MNRAS **000**, 1–2 (2018) Preprint 9 February 2018 Compiled using MNRAS LaTeX style file v3.0

TESS Asteroseismic Predictions for Red Giants using Machine Learning

M. Schofield[1,2]*, G. R. Davies[1,2], W. J. Chaplin[1,2], M. F. Randrianandrasana
[1]*Department of Physics and Astronomy, the University of Birmingham, Birmingham B15 2TT, UK*
[2]*Stellar Astrophysics Centre (SAC), Department of Physics and Astronomy, Aarhus University, Ny Munkegade 120, DK-8000 Aarhus C, Denmark*

Last updated 2017 May 31; in original form 2017 May 31

ABSTRACT
Summary: This paper presents a method to predict Red Giant mode detectability with TESS, using the Machine Learning algorithm Classification. It requires only the global parameters $\Delta\nu$, $\nu_{\rm max}$, the stellar magnitude and length of observation.
Method: Lightcurves for *Kepler* stars with fitted radial mode frequencies were used to generate equivalent TESS lightcurves. The lightcurves were cut down, *Kepler* white noise was removed, the bandpass was adjusted, and TESS white noise was added. A detection test was run on the observed modes in these 'TESS-like' lightcurves. Classifiers were then used to predict mode detectability with TESS based upon the global asteroseismic parameters $\nu_{\rm max}$ and $\Delta\nu$, as well as the stellar magnitude.
Application: By changing only the length of dataset and instrumental noise level, this tool can make predictions for any future mission such as K2, PLATO or CoRoT. This is therefore an ideal tool for target selection.

1 INTRODUCTION

Satellites such as *Kepler* have allowed asteroseismology of solar-like and Red Giant stars to advance rapidly since the last century Chaplin & Miglio (2013). Power spectra can now be resolved to detect individual modes of oscillation in Main Sequence and Red Giant stars (Lund et al. (2017) and Davies & Miglio (2016), respectively).

Future space missions such as TESS (Ricker et al. 2014), K2 (Howell et al. 2014), CoRoT (Baglin et al. 2006) and PLATO (Rauer et al. 2014) will add to our understanding of stellar structure and evolution. These missions will provide a large amount of high-precision data. More than ever, the field of stellar astrophysics will require tools to perform big-data analysis (Kremer et al. 2017).

One of the tools than can be used to handle this larger amount of data is Machine Learning. In this work, Machine Learning was used to create a TESS target selection function using the set of *Kepler* Red Giant stars from Davies & Miglio (2016). This publicly-available algorithm can be used to select targets for any future space mission.

In most situations, Machine Learning is used to solve problems in one of two ways; Supervised or Unsupervised Learning. Supervised Learning involves problems where there is a known result.

Supervised Learning has been used to classify types of variable star using previously labelled data (Nun et al. (2014), Elorrieta et al. (2016)). This previously labelled data is known as training data: this is used to train the Machine Learning algorithm. In the problem of variable stars, lightcurves that had already been classified were used to train the algorithm (this is the training dataset). This algorithm was then used to classify the lightcurves of unidentified stars (this is known as the testing dataset).

In Unsupervised Learning, there are no known results (labels). The aim of Machine Learning in this case would be to find trends between variables. This could be used to identify similar stars by analysing their lightcurves without using previously labelled data (for example, Valenzuela & Pichara (2018)).

The aim of this work is to use Machine Learning to make predictions about mode detection probability ($P_{\rm det}$). In this case, $P_{\rm det}$ is a known label, so this is a Supervised Learning problem[1].

Within Supervised Learning, two common algorithms are that are used are Classification and Regression. In Regression, the relationship between variables is interpreted using a measure of uncertainty (such as using χ^2 tests). Models are fitted using the independent data, and uncertainty is measured. The models are then improved by reducing this uncertainty. Note that regression is used when the label is continuous. For example, predicting the magnitude of a star is a problem suited to regression, as a star can have any magnitude.

Conversely, Classification algorithms work by assessing similarity[2]. In Classification, the training set is separated into groups based on the similarity of the data. The more information that was gained by splitting the data, the better. For example, if the problem were to separate Red Giant stars from Main-Sequence stars, a star could be classified as either a Red Giant (1), or not a Red Giant (0). For example, having a Luminosity above $\sim 10L_{\odot}$ would be a strong indicator that the star was a Red Giant, so the data could be separated into groups here. The Classifier would continue to separate the dataset until the Red Giant and Main Sequence samples were distinct. This is not the only problem where Classification can be used on Red Giant stars (Ness et al. (2015), Wu et al. (2017)).

In this work, individual fitted modes from Davies & Miglio

[1] https://machinelearningmastery.com
[2] http://www.simafore.com

(2016) were used to make asteroseismic predictions for TESS. By separating the targets into those with detected modes and those without, Machine Learning was used to select targets for future observation. Section **??** describes how the timeseries of every star was treated before transforming it to a power spectra. Section **??** then goes on to describe the detection test that was run on every fitted mode. Every mode was grouped into observed (1), or unobserved (0).

Lastly, Section **??** describes the classification of stars into a group with detected modes, and a group without. This was done by giving a Supervised Classifier information on every target: the frequency of maximum solar-like amplitude $\nu_{\rm max}$, the large separation between modes of the same angular degree $\Delta\nu$, the stellar magnitude, and which modes were detected inside the star. 70%???? of the stars were used to train the Classifier; 30%????? of the sample was kept to test the algorithm. The Classifier recognised patterns between the variables in the training set, and made predictions on the testing set with a ?????????% accuracy.

REFERENCES

Baglin A., et al., 2006. p. 3749, http://adsabs.harvard.edu/abs/2006cosp...36.3749B

Chaplin W. J., Miglio A., 2013, Annual Review of Astronomy and Astrophysics, 51, 353

Davies G. R., Miglio A., 2016, arXiv:1601.02802 [astro-ph]

Elorrieta F., et al., 2016, Astronomy and Astrophysics, 595, A82

Howell S. B., et al., 2014, Publications of the Astronomical Society of the Pacific, 126, 398

Kremer J., Stensbo-Smidt K., Gieseke F., Steenstrup Pedersen K., Igel C., 2017, preprint, 1704, arXiv:1704.04650

Lund M. N., et al., 2017, The Astrophysical Journal, 835, 172

Ness M., Hogg D. W., Rix H.-W., Ho A. Y. Q., Zasowski G., 2015, The Astrophysical Journal, 808, 16

Nun I., Pichara K., Protopapas P., Kim D.-W., 2014, The Astrophysical Journal, 793, 23

Rauer H., et al., 2014, Experimental Astronomy, 38, 249

Ricker G. R., et al., 2014, Journal of Astronomical Telescopes, Instruments, and Systems, 1, 014003

Valenzuela L., Pichara K., 2018, Monthly Notices of the Royal Astronomical Society, 474, 3259

Wu Y., et al., 2017, preprint, 1712, arXiv:1712.09779

This paper has been typeset from a TEX/LATEX file prepared by the author.