



ALeRCE

Automatic Learning for the
Rapid Classification of Events



The Universe in a stream: challenges and progress of the ALeRCE broker

Francisco Förster; G. Cabrera-Vives; E. Castillo-Navarrete; P. A. Estévez; P. Sánchez-Sáez; J. Arredondo; F. E. Bauer; R. Carrasco-Davis; M. Catelan; F. Elorrieta; S. Eyheramendy; P. Huijse; G. Pignata; E. Reyes; I. Reyes; D. Rodríguez-Mancini; D. Ruz-Mieres; C. Valenzuela; I. Álvarez-Maldonado; N. Astorga; J. Borissova; A. Clocchiatti; D. De Cicco; C. Donoso-Oliva; M. J. Graham; L. Hernández-García; R. Kurtev; A. Mahabal; J.C. Maureira; R. Molina-Ferreiro; A. Moya; A. Muñoz Arancibia; W. Palma; M. Pérez-Carrasco; A. Papageorgiou; P. Protopapas; M. Romero; L. Sabatini-Gacitua; A. Sánchez; J. San Martín; C. Sepúlveda-Cobo; E. Vera; J. R. Vergara

Förster+2020, Carrasco-Davis+2020, Sánchez-Sáez+2020, <http://alerce.science/>

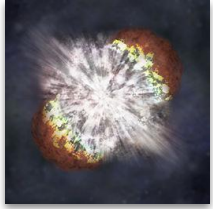




ALeRCE is a Chilean-led initiative
to build a community broker for
LSST and other large etendue
survey telescopes



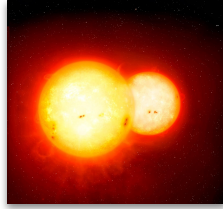
Scientific questions



Transients

Progenitors of stellar explosions (outermost layers) & explosion physics (ejecta structure).

How does the Universe expand?



Variable stars

Microlensing events, changing mode stellar pulsators, rapid reaction to eclipsing events, eruptive events.

How is matter distributed in our galaxy?



Supermassive black holes

Changing state AGNs, reverberation mapping, detection of intermediate mass black holes, tidal disruption events.

How do black holes affect the evolution of galaxies?



Asteroids

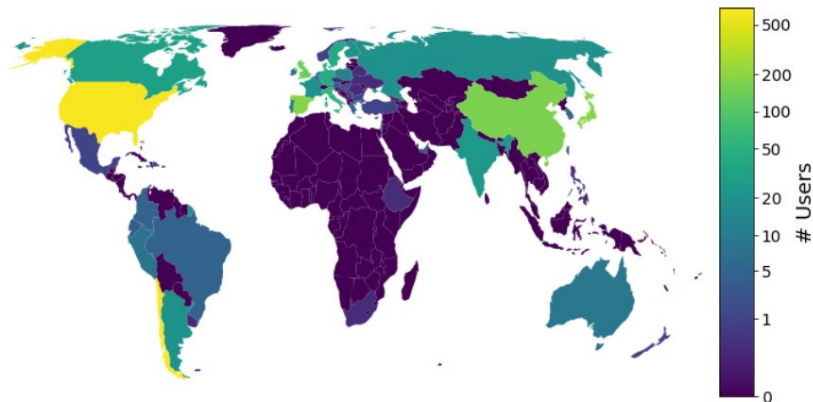
Near earth objects, size distribution, collisional or rotational disintegration.

How was the solar system formed?



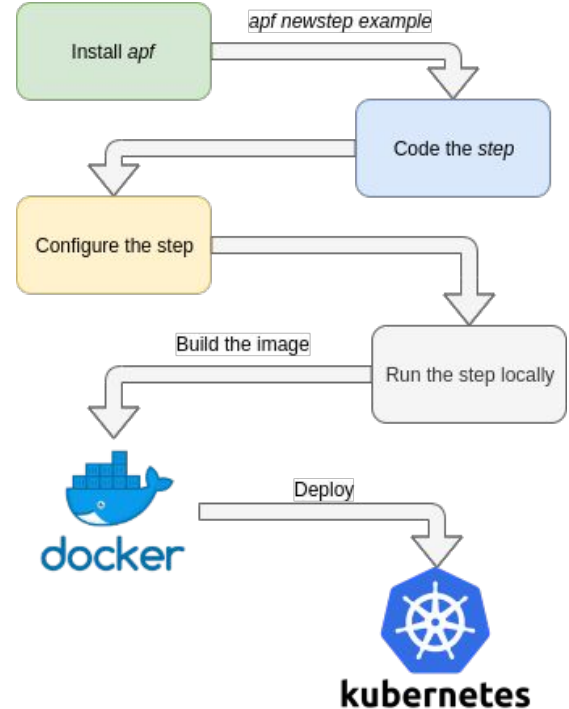
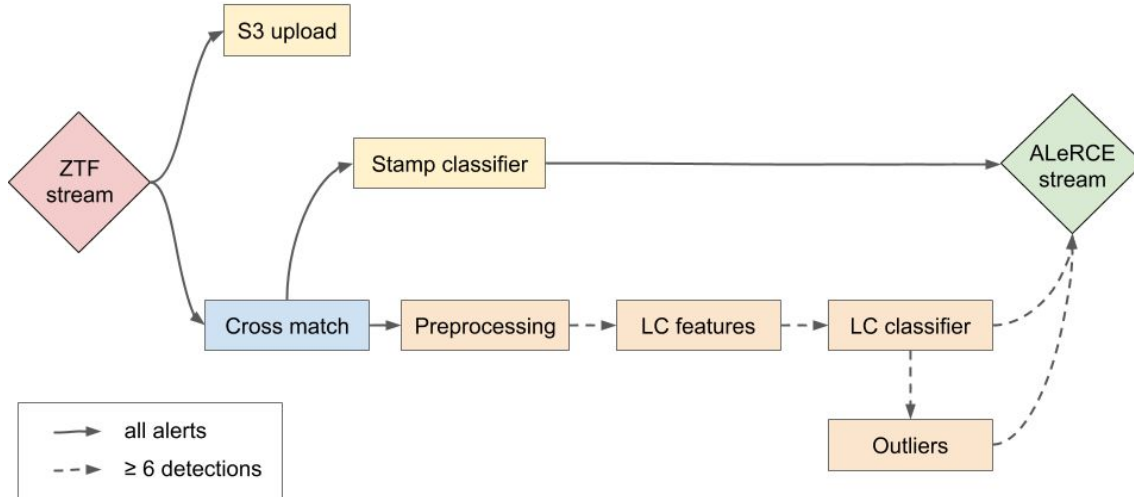
ALeRCE

>2000 users from >50 countries (Google Analytics)





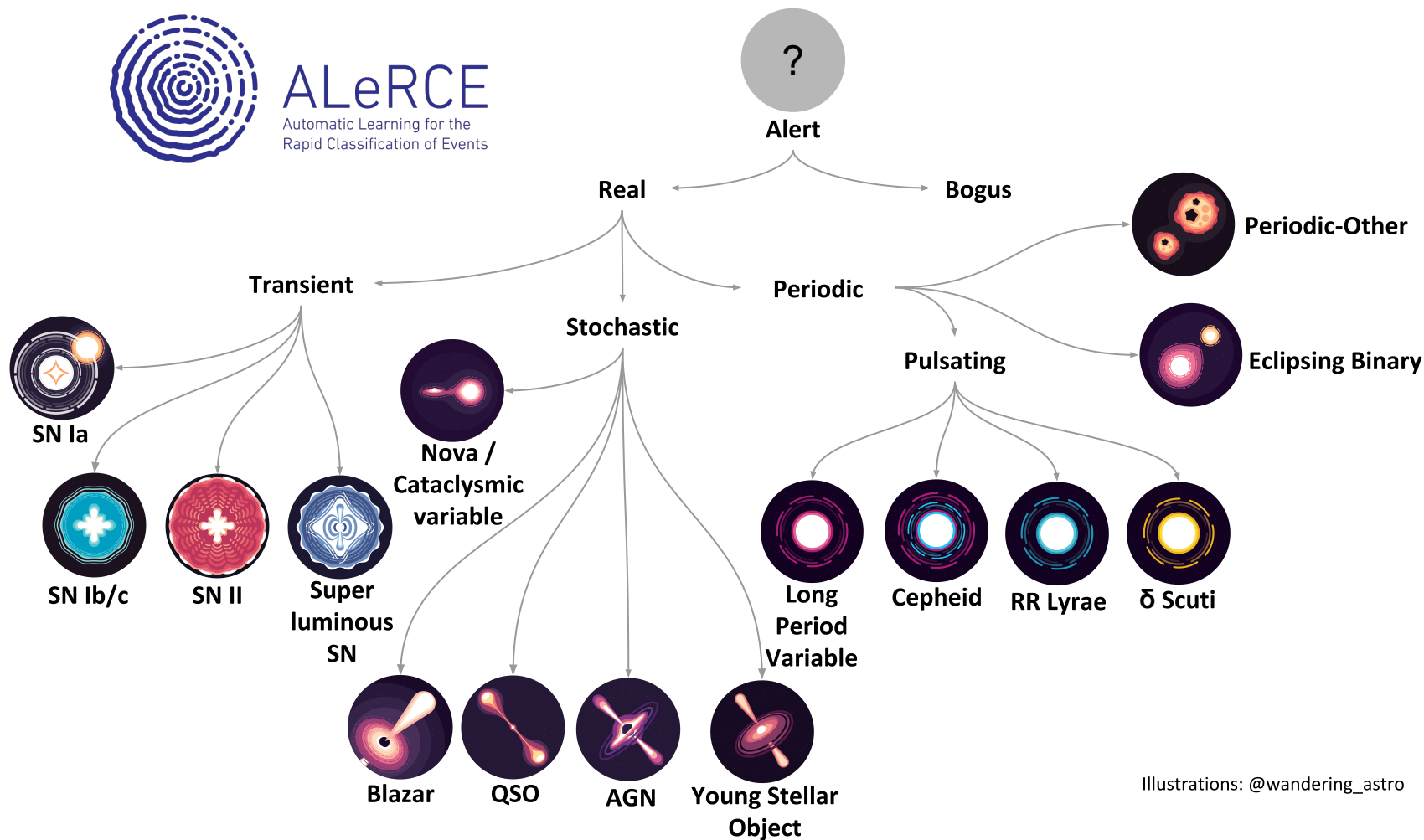
Alert Processing Framework (APF)





ALeRCE

Automatic Learning for the
Rapid Classification of Events



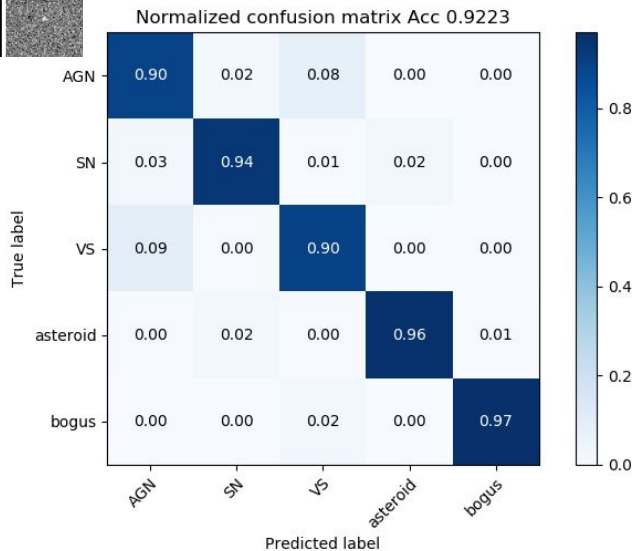
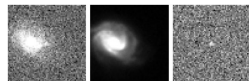
Illustrations: @wandering_astro



Stamp based classifier

Convolutional neural network
Carrasco-Davis et al. 2020

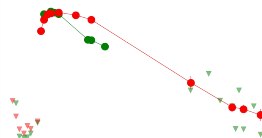
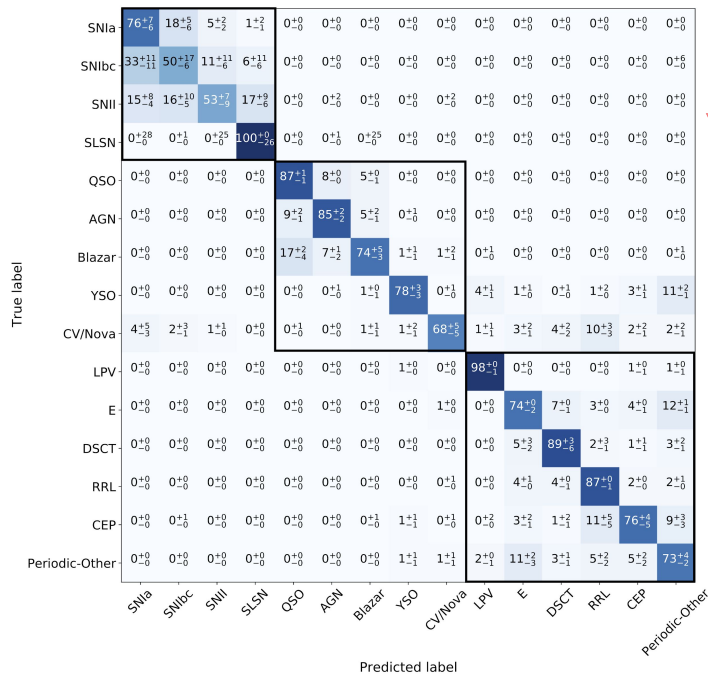
AGN, SN, VS, asteroid, bogus



Light curve based classifier

Balanced Hierarchical Random Forest
Sánchez-Sáez et al. 2020

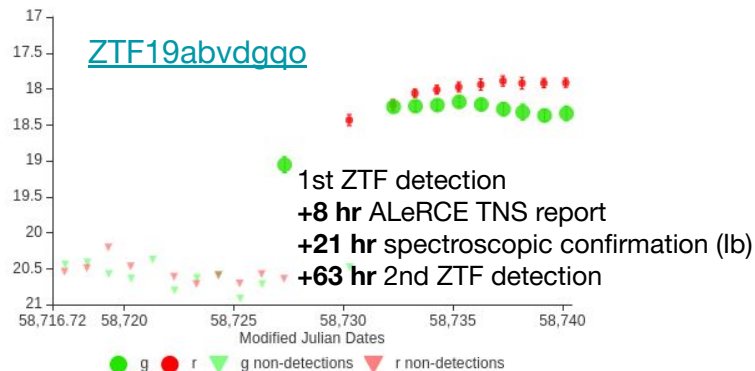
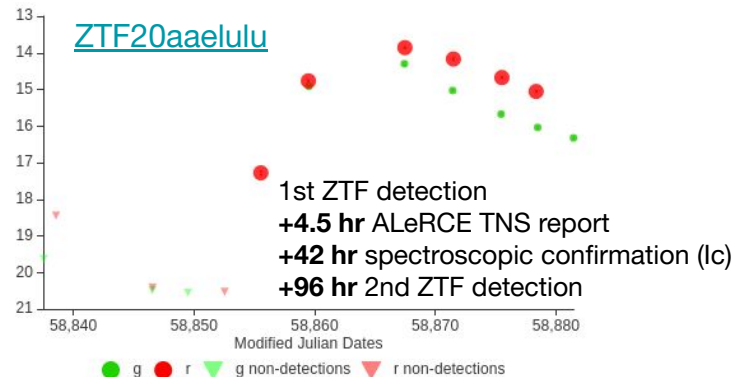
SNla, SNIbc, SNIi, SLSNe, QSO, AGN, Blazar, YSO, CV/Nova, LPV, E, DSCT, RRL, CEP, Periodic-Other





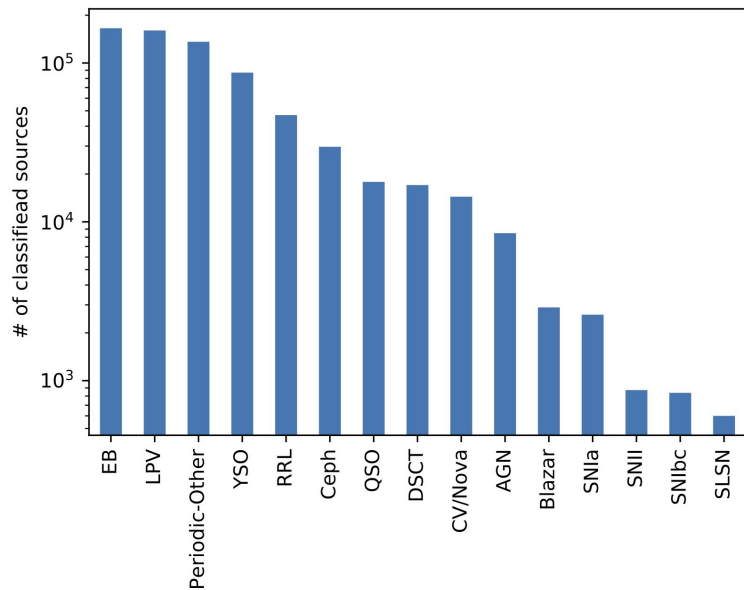
SNe reported by ALeRCE (stamp classifier)

>4400 SNe submitted to TNS
(currently ~70% efficiency, ~80% purity)

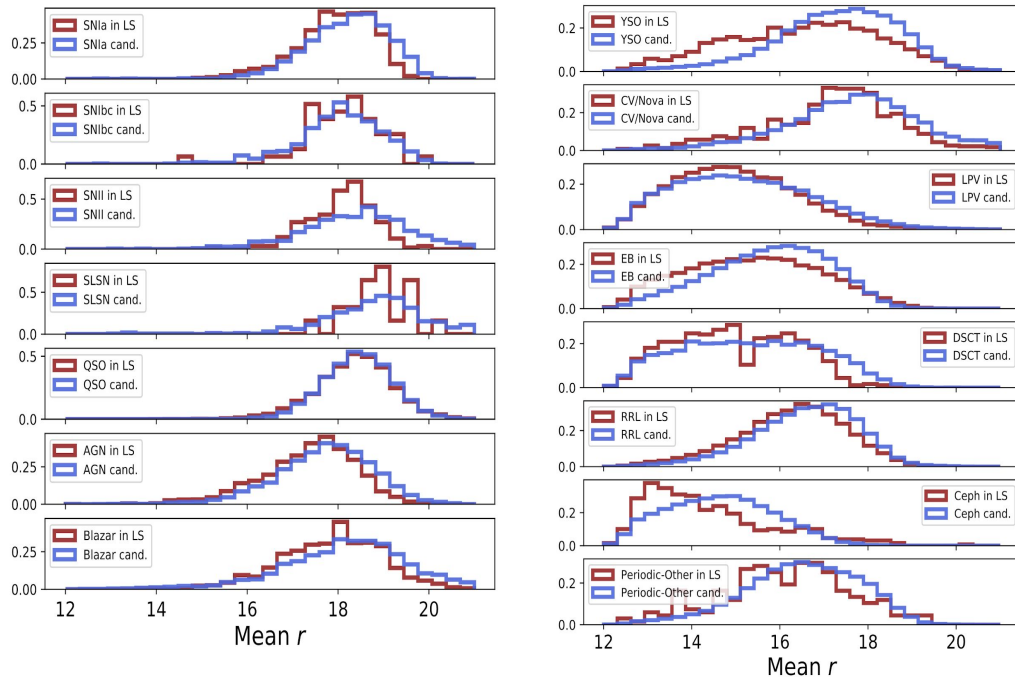




Objects classified by the light curve classifier

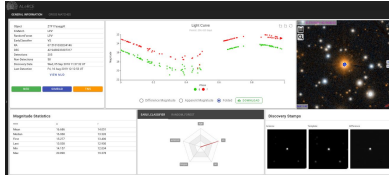


Class distribution in the unlabeled set (Sánchez-Sáez et al. 2020)

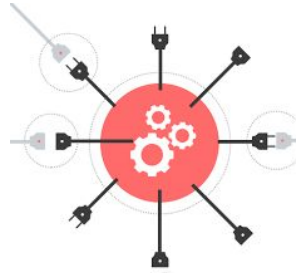


Magnitude distribution in the labeled vs unlabeled sets (Sánchez-Sáez et al. 2020)

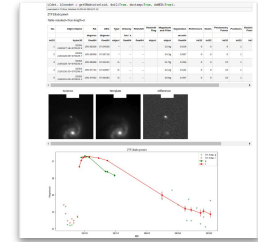
Web Interfaces



API



Jupyter
Notebooks

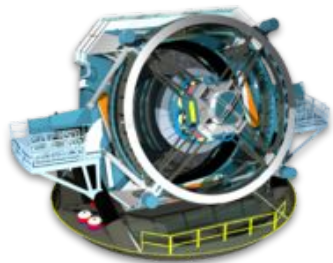


ALeRCE

Automatic Learning for the
Rapid Classification of Events

<http://alerce.science>

Output stream
(real-time follow-up)

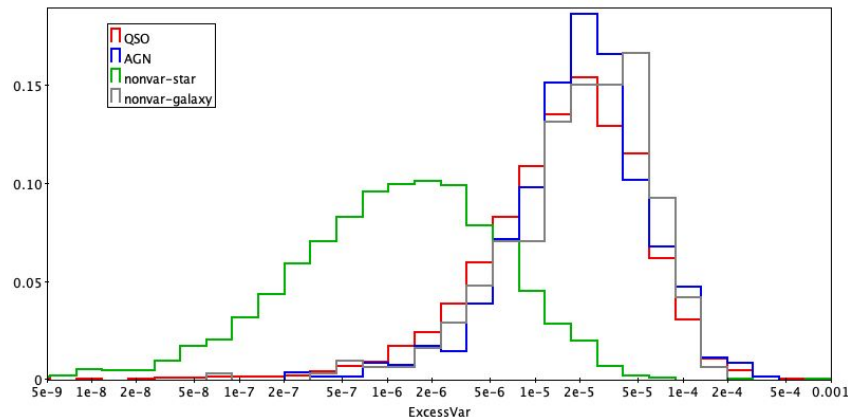
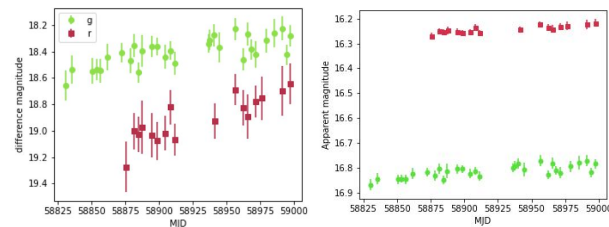
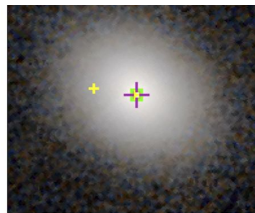


TOMs



New: ZTF DR3 vs ZTF alert stream

- Stripe-82
 - labeled set alerts: labeled set DR3 = 1:10
 - known AGNs in alerts: known AGNs in DR3 = 1:20
- ZTF alert stream is better to study transients and nuclear variability in extended sources.
- Recalibrating ZTF DR3 light curves to better separate variable from non variable stars.
- Using ZTF DR3 autocorrelation structure to separate variable from non variables galaxies



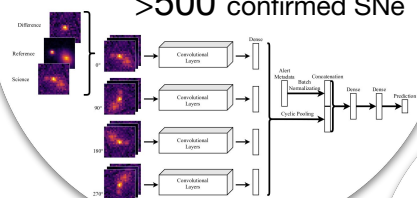


ALeRCE

Automatic Learning for the
Rapid Classification of Events

Stamp Classifier

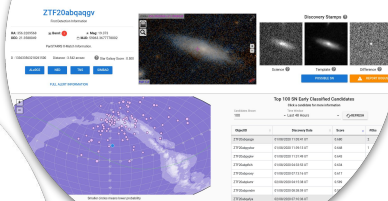
20 million classifications
>4000 TNS candidates
>500 confirmed SNe



SN Hunter

Find recent SNe

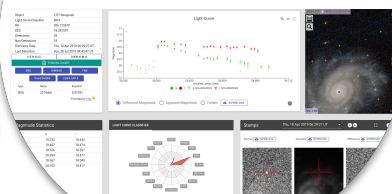
<https://snhunter.alerce.online/>



Explorer

Search and visualize

<https://alerce.online/>

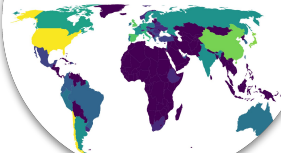


Team

18 astro, 23 stat+ML
6 data engineers

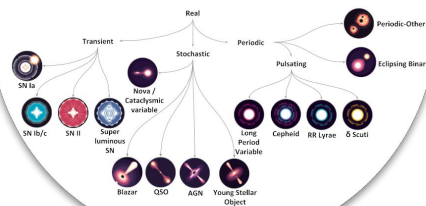


2000 users
50 countries



LC Classifier

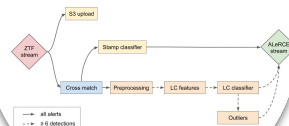
15 class taxonomy
1 million classifications



<http://alerce.science>

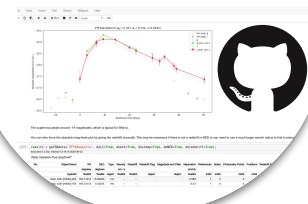
 alercebroker

Pipeline
Modular design



Notebooks

SNe, Variable stars, AGN





ALeRCE's future

- New database & frontend moving from staging to production.
- Xmatch with ZTF DR3. Requires careful calibration.
- Include ZTF Phase II. Forced photometry will enable earlier classification.
- Include ATLAS and other streams. ZTF and ATLAS are complementary!
- Outlier detection in output stream & database.
- Scale infrastructure to Vera Rubin Observatory, xmatch streams in real-time!
- Focus on use cases and user experience.

Thank you ZTF and
congratulations on Phase I!

Links

Homepage: <http://alerce.science>

Explorer: <https://alerce.online/>

New explorer: <http://dev.alerce.online>

SN Hunter: <https://snhunter.alerce.online/>

ALeRCE tutorial: [videos, slides & notebooks.](#)



ALeRCE infrastructure



Feature
computation

Model
training



Distributed
messaging



Distributed
Storage

Services

Container
orchestrator



Distributed
database



Training

Deployment



New: redesigned database, API & client

Database

Detailed description of new schema [here](#)

API

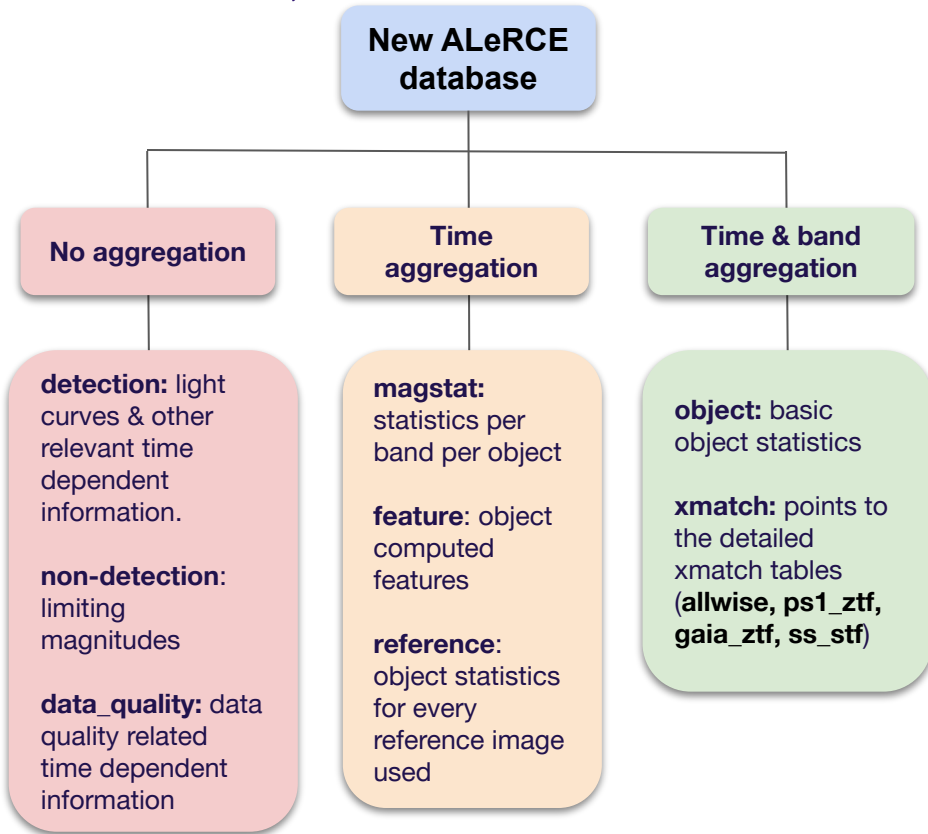
Documentation in <http://dev.api.alerce.online>

Client

Clone new [client](#) and `pip install -e .`

Notebooks

[Supernova](#), [AGN](#) and [variable stars](#)





New: Redesigned ALeRCE explorer

Modular design

Open source

AWS based

New DB connection

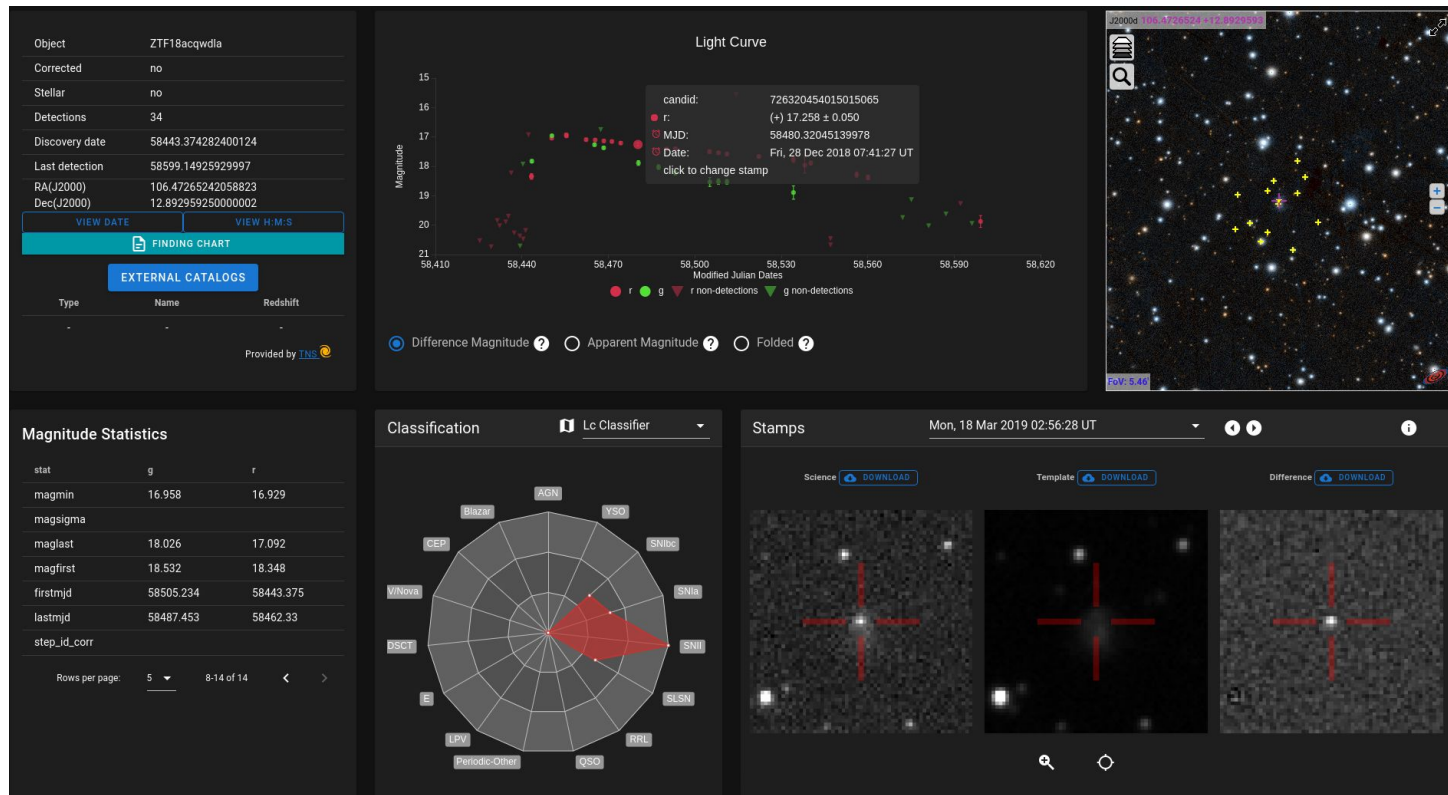
New taxonomy

New classifiers

More statistics

Try the beta version!

<https://dev.alerce.online>





Stamp Classifier vs Light Curve Classifier

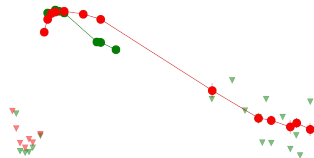
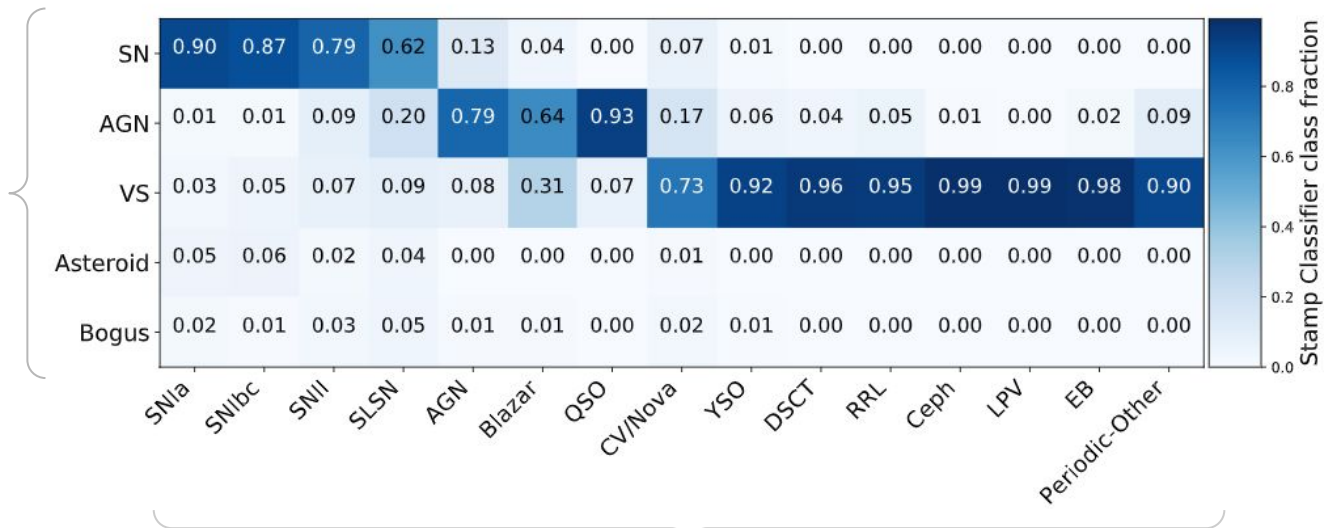
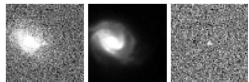
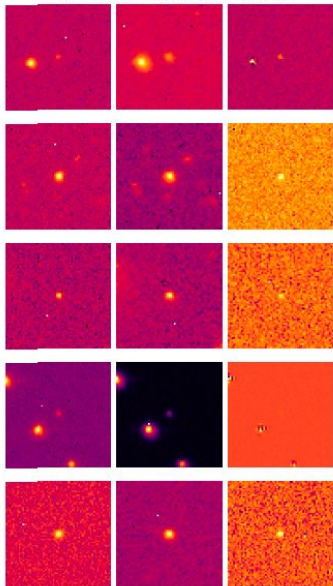


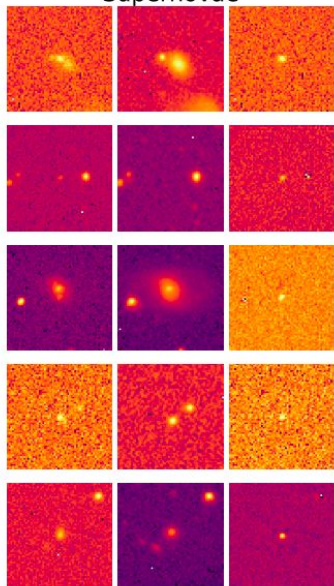


Image stamps

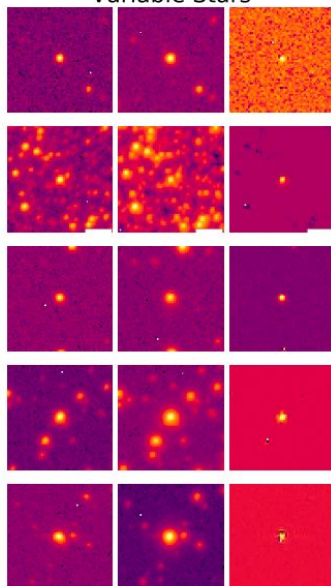
Active Galactic Nuclei



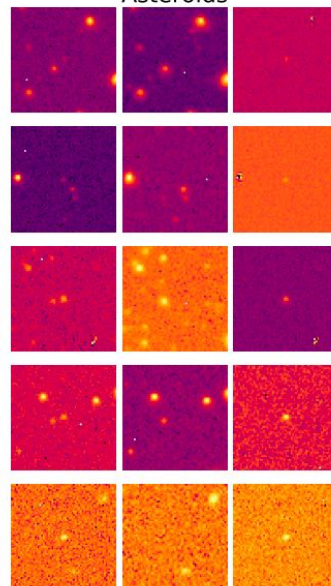
Supernovae



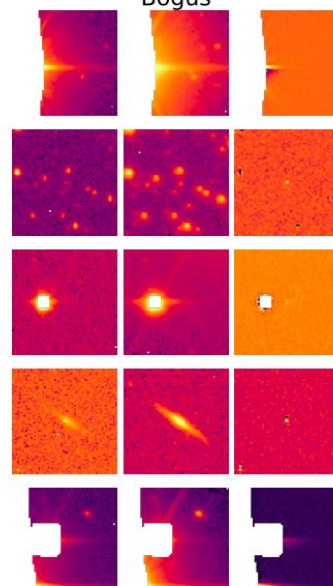
Variable Stars



Asteroids



Bogus

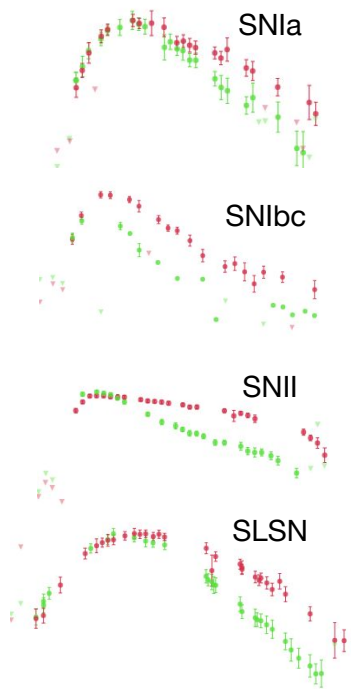




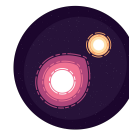
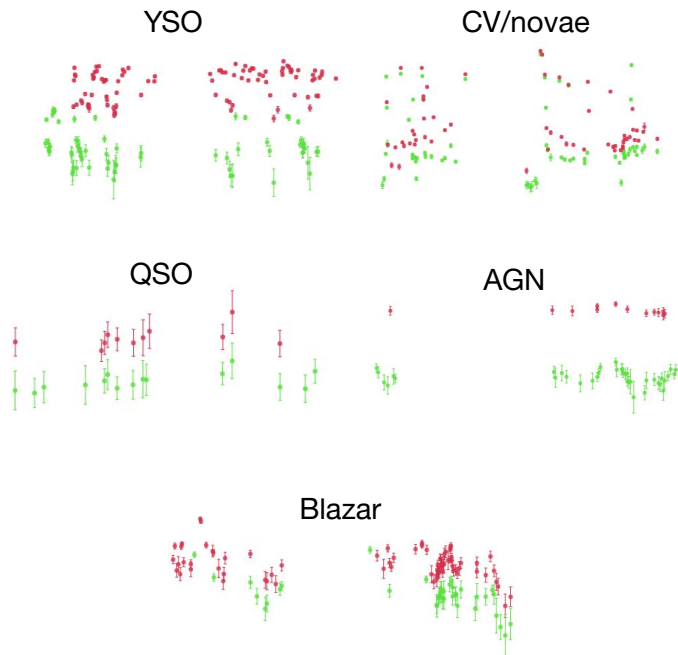
Light Curves



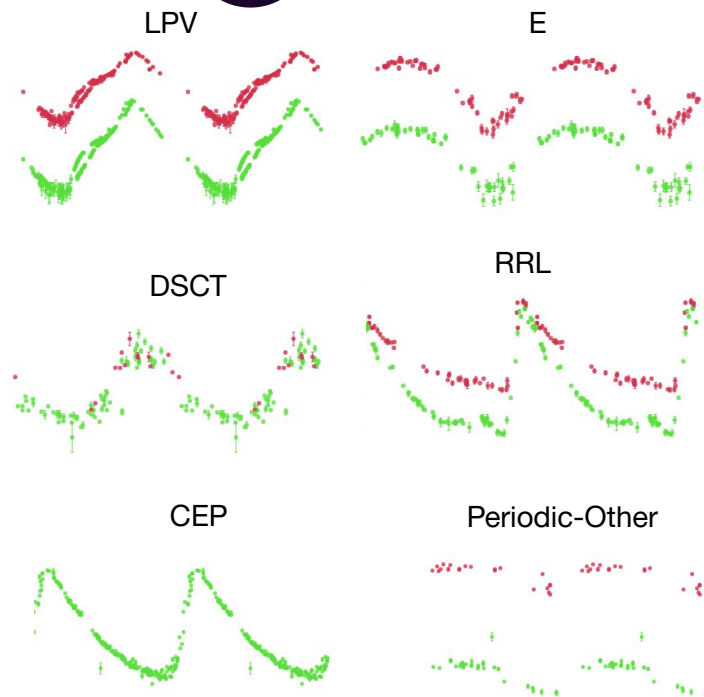
Transient



Stochastic



Periodic
(folded LCs)

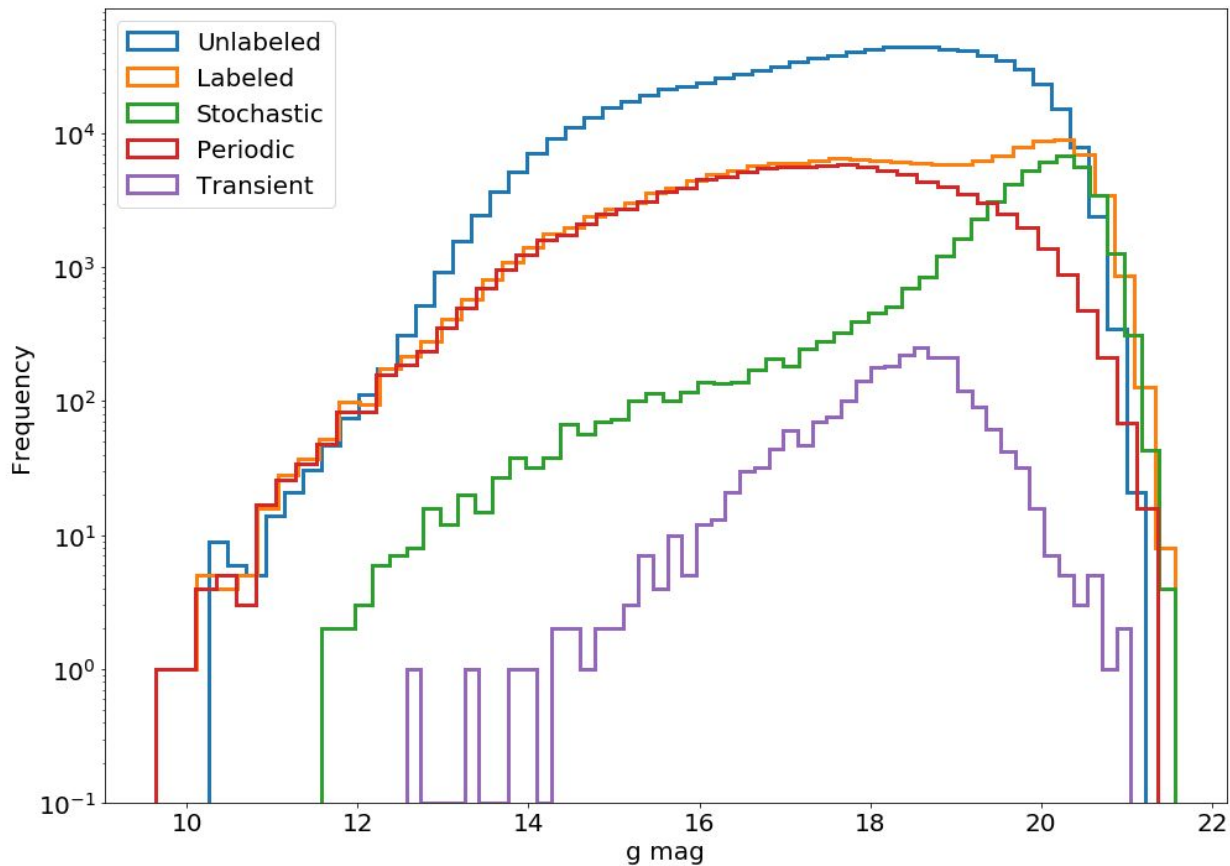


CEP

Periodic-Other



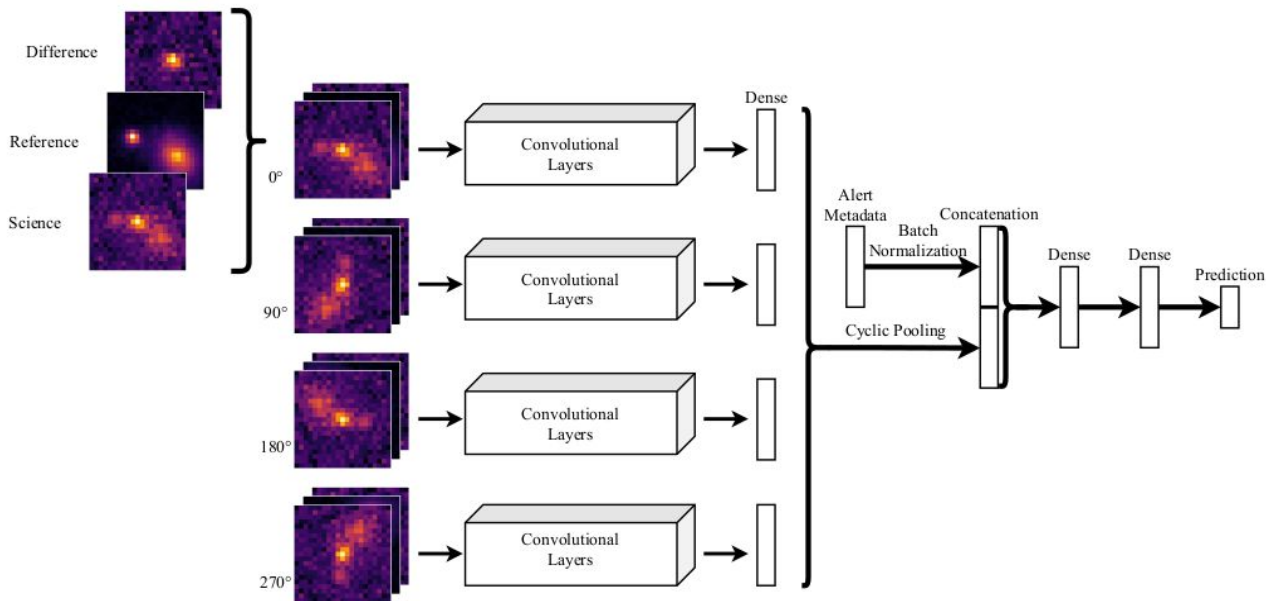
Training set: magnitude distribution



LSST
25th mag!



Stamp based classifier (Carrasco-Davis et al. 2020)

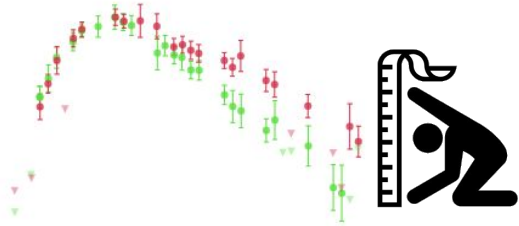


Layer	Layer Parameters	Output Size
Input	-	$21 \times 21 \times 3$
Zero padding	-	$27 \times 27 \times 3$
Rotation augmentation	-	$27 \times 27 \times 3$
Convolution	$4 \times 4, 32$	$24 \times 24 \times 32$
Convolution	$3 \times 3, 32$	$24 \times 24 \times 32$
Max-pooling	2×2 , stride 2	$12 \times 12 \times 32$
Convolution	$3 \times 3, 64$	$12 \times 12 \times 64$
Convolution	$3 \times 3, 64$	$12 \times 12 \times 64$
Convolution	$3 \times 3, 64$	$12 \times 12 \times 64$
Max-pooling	2×2 , stride 2,	$6 \times 6 \times 64$
Flatten	-	2304
Fully connected	2304×64	64
Rotation concatenation	-	4×64
Cyclic pooling	-	64
Concat with BN ^a features	-	$64 + 26$
Fully connected	90×64	64
Fully connected	64×64	64
Output softmax	64×5	5 (n° classes)



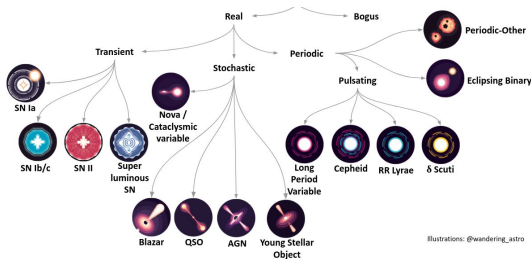
Light Curve based classifier (Sánchez-Sáez et al. 2020)

Characterize



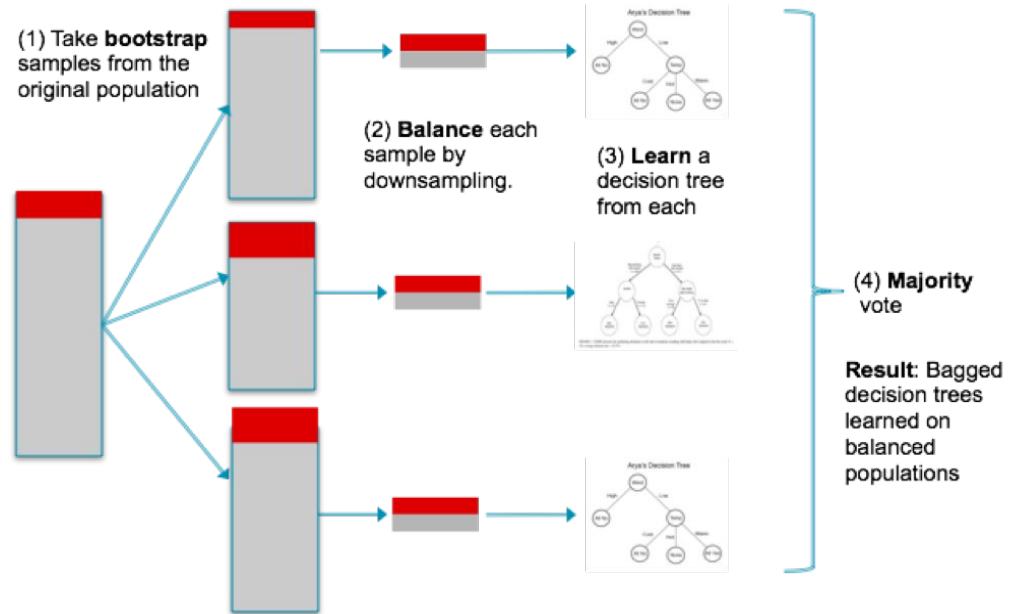
Nun+2015, Eyheramendy+2017/2018, Villar+2019

Define hierarchical structure



Sánchez-Sáez+2020

Train each classifier using bootstrap & balance.

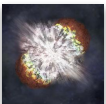
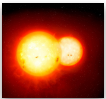


Chen, Chao, Liaw, Breiman (2004)



Jupyter Notebooks

<https://github.com/alercebroker/usecases>

-  API
-  Transients
-  Variable Stars
-  Active Galactic Nuclei
-  Asteroids
-  PanSTARRS
-  Avro inspection