EPSC Abstracts
Vol. 14, EPSC2020-833, 2020
https://doi.org/10.5194/epsc2020-833
Europlanet Science Congress 2020

Classifying Exoplanets with Machine Learning

Ana Barboza[1], Solène Ulmer-Moll[1,2], and João Faria[1,2]
[1]Department of Physics and Astronomy, University of Porto, Porto, Portugal
[2]Instituto de Astrofísica e Ciências do Espaço, Universidade do Porto, CAUP Rua das Estrelas, 4150-762 Porto, Portugal

More than 4200 exoplanets have been detected and their diversity is remarkable, ranging from very small rocky planets, to puffed gas giants. Several of their types are unknown in our Solar System, hence new classes have been defined to understand this diversity and the similarities within each group, such as their formation mechanism or core composition.

We aimed to determine the main types of exoplanets, develop a method that automatically associates exoplanets to their type, classifying them into labels with a machine learning algorithm. We also worked to further understand each group, analysing their characteristics, and exploring correlations within each group.

Given the planetary mass and orbital period of a large number of exoplanets, we used a K-Means clustering algorithm to classify three large groups: Hot Jupiters, Long Period Giants and Small Planets. In order to take into account more planetary and stellar parameters, we also work with the Uniform Manifold Approximation and Projection (UMAP) technique to visualize data on a 2D map, aiming to find structures within the high dimensional parameter space. We identified different clusters of exoplanets on this map with the help of groups already described in the literature. We explored how different sets of input parameters impact the clustering of exoplanets and studied, in particular, the effect of stellar metallicity.

We were able to identify 5 different groups: Hot Jupiters, Longer Period Giants, sub-Jupiters, sub-Neptunes and Rocky Planets.

We described these groups in terms of values for each parameter, and discussed outliers. We also analysed metallicity separately and verified that, on average, giant planets orbit around higher metallicity stars than non giant planets.

The well known groups of giant exoplanets, such as Hot Jupiters and Longer Period giants, are clearly identified in the resulting UMAP 2D parameter space. For smaller planets, several groups were also visible but less separated. We also verified that the global structure is preserved, noticing, for example, the smaller planets ($< 8R_{\square}$) are grouped together and well separated from the Hot Jupiters. Adding more samples of well characterized small planets would certainly help their classification.