Contribution ID: **36**   Type: **Talk**

Extending the Reach of Gaia DR3 with Self-Supervision

Thursday, November 30, 2023 3:44 PM (15 minutes)

The Gaia Collaboration's 3rd data release (DR3) provides comprehensive information including photometry and kinematics on more than a billion stars across the entire sky up to $G \approx 21$, encompassing approximately 220 million stars with supplementary low-resolution spectra ($G < 17.6$). These spectra offer derived valuable stellar properties like [Fe/H], $\log g$, and T_{eff}, serving as proxies to identify and characterize significant stellar structures, such as stellar streams formed from past minor galaxy mergers with the Milky Way.

In pursuit of constraining the chemo-dynamical history of the Galaxy with data-driven algorithms, we propose a novel self-supervised approach implementing masked stellar modelling (MSM) exploiting multiple spectroscopic and photometric surveys to extend beyond the limitations of DR3's low-resolution spectra. We incorporate diverse imaging surveys that span ultraviolet to near-infrared wavelengths across the celestial sphere. The MSM employs a powerful encoder to generate informative embeddings, containing crucial information for downstream tasks, facilitated by an extensive training sample. By leveraging these embeddings, similarity searches on the complete database of embeddings can be conducted instantly. Moreover, spectroscopic surveys often exhibit inconsistencies due to varying assumptions in their respective derivations of stellar characteristics. The MSM method offers the ability to fine-tune any survey on specific stellar astrophysics tasks with much fewer labels, and thanks to its extensive training set, is more robust to misrepresentativity. The stellar embeddings result in a self-consistent dataset, effectively establishing a comprehensive stellar model.

Overall, this research showcases an innovative data-driven approach to utilize various surveys and spectral products, empowering researchers to make significant strides in understanding the Milky Way's history and dynamics. The methodology's effectiveness in regression tasks and its scalability will be highlighted, shedding light on its broader applicability.

Primary author: MCKAY, Aydan (University of Victoria)

Co-author: Dr FABBRO, Sébastien (NRC Herzberg Astronomy and Astrophysics)

Presenter: MCKAY, Aydan (University of Victoria)

Session Classification: Contributed talks

Track Classification: Paris