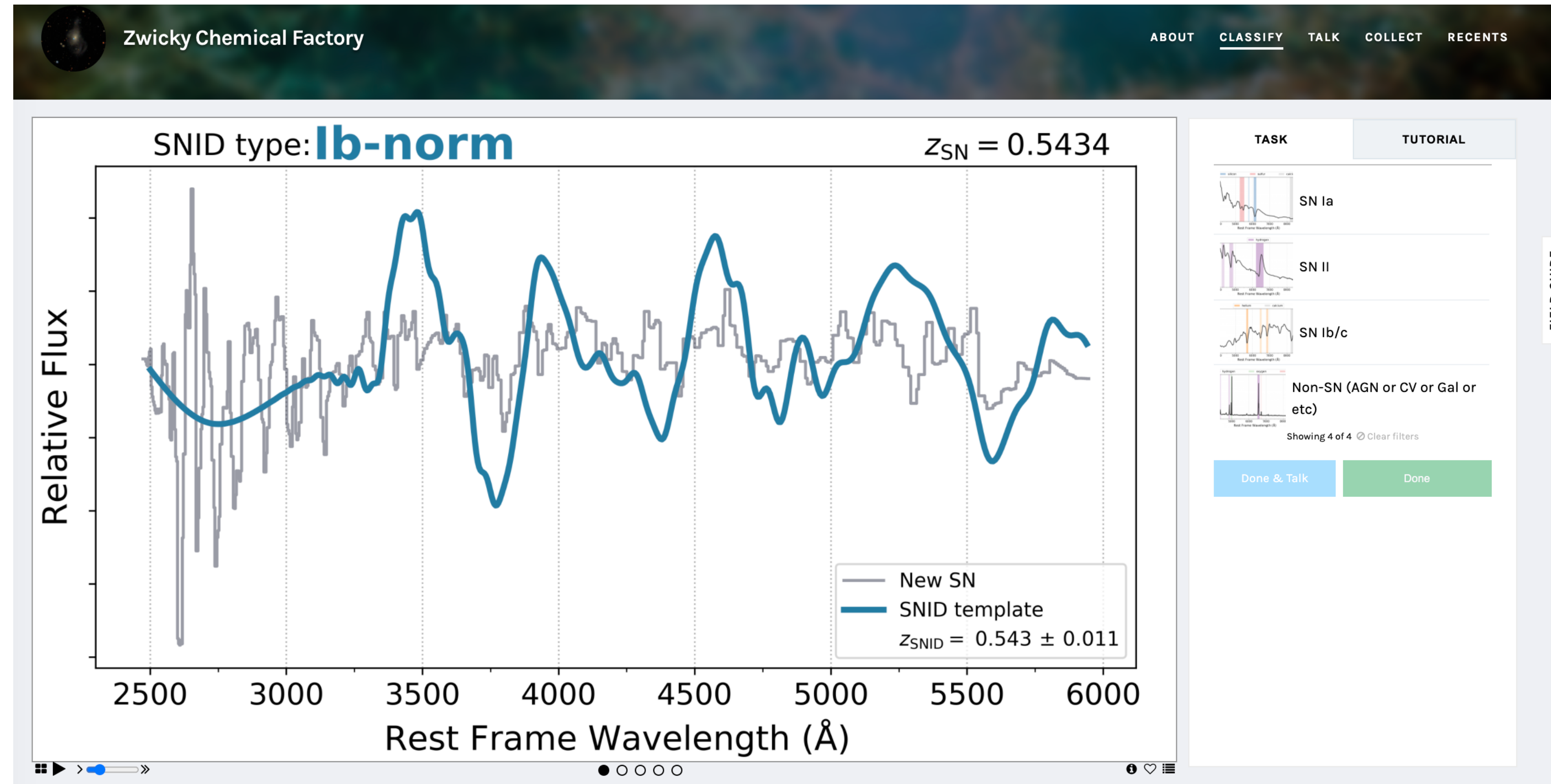
Variable - non-variable

10 classes within the variables



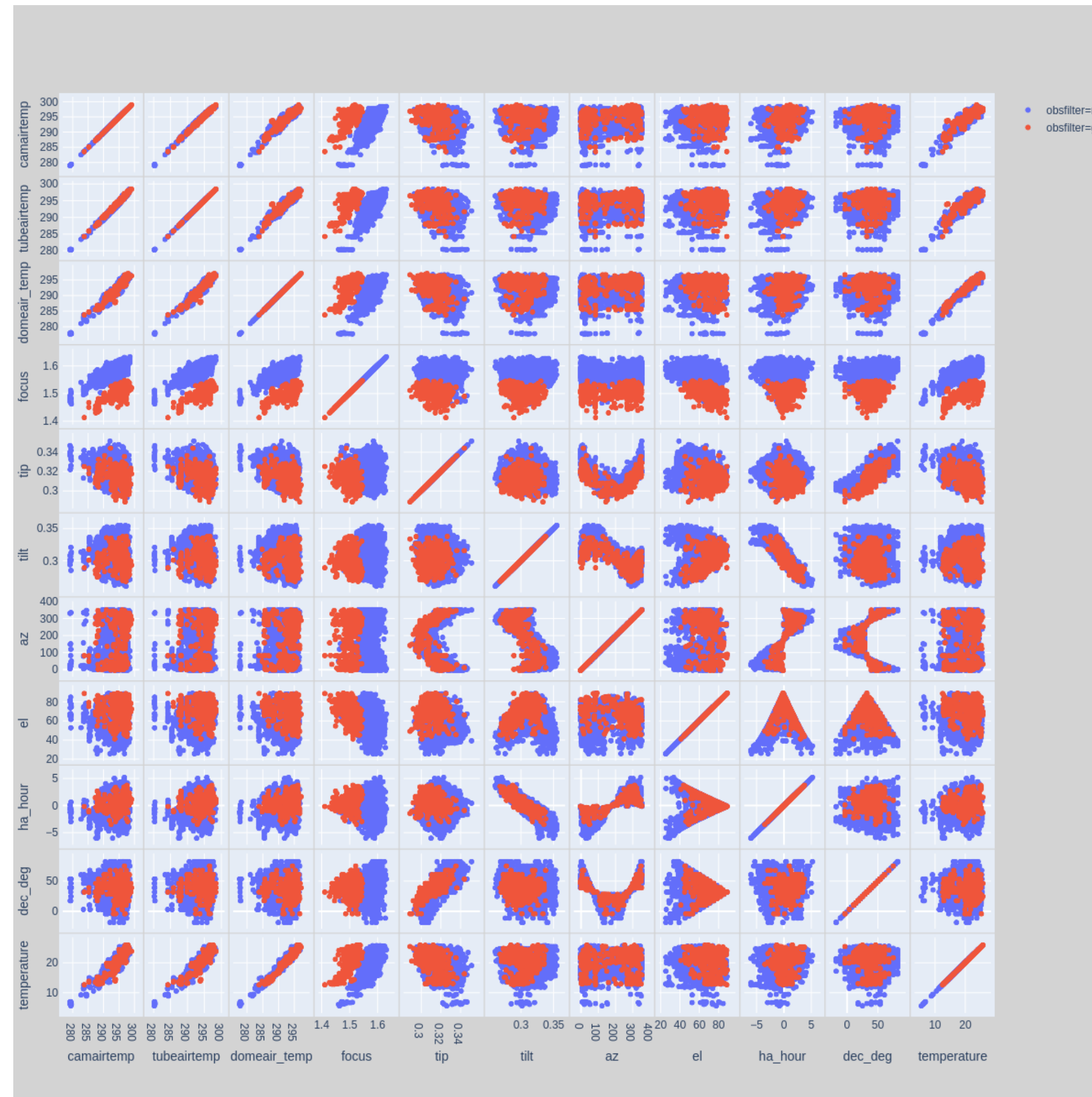
Zooniverse initiatives

- Zwicky's Quirky Transients (Walters, Mahabal, Karmarkar, ...)
- Variables (Coughlin/Miller, ...)
- Zwicky Chemical Factory (Fremling/Miller, ...)



Focus improvements

Camera Temp
Tube Air Temp
Dome Air Temp
Tip
Tilt
Az
El
HA
Dec
Temperature
Focus



with
Richard Walters
Reed Riddle
Matthew Graham

More classes and meta-classification

- Combining DNN and XGBoost
- Online (incremental) classification

For phenomenological

variable - not variable

light curve features

raw (mean, median, skew, ...)

phased (period, significance, ...)

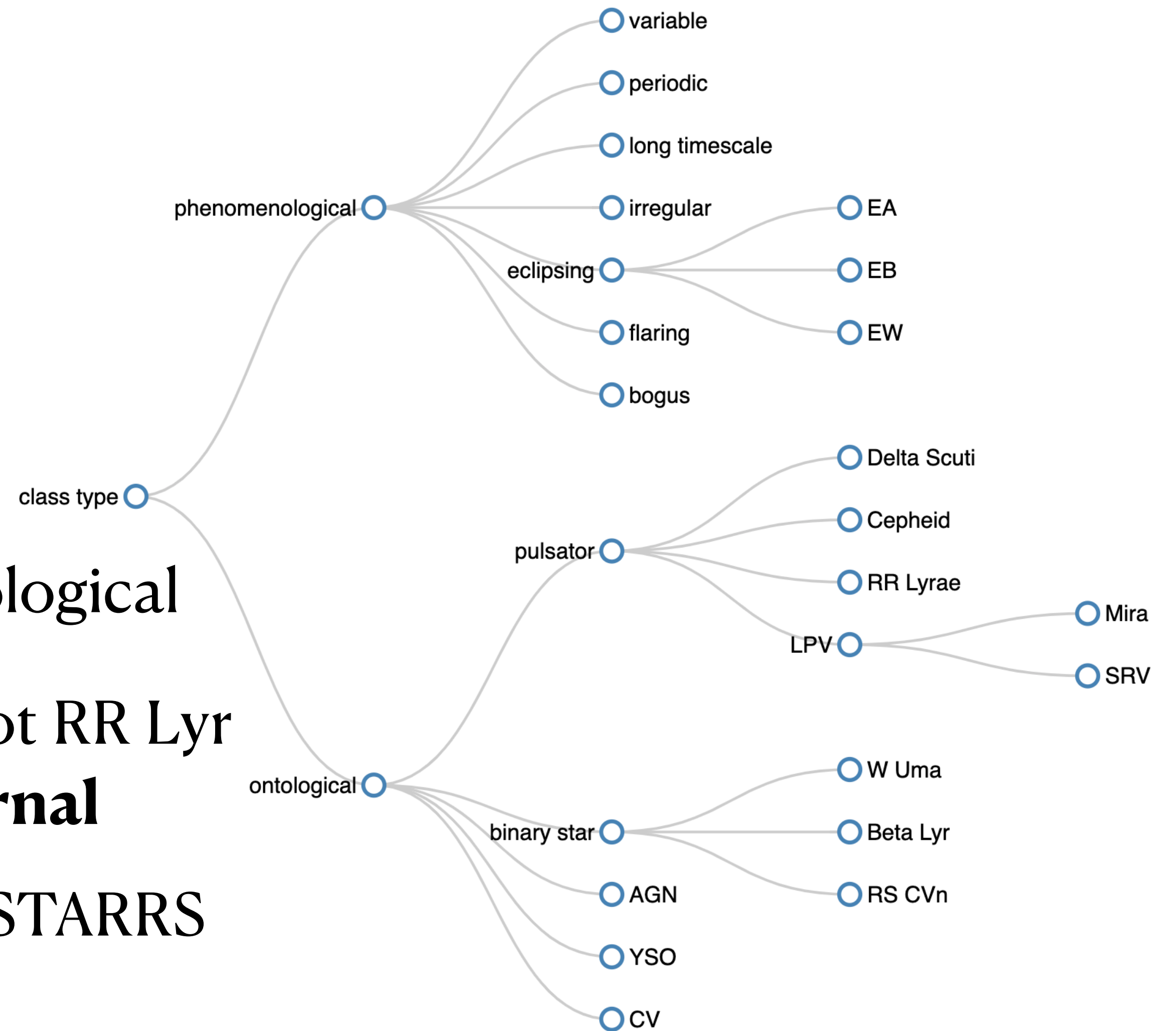
For ontological

RR Lyr - not RR Lyr
external

Pan-STARRS

Gaia

ALLWISE



Finding anomalies

- Aims:
 - Identify rare objects
 - Identify subclasses
- Data:
 - Raw Time series
 - Features derived from time series and metadata
 - Archival data
- Methods: various
 - dimensionality reduction (e.g. t-SNE, UMAP),
 - density based (e.g. HDBSCAN),
 - specially designed deep learning representations of time series (e.g. dmdt)

Statistical studies

- It is important to spend some resources not on individual sources but on ensembles
- Take all of ZTF I data and use that

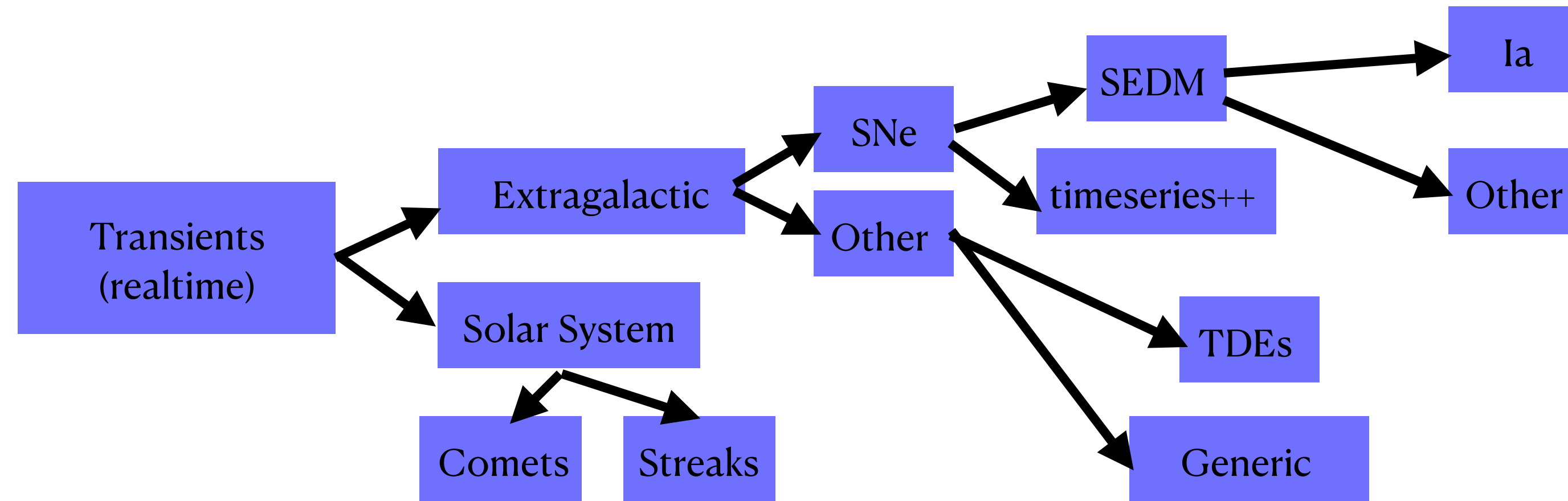
(New) science of interest to (new) partners

ZTF ML Thu 2 PM PT meetings

Contact Ashish Mahabal aam@astro.caltech.edu

ML update - getting ready for ZTF II

Ashish Mahabal++, 2020-10-19



Infrastructure
Fritz++
Notebooks
GitHub

Tools
DNN (CNN/dmdt)
XGBoost
DBScan
Tails++

Anomaly detection
(time series, features)

stats/populations
fainter sources/subclases

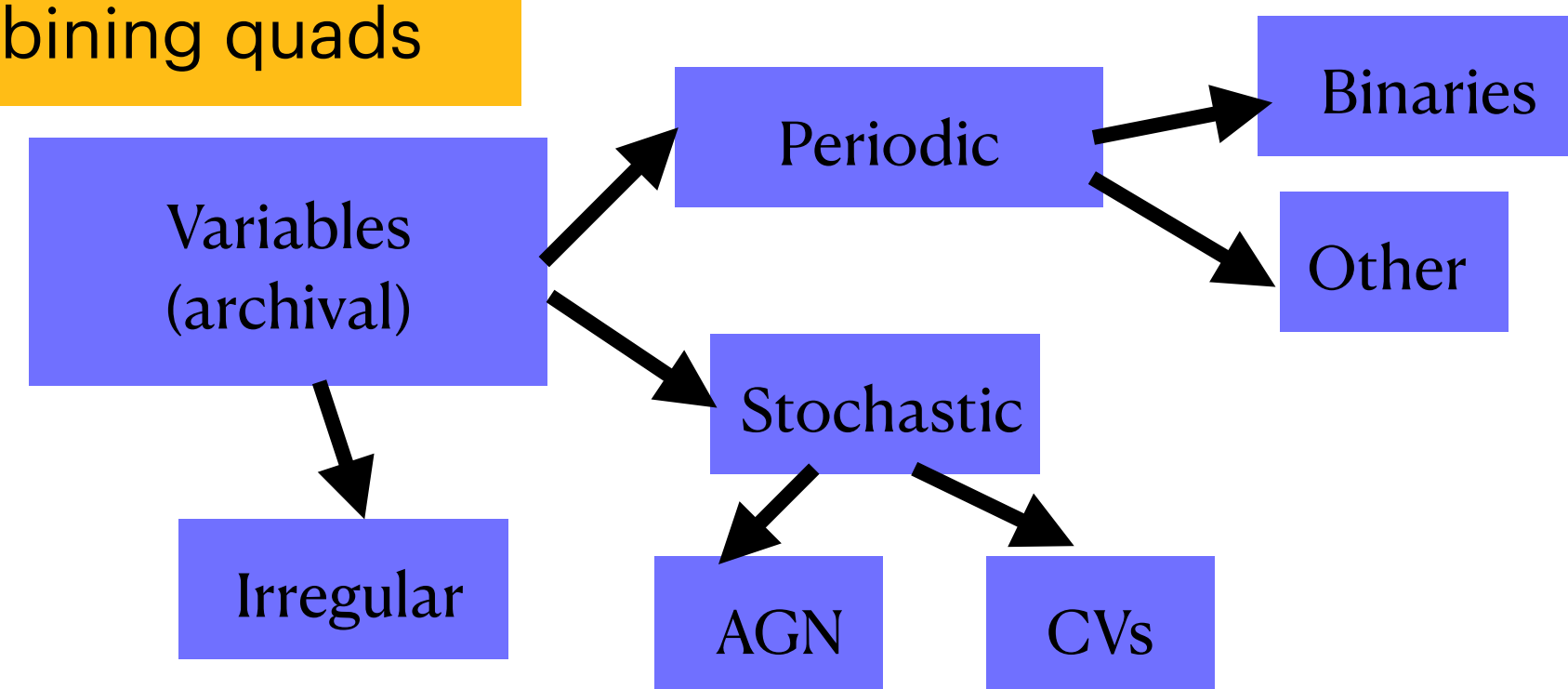
Products
Catalogs
Deep coadds

Focus
Deconvolution

Meta-classification
External data
Feature importance

Specific science
YSOs
microlensing
Partner interests

Combining filters
Combining quads



Recent and ongoing contributions:

Dima, Michael, Matthew, Jan, Przemek, Andrew, Bryce, Christoffer,
Dan, Adam, Richard, Lynne, Yashvi, Suvi, Sjoert, Jakob, Rahul, Nikolas
++ (many more)