

Pressemitteilung**Leibniz-Institut für Astrophysik Potsdam****Dr. Janine Fohlmeister**

10.10.2024

<http://idw-online.de/de/news841021>Forschungsergebnisse, Kooperationen
Elektrotechnik, Geschichte / Archäologie, Gesellschaft, Mathematik, Physik / Astronomie
überregional**A sharper view of the Milky Way with Gaia and machine learning**

A group of scientists led by the Leibniz Institute for Astrophysics Potsdam (AIP) and the Institute of Cosmos Sciences at the University of Barcelona (ICCUB) have used a novel machine learning model to process data for 217 million stars observed by the Gaia mission in an extremely efficient way. The results are competitive with traditional methods used to estimate stellar parameters. This new approach opens up exciting opportunities to map characteristics like interstellar extinction and metallicity across the Milky Way, aiding in the understanding of stellar populations and the structure of our galaxy.

With the third data release of the European Space Agency's Gaia space mission, astronomers gained access to improved measurements for 1.8 billion stars, which provides a vast amount of data for researching the Milky Way. However, analysing such a large dataset efficiently presents challenges. In the now published study, researchers explored the use of machine learning to estimate key stellar properties using Gaia's spectrophotometric data. The model was trained on high-quality data from 8 million stars and achieved reliable predictions with small uncertainties.

"The underlying technique, called extreme gradient-boosted trees allows to estimate precise stellar properties, such as temperature, chemical composition, and interstellar dust obscuration, with unprecedented efficiency. The developed machine learning model, SHBoost, completes its tasks, including model training and prediction, within four hours on a single GPU - a process that previously required two weeks and 3000 high-performance processors," says Arman Khalatyan from AIP and first author of the study. "The machine-learning method is thus significantly reducing computational time, energy consumption, and CO₂ emission." This is the first time such a technique was successfully applied to stars of all types at once.

The model trains on high-quality spectroscopic data from smaller stellar surveys and then applies this learning to Gaia's large third data release (DR3), extracting key stellar parameters using only photometric and astrometric data, as well as the Gaia low-resolution XP spectra. "The high quality of the results reduces the need for additional resource-intensive spectroscopic observations when looking for good candidates to be picked-up for further studies, such as rare metal-poor or super-metal rich stars, crucial for understanding the earliest phases of the Milky Way formation", says Cristina Chiappini from AIP. This technique turns out to be crucial for the preparation of future observations with multi-object spectroscopy, such as 4MIDABLE-LR, a large survey of the Galactic Disc and Bulge that will be part of the 4MOST project at the European Southern Observatory (ESO) in Chile.

"The new model approach provides extensive maps of the Milky Way's overall chemical composition, corroborating the distribution of young and old stars. The data shows the concentration of metal-rich stars in the Galaxy's inner regions, including the bar and bulge, with an enormous statistical power," adds Friedrich Anders from ICCUB.

The team also used the model to map young, massive hot stars throughout the Galaxy, highlighting distant poorly studied regions in which stars are forming. The data also reveal that there exist a number of "stellar voids" in our Milky Way, i.e. areas that host very few young stars. Furthermore, the data demonstrate where the three-dimensional distribution of interstellar dust is still poorly resolved.

As Gaia continues to collect data, the ability of machine-learning models to handle the vast datasets quickly and sustainably makes them an essential tool for future astronomical research. The success of the approach demonstrates the potential for machine learning to revolutionise big data analysis in astronomy and other scientific fields while promoting more sustainable research practices.

wissenschaftliche Ansprechpartner:

Dr. Arman Khalatyan, +49 331 7499 528, akhalatyan@aip.de

Dr. Friedrich Anders, fanders@icc.ub.edu

Dr. Cristina Chiappini, +49 331 7499 454, cristina.chiappini@aip.de

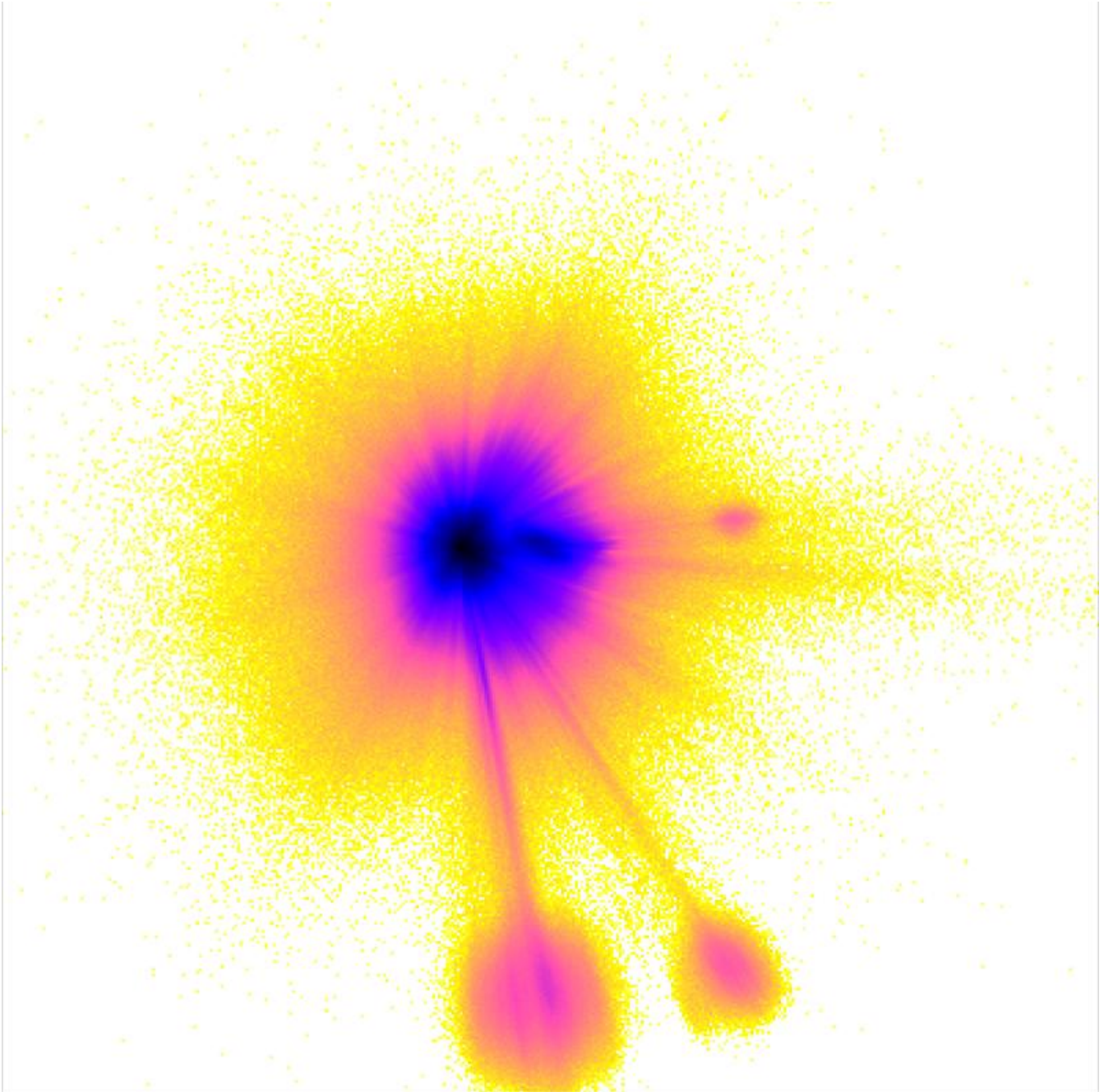
Originalpublikation:

<https://www.aanda.org/component/article?access=doi&doi=10.1051/0004-6361/202451427>

doi:10.1051/0004-6361/202451427

<https://arxiv.org/abs/2407.06963>, doi:10.48550/arXiv.2407.06963

URL zur Pressemitteilung: <https://www.aip.de/en/news/milky-way-gaia-machine-learning/>



Large-scale map (330,000 light years side length) of the density of the 217 million stars from the Gaia DR3 XP sample in Galactocentric Cartesian co-ordinates.

F. Anders (Uni.de Barcelona)
Universitat de Barcelona