Supernovae Demography and Rates based on Machine Learning Classification

M.T. Botticella, M. Brescia,S. Cavuoti,A. Mercurio

7 December 2018

Abstract

The scientific goal of this project is the analysis of demography and rates of Supernovae as a function of different galaxy properties as such as mass, colours and specific star formation rate. We aim to exploit the machine learning approach to both automatically classify transients, perform Supernova typing and measure the galaxy photometric redshifts and properties. The activity of the quintuplet and external collaborators in the last year consisted in performing simulations for both Supernova discovery and classification, as well as to participate to the photo-z challenge. We plan also to perform the study on the derivation of galaxy properties, in terms of reliability and accuracy as a function of the expected data.

1 Quintuplet Information 2018

Maria Teresa Botticella researcher INAF
Amata Mercurio researcher INAF
Stefano Cavuoti researcher University Federico II (associated INAF)
Giulia de Somma PhD student

2 Quintuplet Information 2019

Maria Teresa Botticella researcher INAF
Massimo Brescia researcher INAF
Stefano Cavuoti researcher University Federico II (associated INAF)
Amata Mercurio researcher INAF
Fabio Ragosta PhD student (supervisor M.T. Botticella)

Marco Vicedomini additional (officially out of quintuplet, but active in the project) M.Sc. student (supervisors M. Brescia, S. Cavuoti)

1. **Scientific Collaborations:**
   Maria Teresa Botticella TVS
   Amata Mercurio Galaxies (Subgroup Demographics of galaxy populations)
   Stefano Cavuoti DESC, TVS

3 Scientific Activity

Main scientific goal of our project is the analysis of demography and rates of Supernovae (SNe) in host galaxies. Key questions are: which fraction of SNIa events comes from a single star system or by a double degenerative one? Is there any systematic difference between SNIa progenitors in old/young populations? What is the mass interval of SN Core-Collapse progenitors? The activity of the quintuplet and external collaborators in the last year consisted in performing simulations for both SN discovery and classification.

We simulated both SNIa and SN Core-Collapse as they would be observed in DDF survey and in the Main survey for two different time cadences: the baseline 2018a and Altsched one. In order to evaluate the performance of a given survey in achieving our desired science goals, we adopted a direct MAF metric: the number of well-observed SNe and their redshift range. This metric is highly sensitive to changes of observing strategy.

In parallel we have developed a theoretical study on the problem of deriving information on the SNIa progenitors from the observed correlation between the SNIa rate and the parent galaxy properties. We find that (*i*) the currently available surveys do not allow us to draw strong conclusions on this problem, due to both insufficient statistics and systematic effects, most notably related to the difficulty of pinpointing the star formation history in the sample galaxies; (*ii*) the LSST survey will strongly improve the statistical uncertainty, in such a way that we should concentrate the efforts to reduce the systematic uncertainty. (Greggio et al., A&A, submitted).

The application of Machine Learning (ML) and, more in general, of data mining, is mainly motivated by the possibility to perform an optimization of the Parameter Space (PS) characterizing the light curve feature domain, as well as the automatic identification and classification of several Supernova types through different methodologies based on both ML paradigms (supervised and unsupervised learning).

At this preliminary stage we are using two catalogues: (i) the SNPCC (Kessler 2010) catalogue, containing 22,000 Supernovae, with four photometric bands (GRIZ) and divided in three classes, 1 (SN Ia), 2 (SN II) and 3 (SN Ibc); (ii) the Plasticc Challenge catalogue, containing millions of transient objects, like Supernovae, RR Lyrae, AGN etc., divided in fourteen blind classes and with six photometric bands (UGRIZY). On such data we planned to investigate both classifiers (e.g. Random Forest, MLPQNA, Brescia et al. 2015, and fuzzy) and time series predictors (e.g. Long Short Term Memory).

In terms of parameter space, we are analyzing two approaches, statistical and direct (by means of light curves). A statistical PS is obtained from the object light curves, by transforming them into a set of statistical parameters (D'Isanto et al. 2016). The direct approach works directly on the multi-band time sequence of light curves.

The state of art is the following:(i) Fuzzy classifier has failed the light curves classification on SNPCC catalogue; (ii) LSTM has been tested on the SNPCC catalogue, using a variable number of photometric bands. We found that the greater number of bands, the better results; (iii) both Random Forest and MLPQNA algorithms are nowadays running on the Plasticc Challenge catalogue.

4 Scientific and technical deliverables

1. LSST-DESC Photometric Redshift Working Group *An assessment of photometric redshift PDF performance in the context of LSST, in preparation, foreseen to be submitted to MNRAS (co-authors M. Brescia, S. Cavuoti)*

2. Greggio et al., A&A, submitted

5 Other information

We organised the LSST TVS workshop 2018 in Naples (http://eventi.na.astro.it/en/lsst-tvs-2018/) and attended the LSST -Special Programs Workshop in Palermo (MTB as SOC member).
Maria Teresa joined the telecon of TVS group and gave two invited talks on LSST: "Synergies between LSST and SKA: the transient sky" in "The II National Workshop of SKA science and technology" Bologna 3-5 December 2018, and "LSST and SOXS" in the "Primo Congresso Italiano SOXS", Pavia 10-12 December 2018.
Stefano joined the telecons of DESC and the DESC seminars and has been one of the PZ Working Group reviewers for the paper: "*Approximating Photo-z PDFs for Large Surveys*".
M. Brescia gave the talk *Photometric Redshifts for LSST* at the TVS workshop 2018 in Naples, prepared with S. Cavuoti and A. Mercurio.

6 References

Brescia, Cavuoti, Longo, et al. 2015, MNRAS 450 (4), 3893-3903
D'Isanto, Cavuoti, Brescia, et al., 2016, MNRAS 457 (3), 3119-3132
Kessler, Bassett, Belov, et al., 2010, PASP 122 (898), 1415