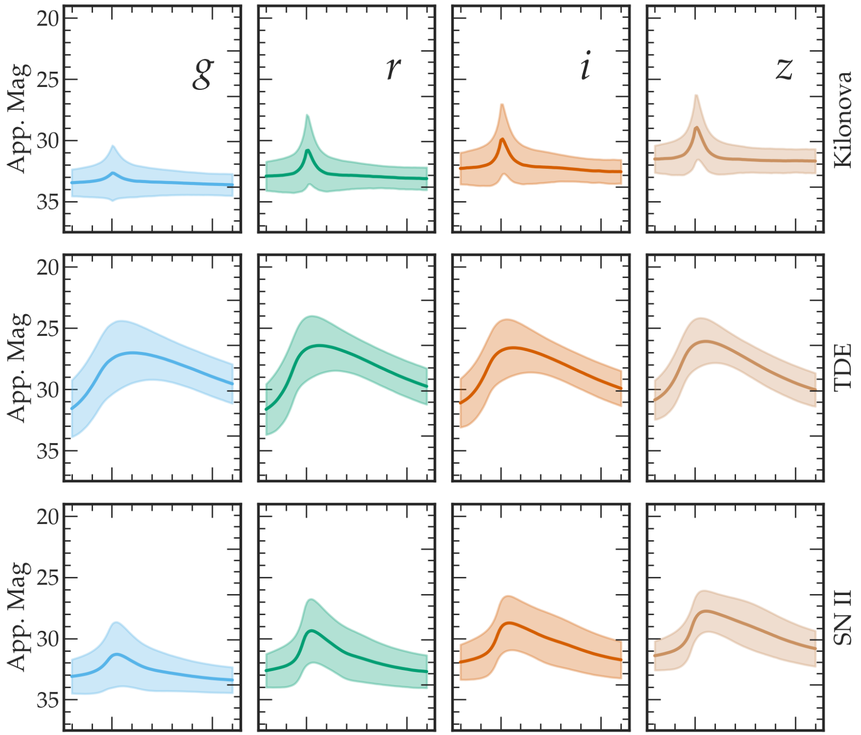
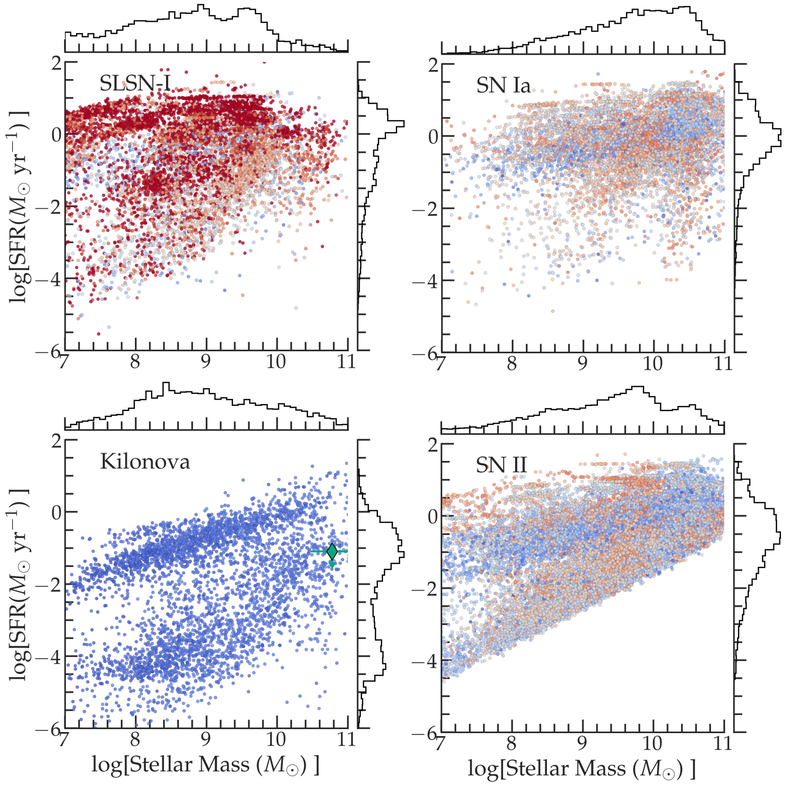
Real-Time Transient Classification With Host-Galaxy Info

Transient Class	NLC
Kilonovae	1,108
SLSN-I	76,686
TDE	54,652
AGN	174,833
SN IIn	267,7674
SNII	41,805,604
SNIb	6,297,024
SNIc	6,297,024
SNIc-BL	67,0240
SNIIb	6,538,537
SNIa	16,193,038
SNIa-91bg	1,079,102
SNIax	3,764,597

Extended ELAsTiCC: 90M Transients
+ Global Host Properties (observed + derived)

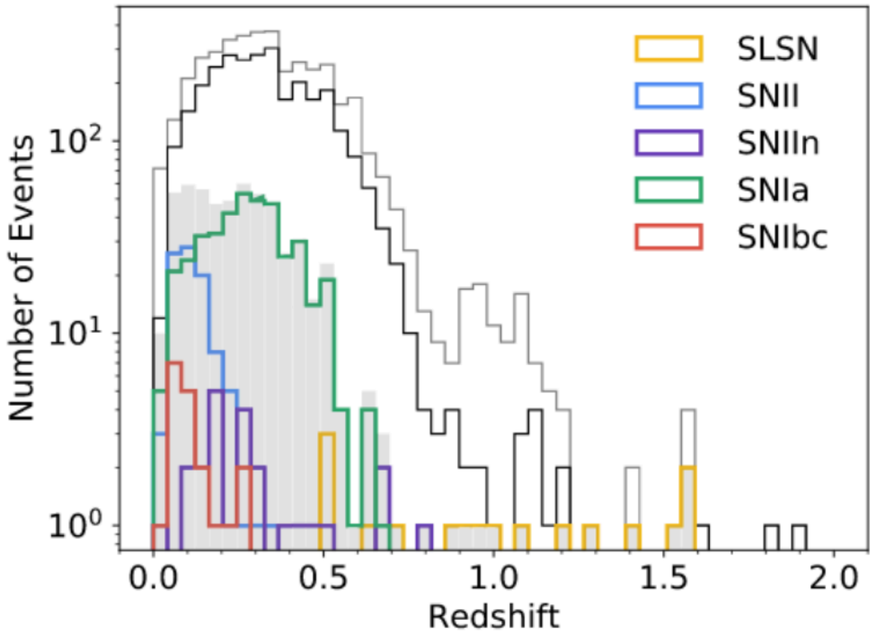


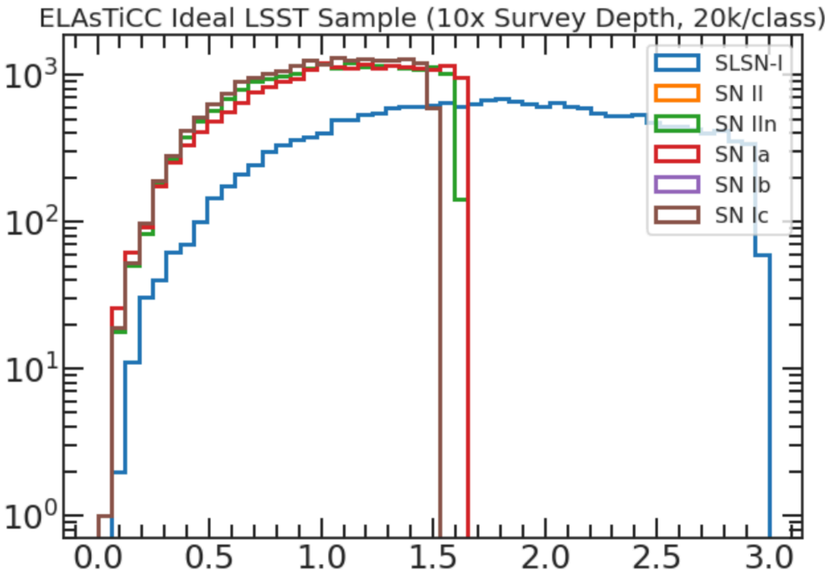


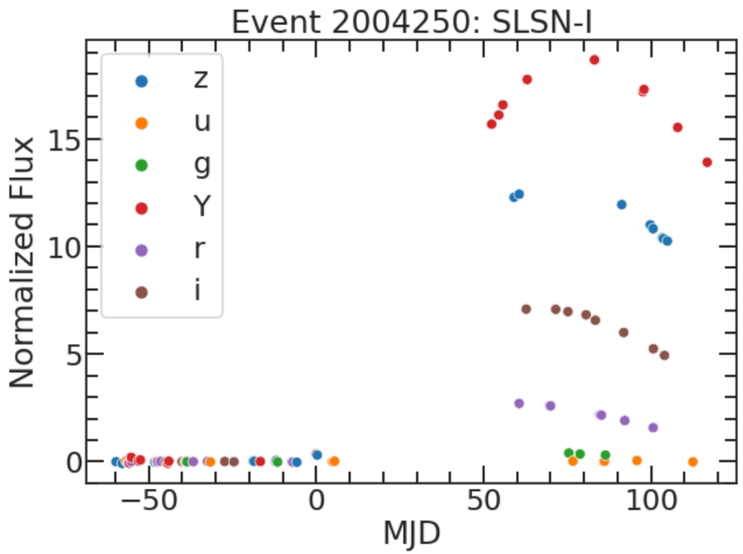
1. Public RAPID (Trained on PLAsTiCC) + Classifier on host information (can be simple)

2. Sim. photometry only
3. Obs. photometry
4. Images

Initial Work Adapting SuperRaenn (Villar+2020)







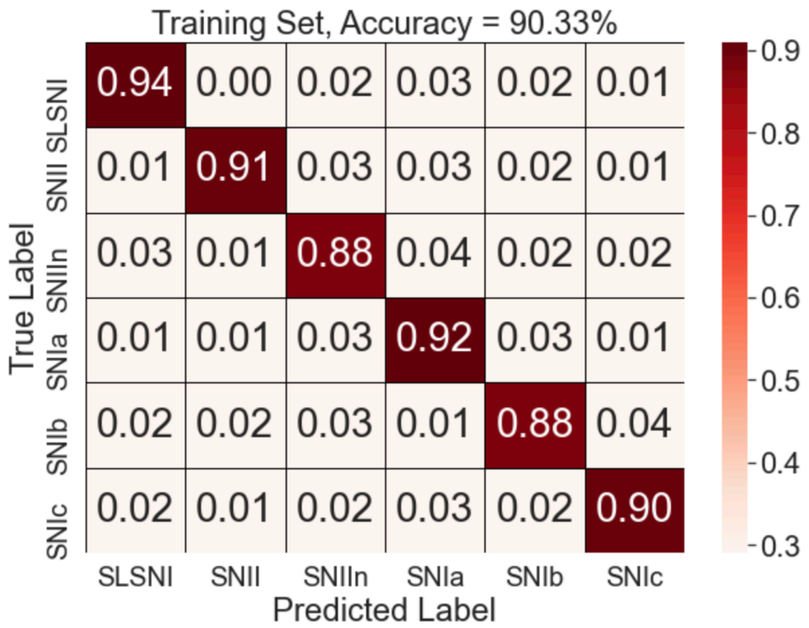
Event Name	SLSNI	SNII	SNIIn	SNIa	SNIb	SNIc	Correct?
1214988	0.0	0.0	0.0	0.0	0.0	1.0	1
2004250	1.0	0.0	0.0	0.0	0.0	0.0	1
1214329	0.0	0.0	0.017	0.0	0.28	0.7	1
1212544	0.0	0.0	0.0	0.0	0.0	1.0	1
1121059	0.0	1.0	0.0	0.0	0.0	0.0	1
1056554	0.0	0.0	1.0	0.0	0.0	0.0	1
1156262	0.0	0.0	0.0	0.0	1.0	0.0	1
1056245	0.0	0.0	1.0	0.0	0.0	0.0	1
1207204	0.28	0.03	0.46	0.1489	0.031	0.0428	0
1155204	0.0	0.0	0.0	0.0	1.0	0.0	1
1172986	0.0	0.0	0.0	1.0	0.0	0.0	1

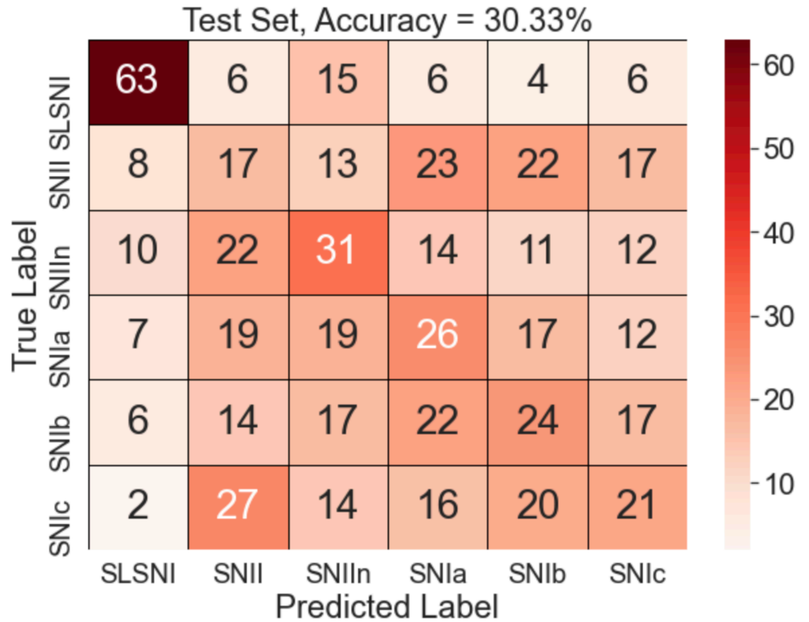
Dataset: LSST Lightcurves in (u)griz(Y)
Classes: SLSNI, SNII, SNIIn, SNIa, (SNIb, SNIc)

Now: Plotting GP and interpreting features

Next steps: Split into labeled/unlabeled, train with full sample

Adapting SuperRaenn (Villar+2020)

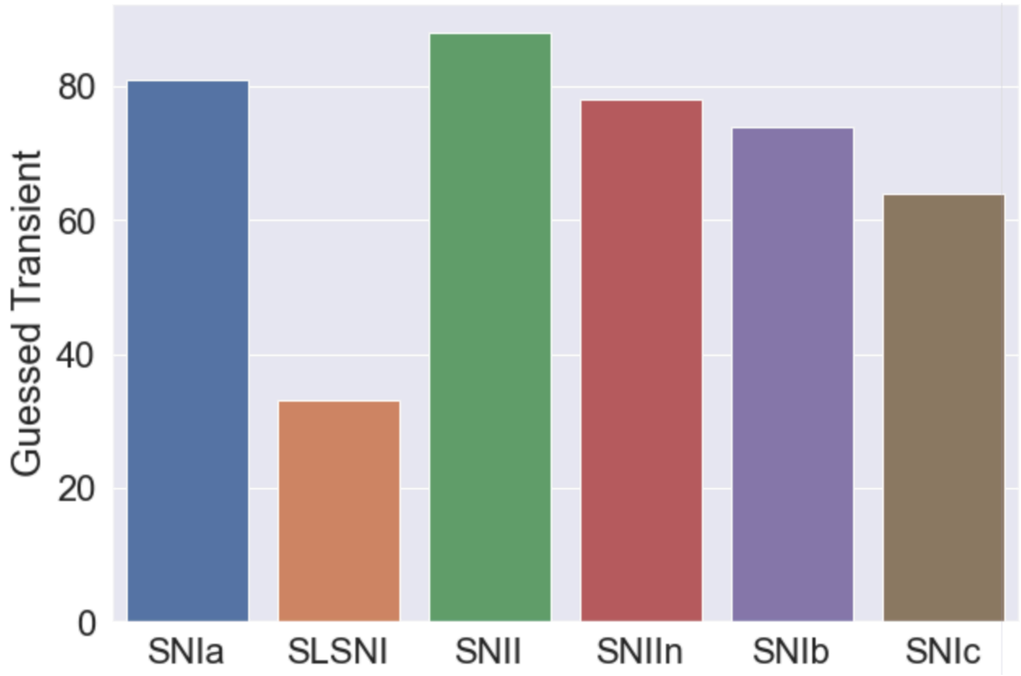




Dataset: LSST Light curves in (u)griz(Y)
Classes: SLSNI, SNII, SNIIn, SNIa, (SNIb, SNIc)

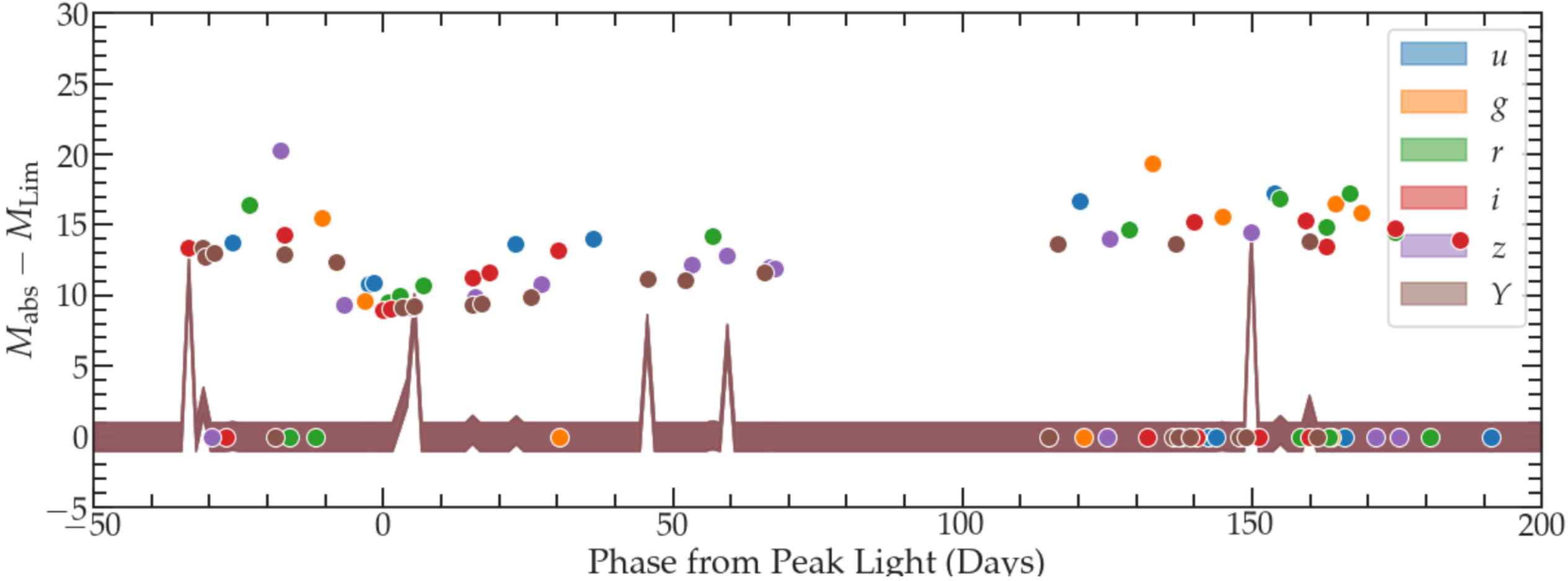
Baseline Test: 500/class, 80/20 train/test, 10 training epochs
2400 in train, 600 in test (100/class)

Resultant model is *heavily* overfit. Will rerun with larger samples (but also need to plot GP fits to LSST photometry)

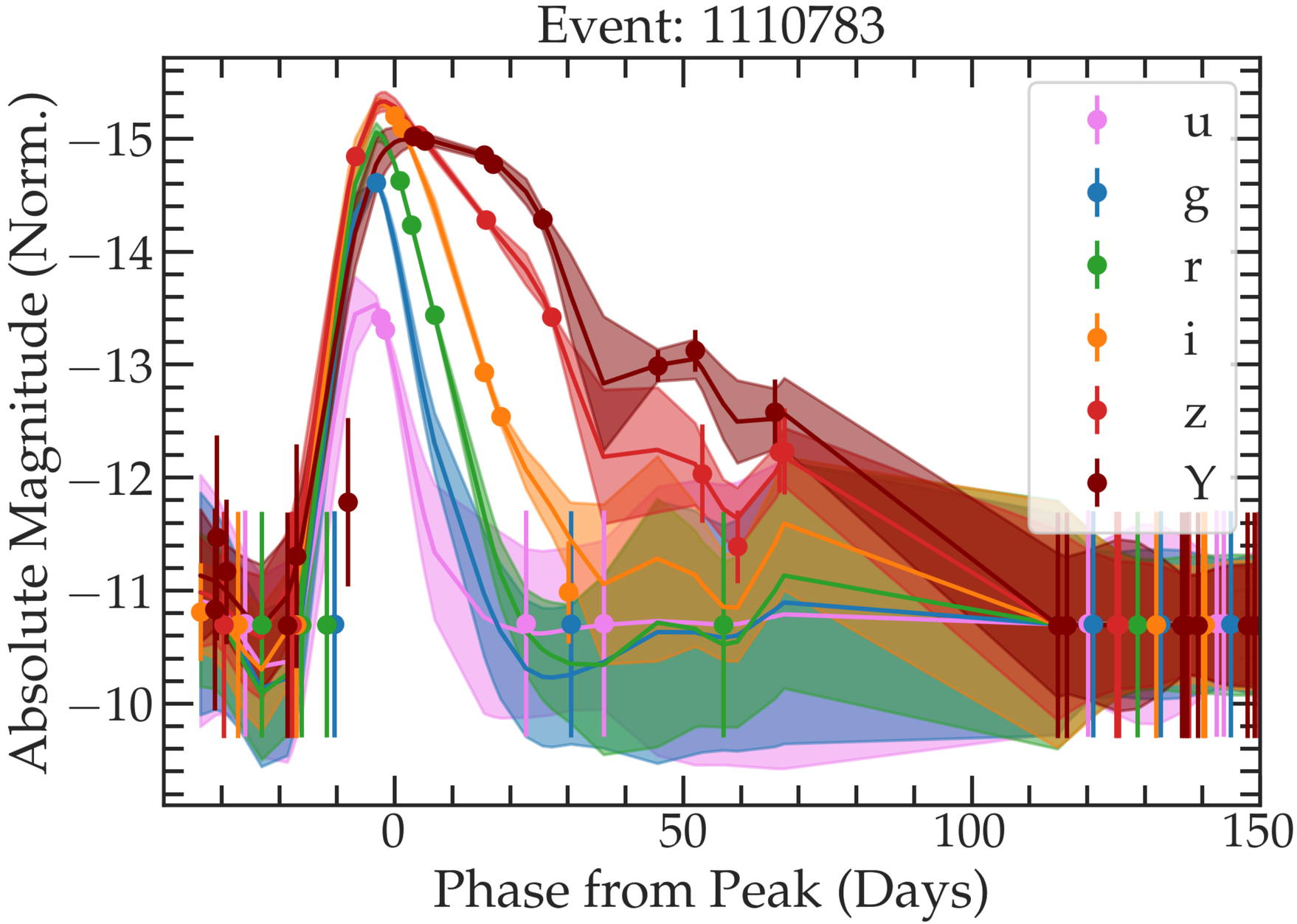


SuperRaenn GP Fits (Villar+2020)

Misjudged the limiting magnitude
(and found a typo in superRAENN)



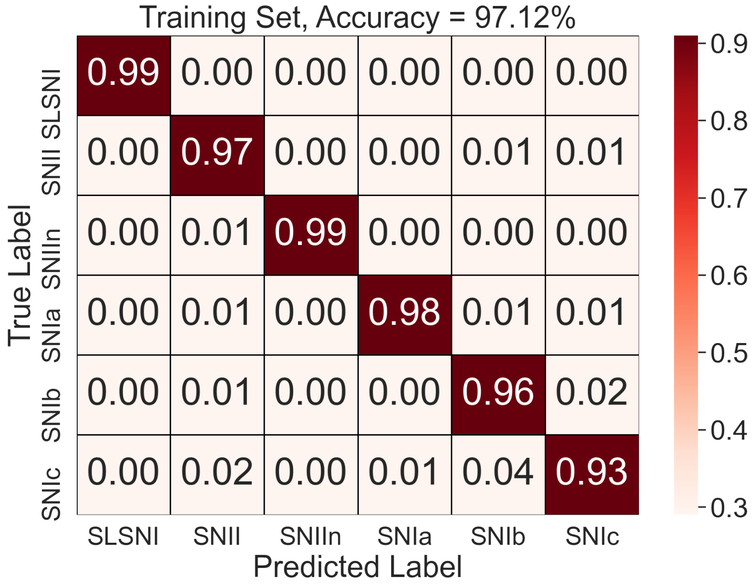
SuperRAENN GP Fits, Take Two



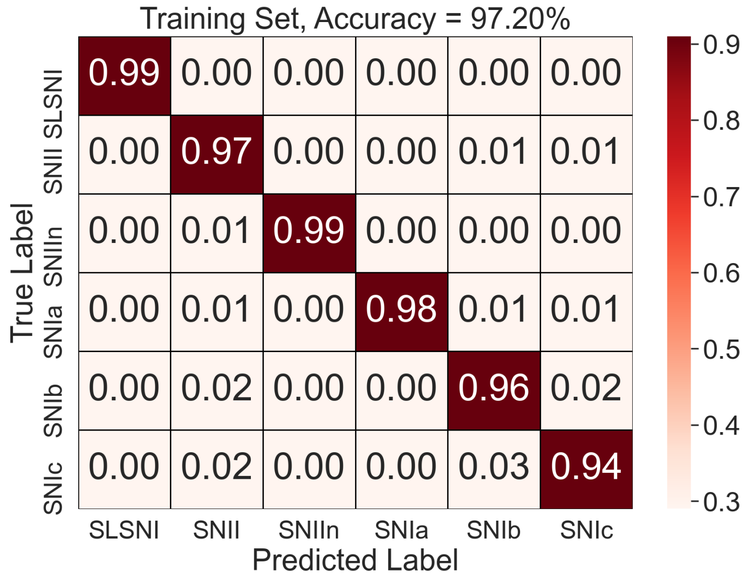
SuperRAENN: A Proof of Concept

5000/class, 80/20 split

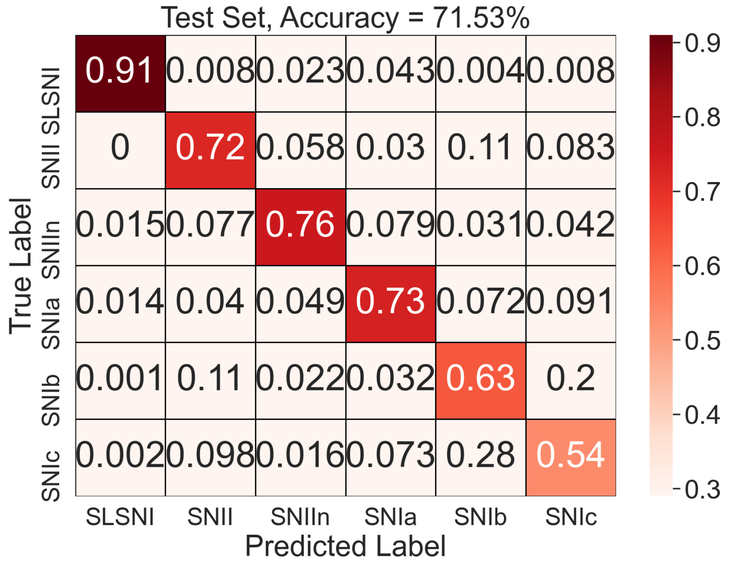
Original



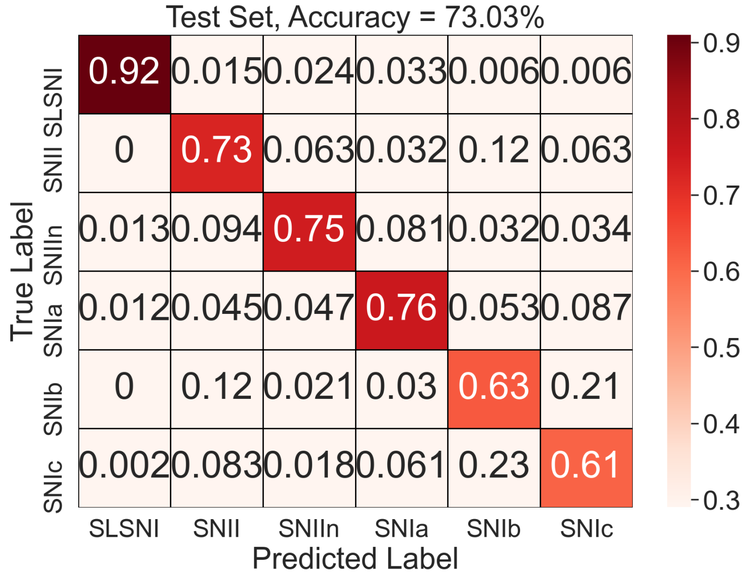
With SFR, Msol



Training



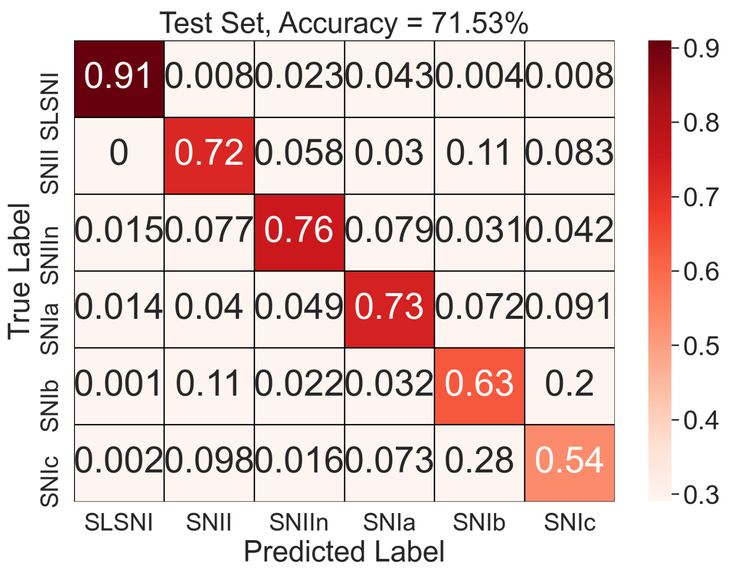
Testing



SuperRAENN: A Proof of Concept

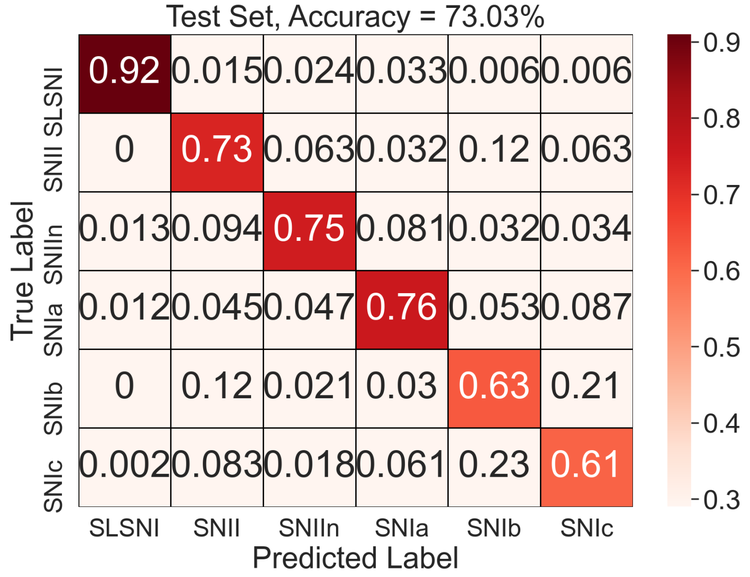
5000/class, 80/20 split

Original



Testing

With SFR, Msol

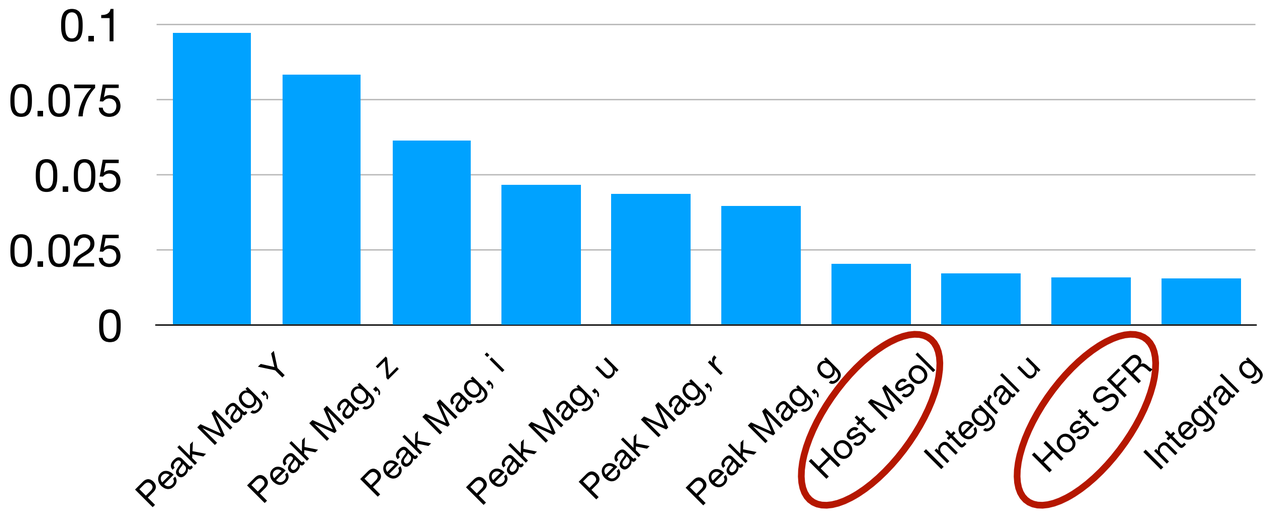


Caveats:

Idealized LCs ($10^5 \times$ LSST exposure)
Simulations (large sample sizes for training + simplified correlations)
Derived host-galaxy information provided (not directly available)
True redshifts provided (not photo-zs)

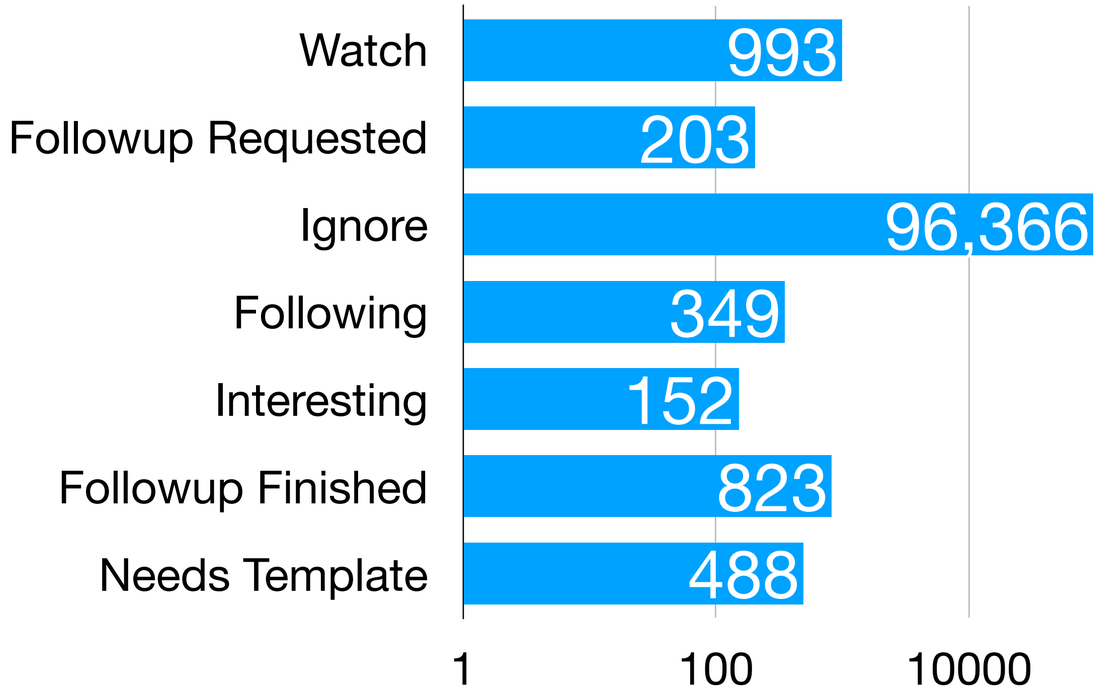
Simulations re-run for true LSST exposure and photo-zs.

Feature Importance Plot



YSE-PZ Transient Stats

Transients Currently on YSE-PZ

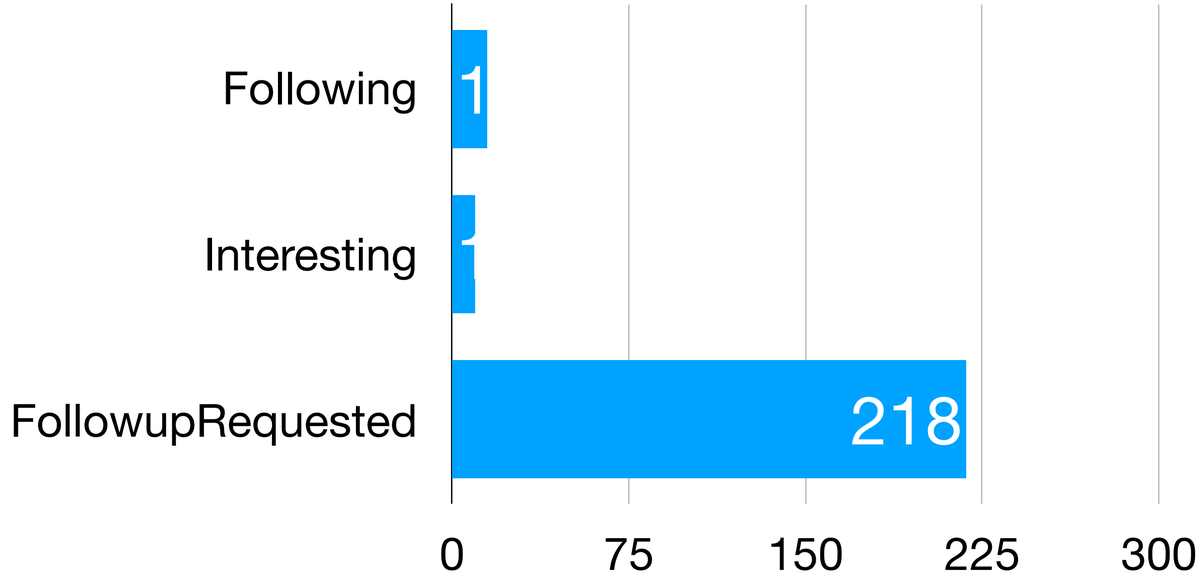


2.5k discovered by YSE

*Includes “uninteresting” and “interesting but not followed”

Transductive (contrastive but for small samples?)
learning to categorize “not followed”?

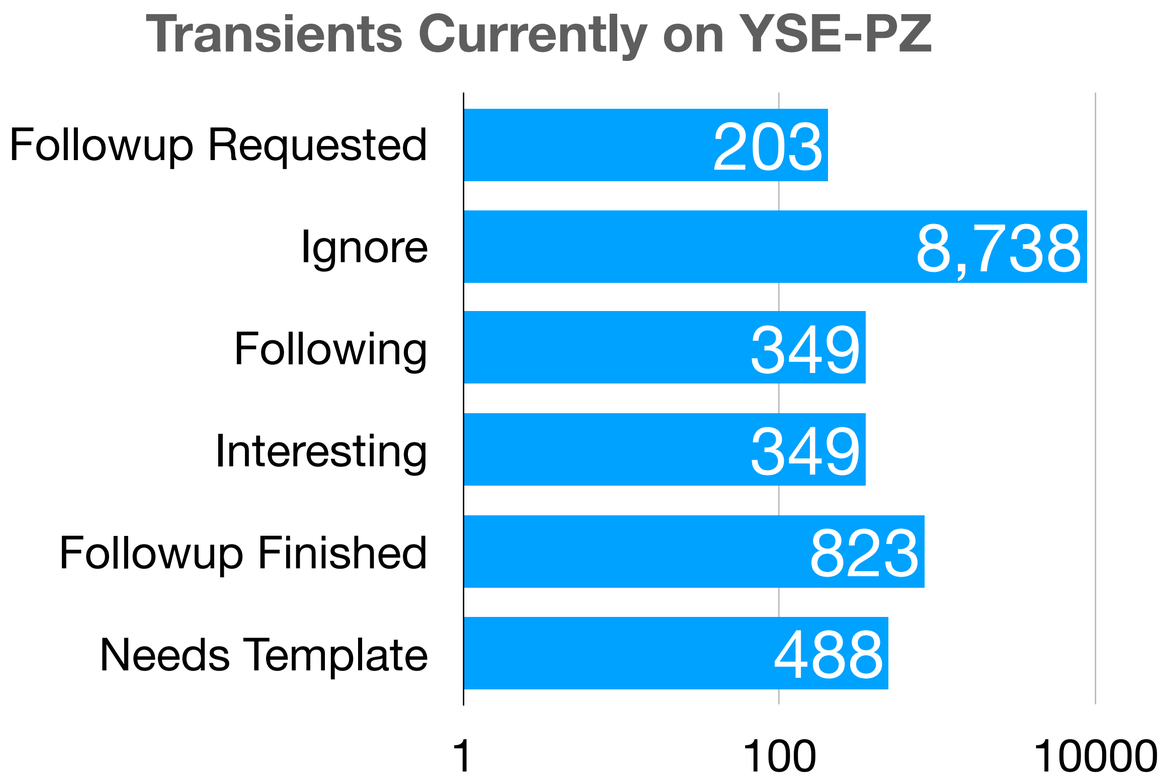
Transients With Time-Stamped Tag Changes



Version Control since 2021-10-22
10k logged status changes
(Mostly Watch —> Ignore, no user info)

For time-series predictions?

YSE-PZ Transient Stats



All photometry and host-galaxy information downloaded from YSE-PZ, extracting light-curve features now with ALeRCE