MEASURING THE METALLICITY OF GALAXIES USING MACHINE LEARNING

Viviana Acquaviva
(CUNY New York City College of Technology)

Prague, June 2017

European Week of Astronomy and Space Science

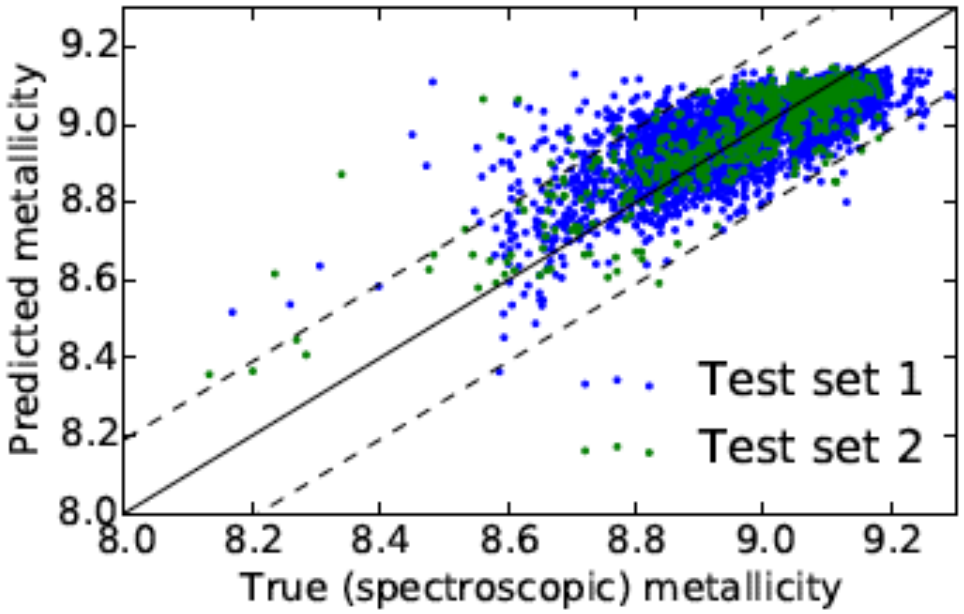
Can we predict Z from photometry?

FEATURES

- SDSS catalog: u, g, r, i, z bands
- also added a couple of derived quantities: stellar mass, photo-zs
- feature "engineering": colors, squared colors

TARGET

- Metallicity determinations from Tremonti et al 04 are used as "ground truth"

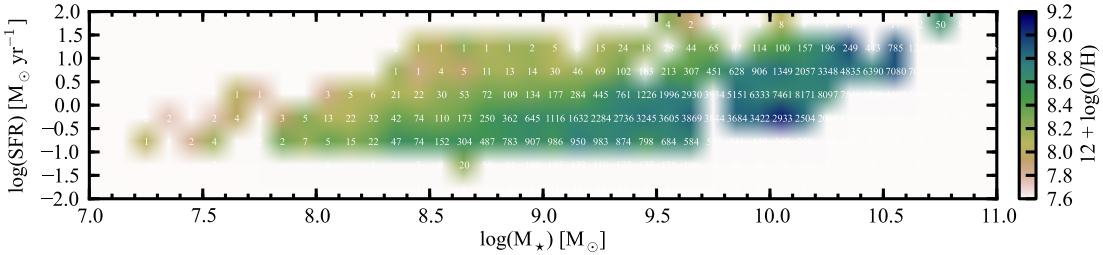


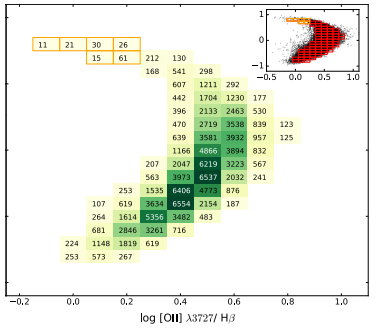
Plot and table from Acquaviva 2016

Data set	Number of objects in training/test set	Average/Median r magnitude	RMSE (all objects)	RMSE (no outliers)	OLF
$0.09 < z < 0.12$	20253/5063	14.2/17.5	0.081 ± 0.001	$0.068 \pm 4e\text{-}4$	0.022 ± 0.002
$0.2 < z < 0.25$	2440/610	18.3/18.3	0.09 ± 0.003	0.069 ± 0.002	0.039 ± 0.007

What next?

Several recent papers have re-examined the relation between metallicity determined via the electron temperature method and strong lines indicators. Auroral lines, which allow one to apply the direct method, cannot be detected often in individual galaxies, so this technique is based on stacking SDSS galaxies with similar line ratios or M, SFR.





Stacks from Andrews and Martini 2013 (above) and Curti et at 2016 (right)

Idea: Combine supervised and unsupervised machine learning to

- Identify an optimal metric to stack spectra of galaxies with similar [O/H] without previous assumptions, either purely from data or from a calibration sample
- Tap the potential in the SDSS data by building smaller stacks (clusters in ML parlance)
- Derive new strong line indicators
- Apply to SDSS + high-z data such as MOSDEF sample
- Test whether the relation between photometric features and metallicity holds