Contribution ID: **7**   Type: **"Classic" talk**

Supervised classification of variable stars from the NASA TESS survey with features originating from the biomedical domain

Thursday 27 May 2021 11:15 (20 minutes)

The currently ongoing NASA TESS space mission is expected to observe tens of millions of stars. The resulting stellar surface brightness measurement time series ("light curves") allow astronomers to search for specific types of stars or planets, as well as to then infer their fundamental physical parameters. Given that we can observe different types of light curves for different types of stars, the first step is to classify the light curves according to their underlying variability type. As we are working with vast amounts of data, it is infeasible to manually classify all observations and we therefore require automated techniques.

Hence, we developed a classification method based on a Random Forest classifier that can successfully classify stars according to their variability type. In order to find the ideal feature sets to characterize the different types of stellar variability, we turned to the biomedical literature on EEG signal processing as these signals share some common characteristics with stellar variability signals. We specifically turned to the field of entropy analysis, from which we then adopted the multiscale entropy from Costa et al. (2005) to characterize the complexity and uncertainty present in stellar variability signals. We used this to complement our more traditional Fourier and statistical feature sets, and discovered that the entropy metrics proved to be important features in our classifier due to their ability to differentiate light curves based on their unpredictability and complexity levels.

We then incorporated our classifier into the larger TESS Data for Asteroseismology (T'DA) classification pipeline to obtain the best results. In the pipeline we first train multiple distinct classifiers with different feature sets on the same data and then pass their results (the class probabilities) on to a meta-classifier that combines the predictions from this ensemble of models and returns a final classification. The benefit of this approach is that the metaclassifier accounts for the strengths and weaknesses of each of the classifiers and in this way returns an optimal classification. We validated our method on data from the previous NASA Kepler mission, given that we already had labelled datasets available here.

Author: AUDENAERT, Jeroen (KU Leuven, Institute of Astronomy)

Co-authors: Dr TKACHENKO, Andrew (KU Leuven, Institute of Astronomy); Prof. AERTS, Conny (KU Leuven, Institute of Astronomy)

Presenter: AUDENAERT, Jeroen (KU Leuven, Institute of Astronomy)

Session Classification: Morning 1

Track Classification: Surveys