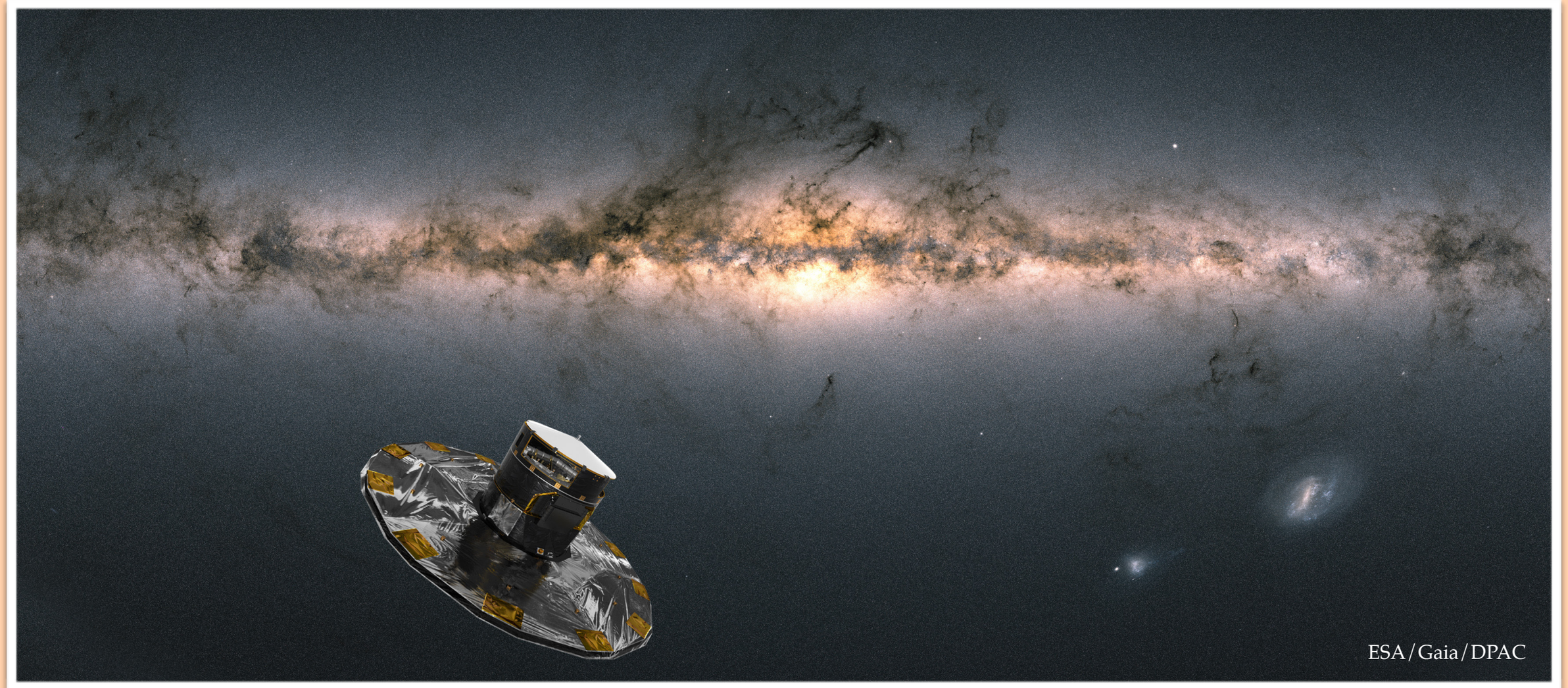


Machine learning algorithms applied to the Gaia data

MiniGrant RSN2 : 15,000€

Tatiana Muraveva,
Gisella Clementini,
Alessia Garofalo,
Maribel Carnerero,
Lorenzo Monti

Project:

Astronomy is entering a new era of Big Data science thanks to exponentially growing data volumes from large surveys, such as the ESA mission Gaia. The extraordinary volume of these data vastly exceeds the discovery capabilities of humans. Thus, exploiting machine learning (ML) techniques, which can provide the accuracy and automation required to exploit large datasets, becomes highly needed. In this project, we apply various ML algorithms to the Gaia data in order to:

- infer the metallicities and distances of RR Lyrae variable stars, contributing to our knowledge of the Milky Way and Local Group galaxies.
- discover new AGN objects.

Expenses:

- Participation in the IAU Symposium 376 “At the cross-roads of astrophysics and cosmology: Period–luminosity relations in the 2020s”, Budapest, Hungary (T. Muraveva)
- Participation in the MW-Gaia 2023 workshop “Science and technology roadmap for μas studies of the Milky Way”, Lund, Sweden (G. Clementini)
- Participation in the “Center on High-Performance Computing, Big Data and Quantum Computing Spoke 3 Technical meeting”, Trieste, Italy (L. Monti)
- Exchange visit with an expert in ML algorithms, Tenerife, Spain (M. Carnerero)
- Exchange visit with an expert in Java, Torino, Italy (L. Monti)

Deliverables:

- Invited talk on the IAU Symposium 376 “At the cross-roads of astrophysics and cosmology: Period–luminosity relations in the 2020s, Budapest, Hungary (T. Muraveva)
- Invited talk on the MW-Gaia 2023 workshop “Science and technology roadmap for μas studies of the Milky Way”, Lund, Sweden (T. Muraveva)
- Contributed talk on the “Center on High-Performance Computing, Big Data and Quantum Computing Spoke 3 Technical meeting”, Trieste, Italy (L. Monti)

Results:

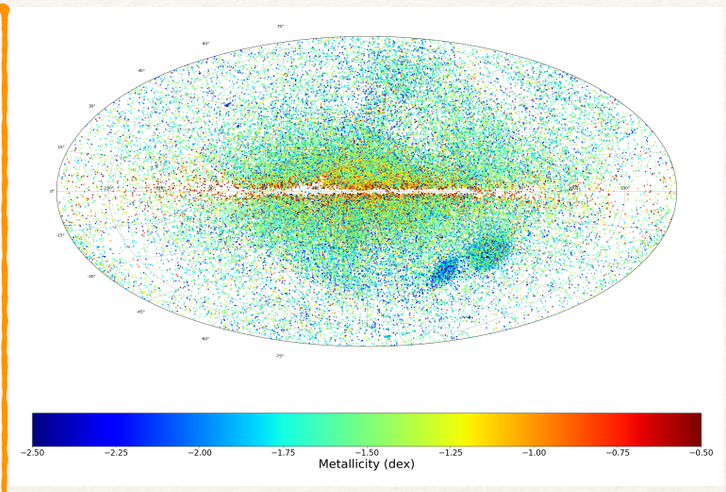
RR Lyrae stars

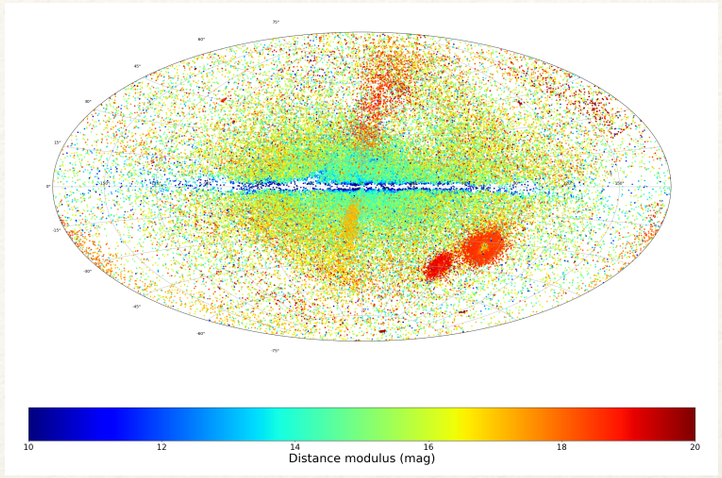
Done:

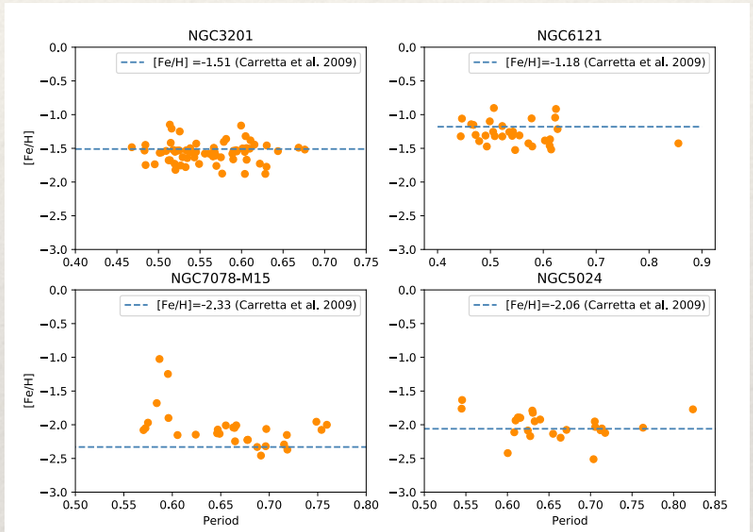
- Collecting of the catalogue of RR Lyrae with accurate metallicity estimates.
- Application of the Sequential Feature Selector and Select K best to identify the most important parameters for determination of metallicity.
- Application of the ML algorithms (e. g. XGBoost, Random Forrest regression) to measure metallicities from the Fourier parameters.
- Bayesian approach to fit the linear relation between Fourier parameters and metallicities.

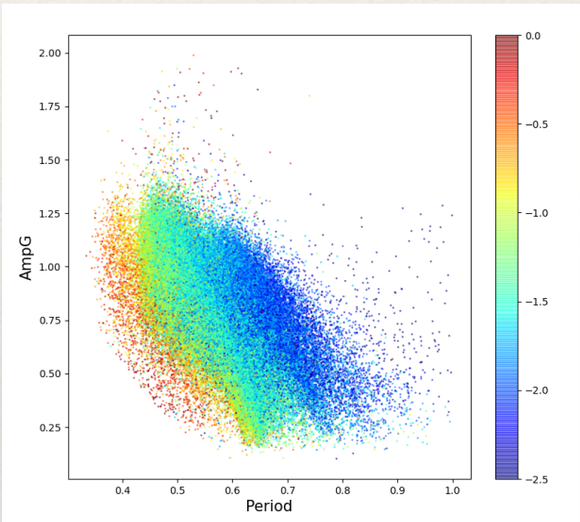
Next steps:

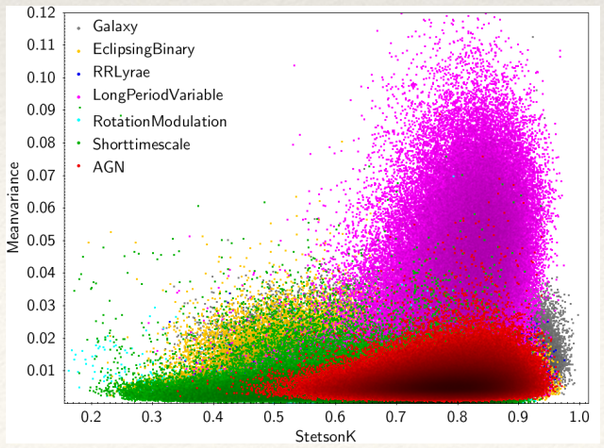
- Time Series Prediction with LSTM Recurrent Neural Networks.
- Transformers architecture applied to the time-series photometry.











AGN

Done:

- Collecting of the catalogue of AGN, stars and galaxies
- Feature selection

Next steps:

- Application of ML (Random Forrest) for classification of AGN