Machine Learning for Stellar SEDs

SUMMARY.

Current space missions, benefiting from technological progresses, have raised the amount of collected data to unprecedented levels, and have highlighted a critical need for the development of new tools capable of processing and analysing big data from space. Artificial Intelligence (AI) is a powerful tool that is becoming more common across a wide range of fields, including astronomy and Earth observations. In this module, the student will use AI techniques for the analysis of stellar SEDs (spectral energy distribution) from the Gaia All Sky Parameters for Stars(GASPS) service. This module is provided by ACRI-ST, an SME of the space sector that provides engineering and data services for space missions.

— OBJECTIVES

The student will analyse spectral energy distributin of stars in the GASPS survey using AI. In particular, the student will learn:

- AI techniques and understand their advantages and limitations.
- Run an XGBooat machine learning model and familiar with the GASPS service.

— PREREQUISITES

- ✗ S1. Data Sciences
- ✗ S1. Numerical methods
- ✗ S2. Statistics

It is recommended to have good coding skills in python. Familiarity with machine learning is also beneficial.

— THEORY

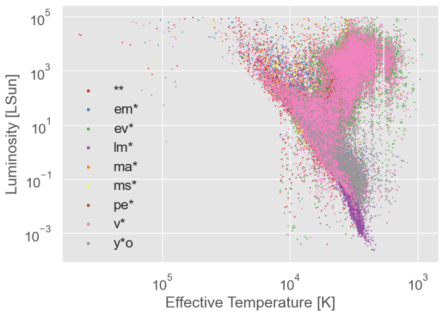
by Jeronimo Bernard-Salas

The module will cover an introduction to machine learning, the Gaia mission, and GASPS which includes data from different missions like Gaia, WISE, 2MASS etc...

— APPLICATIONS

by Nick Cox

The student will analysie GASPS results in the context of supervised machine learning models. The project covers different steps in a data science workflow, from data retrieval, machine learning inference, and data analysis and interpretation.



An HR-diagram from stars of the GASPS service.

— MAIN PROGRESSION STEPS

There module is divided in three main stages as described below:

- `Week 1`: Introduction to machine learning, Gaia and GASPS.
- `Week 2-5`: Model deployment.
- `Week 6-7`: Model Inference, project report and presentation.

— EVALUATION

- `Theory grade [30%]`
  - Written Report (70%): bibliography, thematic and technical description
  - Project Presentation (30%): questions
- `Practice grade [30%]`
  - Progress meetings (30%): progress
  - Project report and presentation (70%): initiative, analysis, results
- `Defense grade [40%]`
  - Oral and slides quality
  - Context
  - Project / Personal work
  - Answers to questions

— BIBLIOGRAPHY & RESOURCES

The links below provide information on basic machine learning and on the methodology at the core of the module.

- Machine learning in python
- Machine learning and data mining for astronomy
- Machine learning based stellar classification with highly sparse photometry data.

Image credits: Title (NASA, ESA, Hubble Legacy Archive, Utkarsh Mishra). HR diagram: S.E. Cody et al. (Open Research Europe, 2024)

— CONTACT

☎ +33.4.92.96.19.56
✉ jeronimo.bernard-salas@acri-st.fr