Charlie Hughes

Supervisors: Ting Li and Josh Speagle

Predicting Metallicity using Broadband Photometry and Machine Learning

UNIVERSITY OF TORONTO

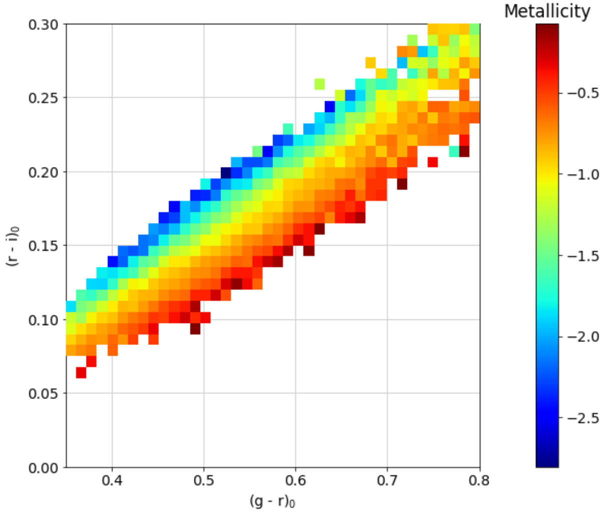


Figure 1: 2d histogram of dereddened (g-r) and (r-i) from DES, color-coded with S5 metallicities. Adapted from Li et al. (2019).

This project uses **machine learning** to generate stellar metallicity predictions using **broadband photometry**.

Context

- Li et al. (2018) first observed a metallicity pattern in the ***griz*** colors in **broadband** photometry (Fig 1).
- Spectroscopy is the most accurate method for determining metallicity, but is limited and resource intensive.
- Photometry advantage: ~1000x more stars. Narrowband photometry is conventionally used, but broadband can observe fainter stars.

Machine Learning

ML has better performance than simple statistics (e.g. polynomial fit) and is highly customizable – one can easily implement and test new data (e.g. Gaia parallax, new color bands). Experimenting with multiple methods, we found random forest regressors and artificial neural networks worked best.

Data Sources

Photometry: *Dark Energy Survey* (DES) and *Dark Energy Camera Legacy Survey* (DECaLS)

Metallicity training set: *Southern Stellar Stream Survey* (S5)

Model Analysis

We carefully evaluated our model's parameter space to ensure it is not a black box!

Data cuts. Our data cuts had three motivations: physical, bad data, or insufficient density of data. Ex: (g-r) > 0.3, because bluer stars → hotter → fewer metallicity lines.

Nearest Neighbour check: Ensure each prediction is in sufficiently dense region of training set parameter space. Reduces error by ~10%.

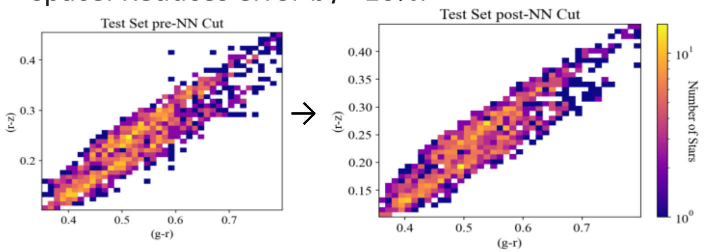


Figure 2: Effect of Nearest Neighbour in color-color space

We used a Monte Carlo technique to determine how photometric uncertainties propagated through our ML model into metallicity uncertainties.

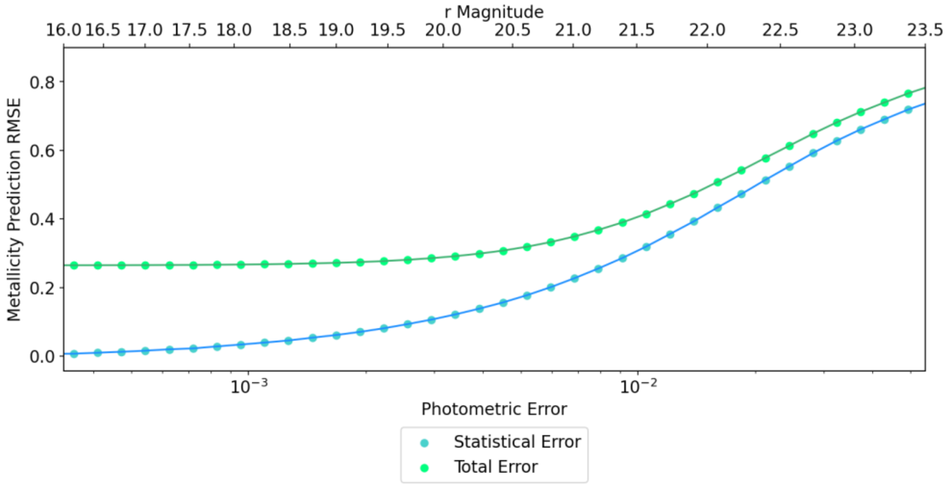
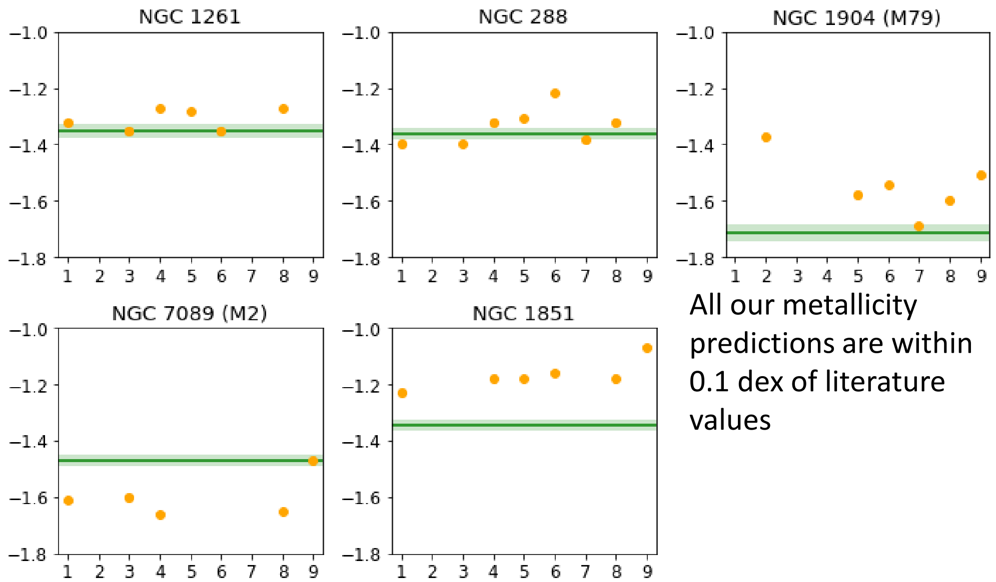


Figure 3: Total and statistical Error in our DES ANN

Model Verification

We compared our isometallicity lines to theoretical ones generated from MIST, using the Brutus library. We found our ML model and the various theoretical iso-curves showed good correspondence, including when broken down into various evolutionary states (MS vs RGB) and the condensing around (g-r) < 0.35 seen in Fig 1 – indicating a limit to the model's parameter space.

As a further test and application, we predicted the metallicity of globular clusters.



All our metallicity predictions are within 0.1 dex of literature values

Figure 4: Mean predicted Fe/H for 5 Globular Clusters from DES compared to literature.

Sources: 1: Marin-Franch et al. (2009). 2: De Angeli et al. (2005). 3: Dotter et al. (2010). 4: VandenBerg et al. (2013). 5 & 6: NGC 1261: Marino et al. (2021), NGC 288 & 1904: Carretta et al. (2009), NGC 1851: Carretta et al. (2011). 7: NGC 288: Costar & Smith (1988), NGC 1904: Geisler et al. (1997). 8: Harris (2010). 9: Pancino et al. (2017).