ELAsTiCC & FINK

Anais Möller, *Julien Peloton, Emille Ishida*
for the Fink collaboration
Special thanks: Marco Leoni, Clécio de Bom

Acknowledging the traditional owners of the land, Wurundjeri People of the Kulin Nation

SWINBURNE UNIVERSITY OF TECHNOLOGY

Two challenges

I. **Infrastructure** Connecting DESC | Rubin | Brokers

Phase I : Repeated 10% sample -> Brokers -> DESC — Now!

Phase II: 100% sample — September/October

II. Classification — September/October

ELAsTiCC & brokers

Rubin Observatory Full-Stream Alert Brokers

- Alerce
- AMPEL
- ANTARES
- BABAMUL
- Fink
- Lasair
- Pitt-Google

Rubin Observatory Downstream Alert Brokers

- SNAPS
- POI: Variables

- A community driven effort, *open to anyone*
- >40 researchers and engineers >7 countries and growing!
- Designed for Rubin with big data technology (e.g. distributed computation)
- Since 2019 processing ZTF II public alert stream available at fink-portal.org

Möller, Peloton, Ishida and the Fink Collaboration 2020

enrich, select, distribute to maximise science

*enrich, select, **distribute** to maximise science*

ELAsTiCC

I. **Infrastructure** Connecting DESC | Rubin | Brokers

Phase I : Repeated 10% sample -> Brokers -> DESC

Status:

✅ Received stream with 10 % sample:

adapted LSST filters, fields, …

✅ Resent stream with “fake” classifications

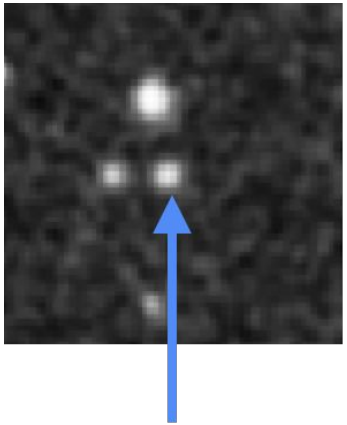

```
{
    "alertId": 31439298006,
    "brokerIngestTimestamp": 1655580540.166,
    "brokerName": "Fink",
    "brokerVersion": "2.1",
    "classifications": [
        {
            "classId": 10,
            "classifierName":
"rf_snia_vs_nonia_1.1.1",
            "classifierParams": "coucou",
            "probability": 0.0
        },
        {
            "classId": 10,
            "classifierName":
"snn_snia_vs_nonia_1.1.1",
            "classifierParams": "coucou",
            "probability":
0.3399367034435272
        },
        {
            "classId": 10,
            "classifierName":
"snn_sn_vs_all_1.1.1",
            "classifierParams": "coucou",
            "probability":
0.028412606567144394
        }
    ],
    "diaSourceId": 31439298006,
    "elasticcPublishTimestamp":
1701653892.48
 }
```

***enrich**, select, distribute to maximise science*

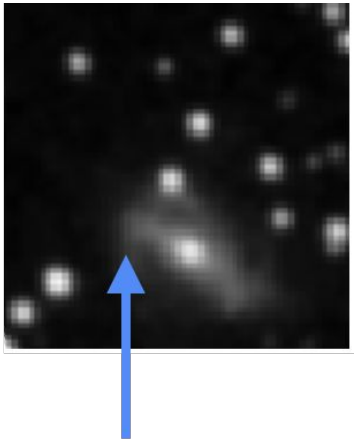
***enrich**, select, distribute to maximise science*

Catalogues

CDS SIMBAD, Gaia DR3,
GCVS, VSX, Mangrove,MPC

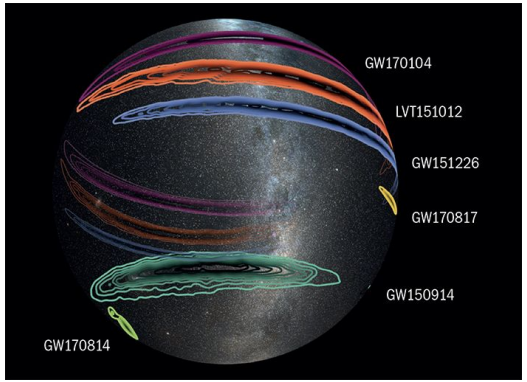


Variable Star



Galaxy close by

VOEvents, GCN, …



features

Color, rates, …

***enrich**, select, distribute to maximise science*

Machine Learning

- Early SNIa: Random Forest (RF) and Active Learning (Ishida+2019, Leoni+2022)
- Supernova: RNNs (Möller+2019)
- Microlensing: RF (Godines, Bachelet+2019)
- Fast transients: RF (Biswas et al. In prep)
- [In progress] Multi-class: Transformers (Allam+2021)
- [In progress] Multi-class: RNN (Burhanudin+2021)
- [In progress] Multi-class: LSTM (Bom+ in prep.)

These algorithms + cuts select candidates! Möller, Peloton, Ishida+2020

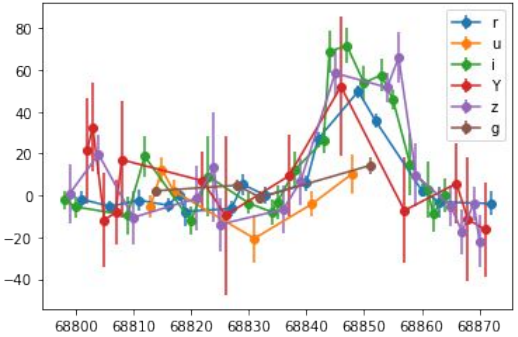
This is different in ELAsTiCC as we can't query catalogues nor have additional info

Machine Learning for ELAsTiCC

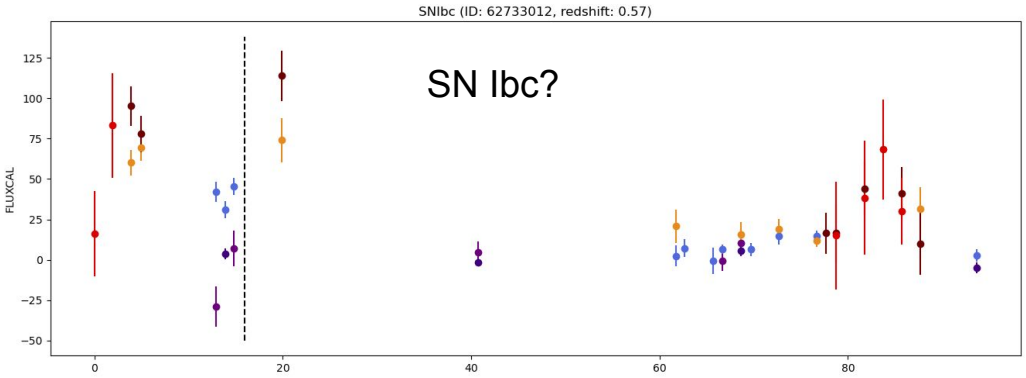
- Size and balance of subclasses affect the potential performance of algorithms when using only training set (no augmentation)
- Taxonomy
- Baselines
- Magnitude limits and detections

```
Alert/Real/Static
├── 0 Static/Other
├── 1 Non-Recurring
│   ├── 10 Non-Recurring/Other
│   ├── 11 SN-like
│   │   ├── 110 SN-like/Other
│   │   ├── 111 Ia
│   │   ├── 112 Ib/c
│   │   ├── 113 II
│   │   ├── 114 Iax
│   │   └── 115 91bg
│   ├── 12 Fast
│   │   ├── 120 Fast/Other
│   │   ├── 121 KN
│   │   ├── 122 M-dwarf Flare
│   │   ├── 123 Dwarf Novae
│   │   └── 124 uLens
│   └── 13 Long
│       ├── 130 Long/Other
│       ├── 131 SLSN
│       ├── 132 TDE
│       ├── 133 ILOT
│       ├── 134 CART
│       └── 135 PISN
└── 2 Recurring
    ├── 20 Recurring/Other
    ├── 21 Periodic
    │   ├── 210 Periodic/Other
    │   ├── 211 Cepheid
    │   ├── 212 RR Lyrae
    │   ├── 213 Delta Scuti
    │   ├── 214 EB
    │   └── 215 LPV/Mira
    └── 22 Non-Periodic
        ├── 220 Non-Periodic/Other
        └── 221 AGN
```

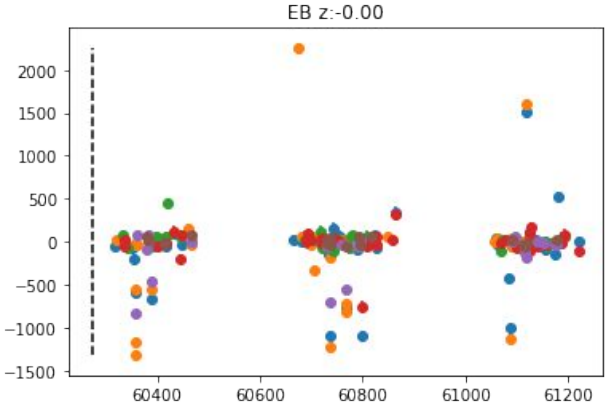
Machine Learning for ELAsTiCC



SN Ia



SN Ibc?



EB

SN Ia!

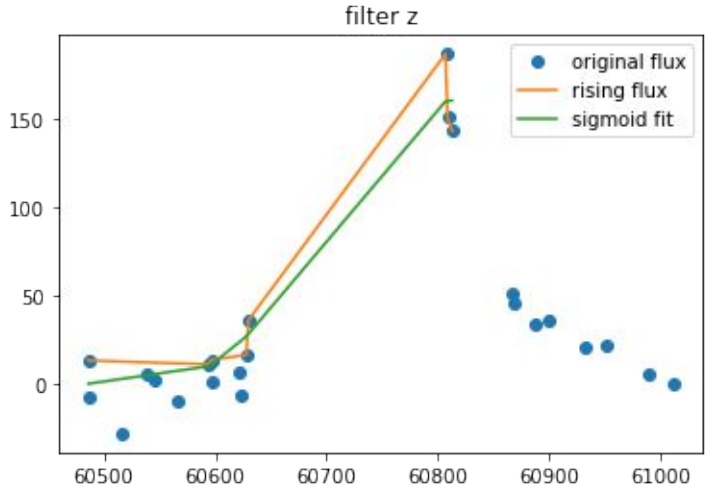
Very very preliminary!

Machine Learning for ELAsTiCC

Early SNIa with Active Learning

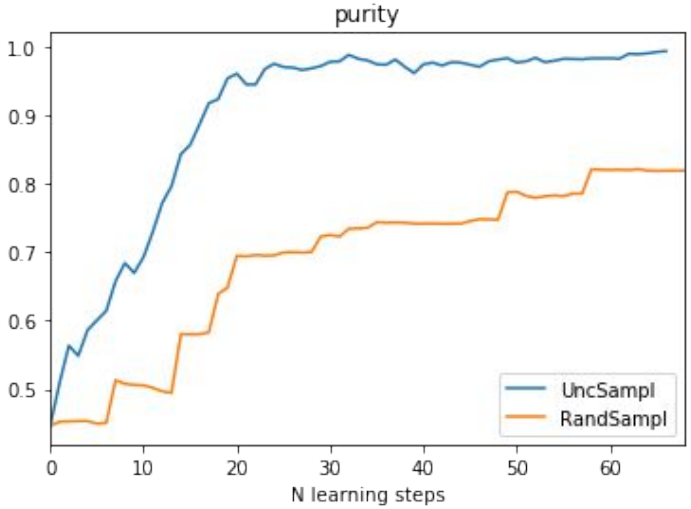
Leoni+2022

Feature extraction



Only rising events and FLUXCAL>100

AL loop



Work by M. Leoni & E. Ishida

Purity : TP/TP+FN

Machine Learning for ELAsTiCC

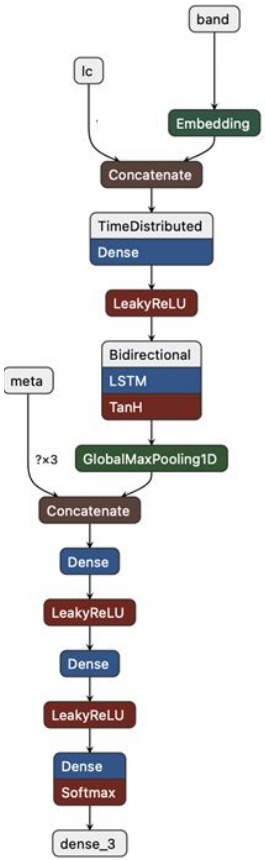
Very very preliminary!

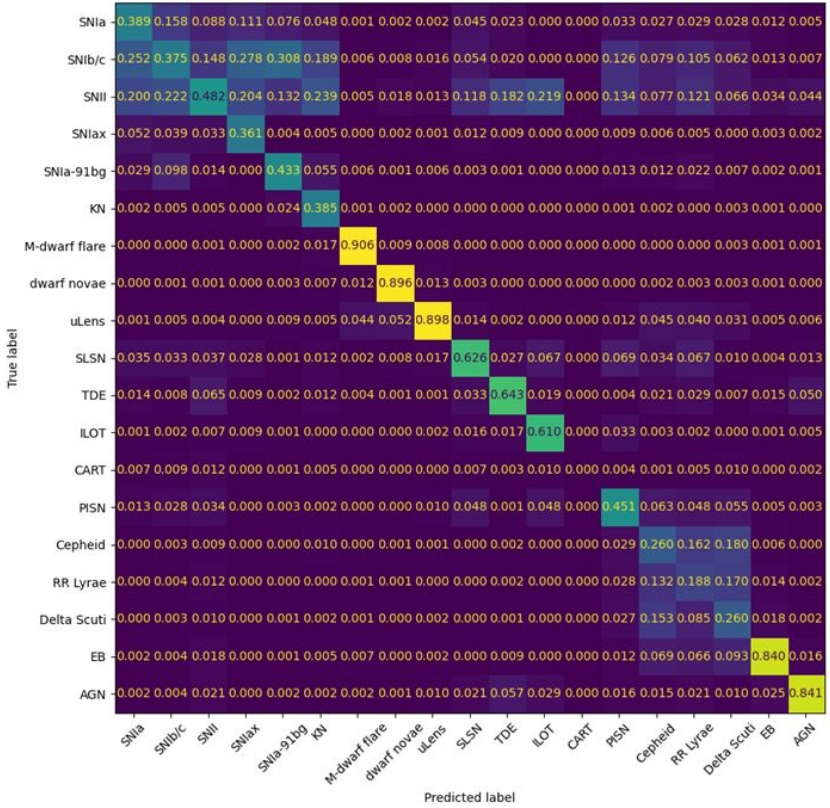
Long Short-term Memory Deep Network

Using only the first alert and forced photometry

Metadata used:

- redshift+error
- host galaxy redshift+error
- MW extinction



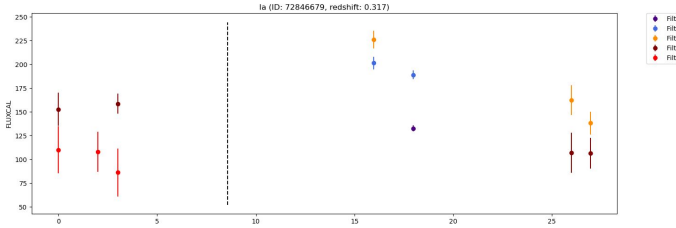


Work by Bom, Fraga et al.

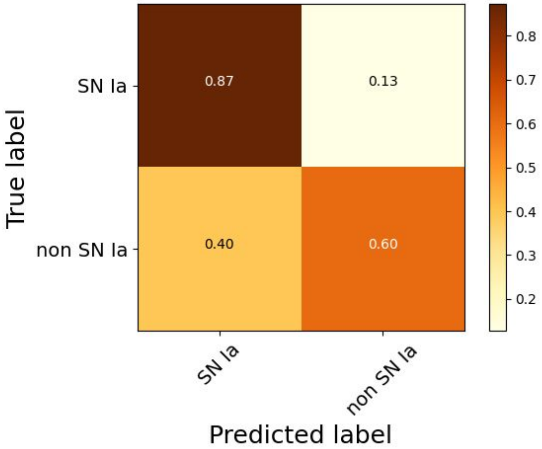
Very very preliminary!

Machine Learning for ELAsTiCC

SNIa vs non-Ia with SuperNNova (RNN) Möller+2019



- ✅ Adapted algorithm to LSST filters/inputs
- 🚧 Training set curation
  - Time window
  - Sampling + magnitude limits



Accuracies SNIa vs non-Ia: ~75% complete-lightcurve, partial are more challenging

Multi-class on the works…

FINK & ELAsTiCC

✅ Infrastructure tests successfully ongoing

🚧 ML phase

❗ Real data uses ML score**s** + selection cuts + catalogues + context

❗ Training set curation is non-trivial

❗ Different ML algorithms -> non-normalized scores…

Other questions….

- Very sparse sampled light-curves, should they be our selection goal?
  - Can we do cosmology with this type of observing strategy?