



Contribution ID: 117

Type: **Poster Presentation**

Identification of ultracool dwarfs in J-PLUS DR2 using Virtual Observatory tools and Machine Learning techniques

Monday, 30 May 2022 14:59 (3 minutes)

The Javalambre Photometric Local Universe Survey Data Release 2 (J-PLUS DR2) covers $2\,176 \text{ deg}^2$ using a unique filter system of 12 optical bands. This large coverage of the electromagnetic spectrum allows a more accurate determination of physical parameters such as, for instance, effective temperatures.

Current surveys like J-PLUS, and others to come in the near future, are causing a data avalanche in Astronomy. In this scenario, the Virtual Observatory makes the difference in what refers to the discovery, access and analysis of scientific data. Moreover, the huge volume of information generated by these surveys goes beyond what traditional processing and analysis methods can offer. To face this situation, machine learning (ML) approaches have gained momentum over the last few years offering a suite of alternatives depending on the proposed case.

We present the search for ultracool dwarfs (UCDs, spectral types later than M7) performed across the entire J-PLUS DR2 data set. For this purpose, we apply a methodology driven by the use of multiple VO tools and services that combines J-PLUS data with astrometric information from Gaia EDR3. Furthermore, we explore the ability to reproduce this search with a purely ML-based methodology that relies solely on J-PLUS optical photometry, with a two-step ML method based on Principal Component Analysis (PCA) and Support Vector Machine (SVM) algorithms.

Our methodology starts with a pre-screening process in which we use three different approaches to obtain a shortlist of candidates. Two of these approaches rely on astrometric constraints to preselect the candidates, while the third uses only J-PLUS photometry. Finally, we use VOSA, a Virtual Observatory tool which fits observational data to different collections of theoretical models, to estimate the effective temperature of the candidates and keep only those with $T_{eff} > 3\,000 \text{ K}$.

After this process, we ended up with $9\,811$ candidate UCDs across the entire sky coverage of J-PLUS DR2. We conducted an in-depth analysis of the kinematics and binarity of these candidates. Also, we developed a Python algorithm to detect flares on $H\alpha$ and Ca II H and K emission lines, using only J-PLUS photometry, detecting 8 objects with relevant emission peaks in these lines.

When reproducing this search with the ML-based methodology, we were able to remove the hottest objects ($T_{eff} > 4\,100 \text{ K}$) in the PCA step, using as variables multiple J-PLUS colours. Then, we trained a grid of classification SVMs using as labels the candidate UCDs obtained with the previous methodology. The best recall score (98%) was obtained with a radial basis function (RBF) kernel and hyperparameters $C = 1000$ and $\gamma = 0.001$. Using this model to predict on unseen data, we were able to recover 96% of the candidate UCDs. In contrast with the VO methodology, we deduced that the ML methodology is more efficient in the sense that it allows a greater number of true negatives to be discarded prior to analysis with VOSA, although it is a more restrictive method as it requires objects with good photometry in all the J-PLUS filters used to build the variables.

Main Topic

Classification and regression

Secondary Topic

Supervised/Unsupervised/Semi-supervised Learning

Participation mode

Remote

Primary authors: Dr GONZÁLEZ MARCOS, Ana (Universidad de La Rioja); Dr SOLANO MÁRQUEZ, Enrique (Centro de Astrobiología (INTA-CSIC)); MAS BUITRAGO, Pedro (Centro de Astrobiología (INTA-CSIC))

Presenter: MAS BUITRAGO, Pedro (Centro de Astrobiología (INTA-CSIC))

Session Classification: Poster Session Day 1