# Analysis and Classification of Hot Subdwarfs Using Artificial Intelligence Techniques and Gaia DR3 data

VISCASILLAS VÁZQUEZ, Carlos (Vilnius University); SOLANO, Enrique (Centro de Astrobiología); ULLA-MIGUEL, Ana (Universidade de Vigo); ÁLVAREZ GONZÁLEZ, Marco Antonio (University of A Coruña); SANTOVEÑA GÓMEZ, Raúl (University of A Coruña); MANTEIGA OUTEIRO, Minia (University of A Coruña); DAFONTE VÁZQUEZ, José Carlos (University of A Coruña); PEREZ-FERNÁNDEZ, Esther (Universidade de Vigo); RODRIGO, Carlos (Centro de Astrobiología); ALLER, Alba (Centro de Astrobiología); OREIRO, Raquel (Universidade de Vigo).

Vilnius University

Universidade de Vigo

UNIVERSIDADE DA CORUÑA

carlos.viscasillas@ff.vu.lt

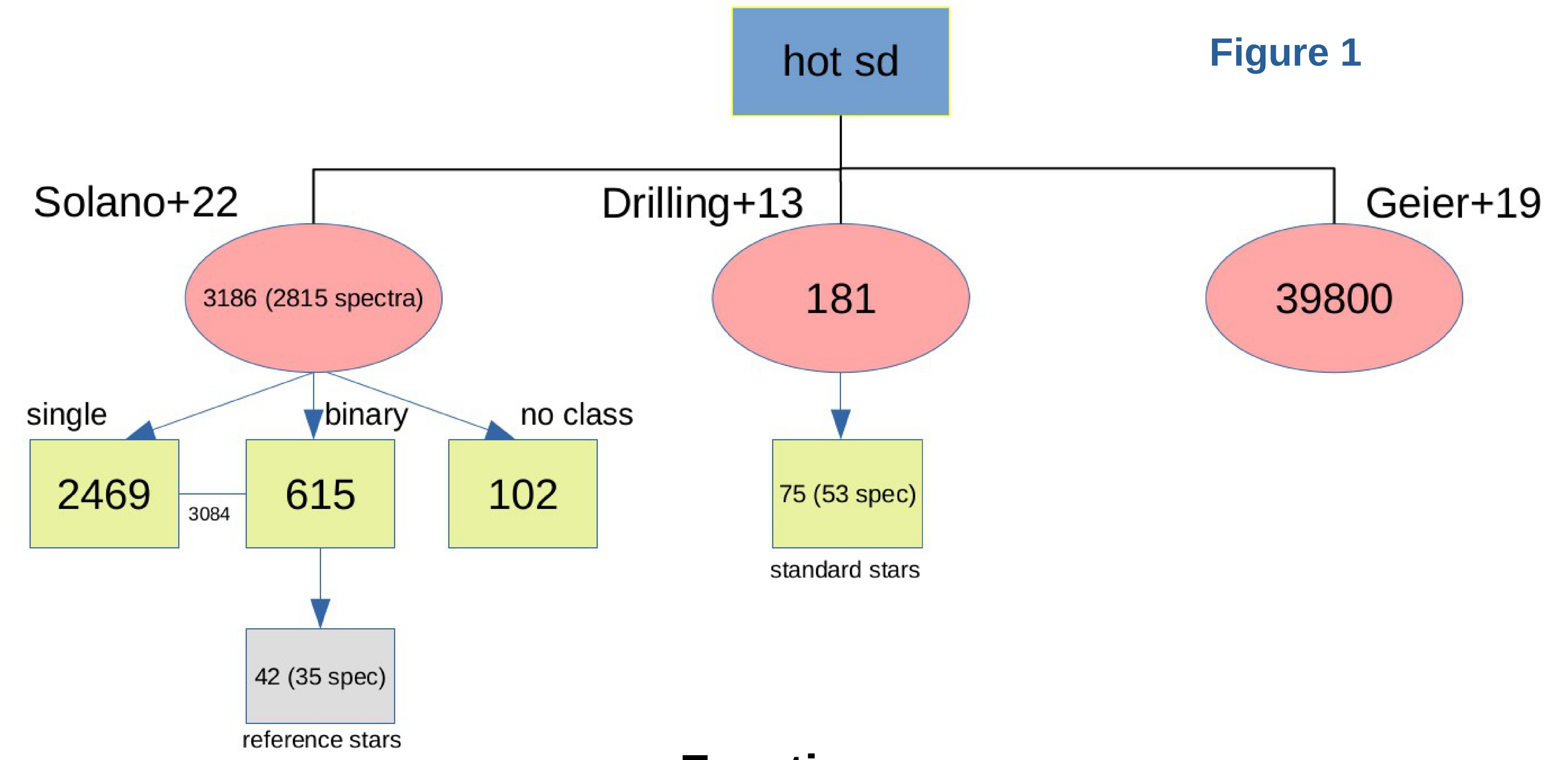
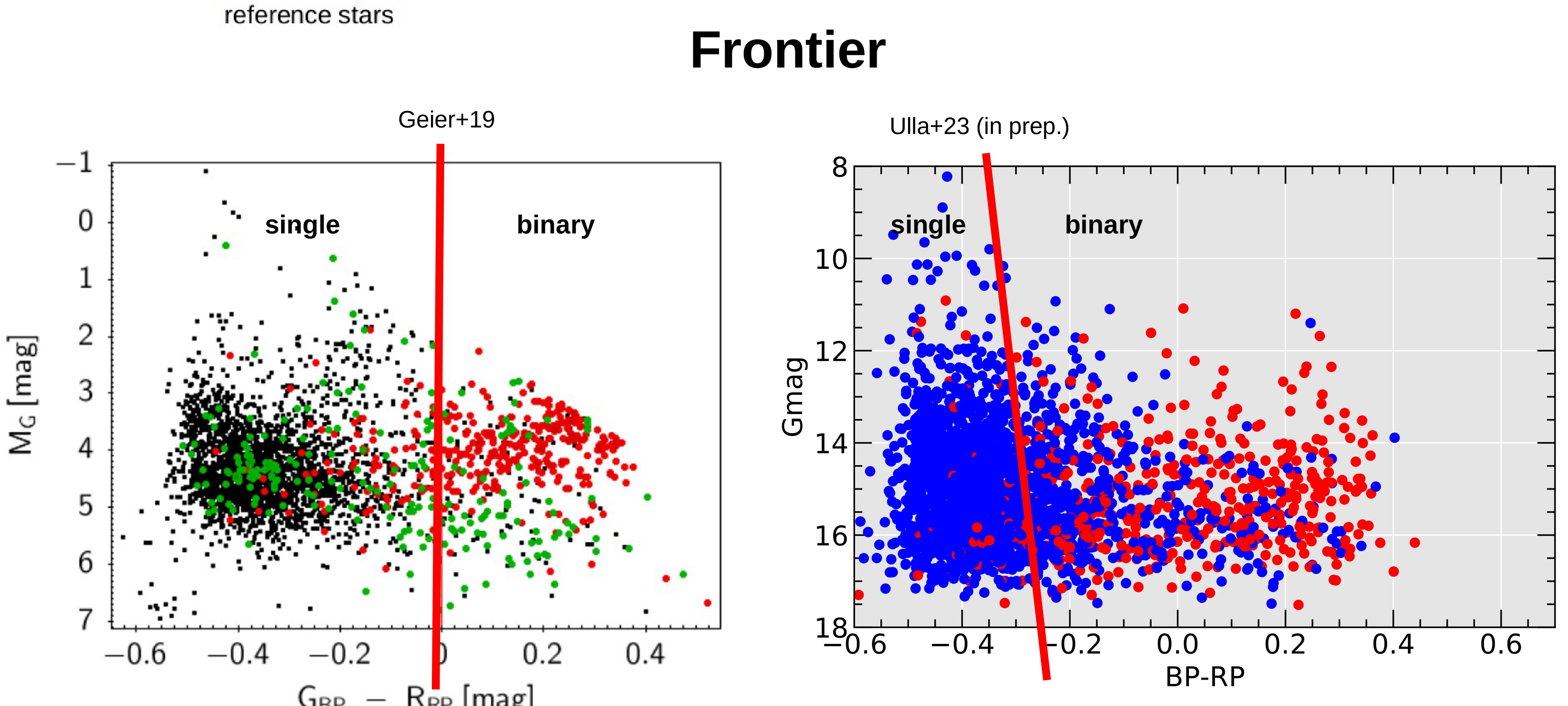



Figure 1



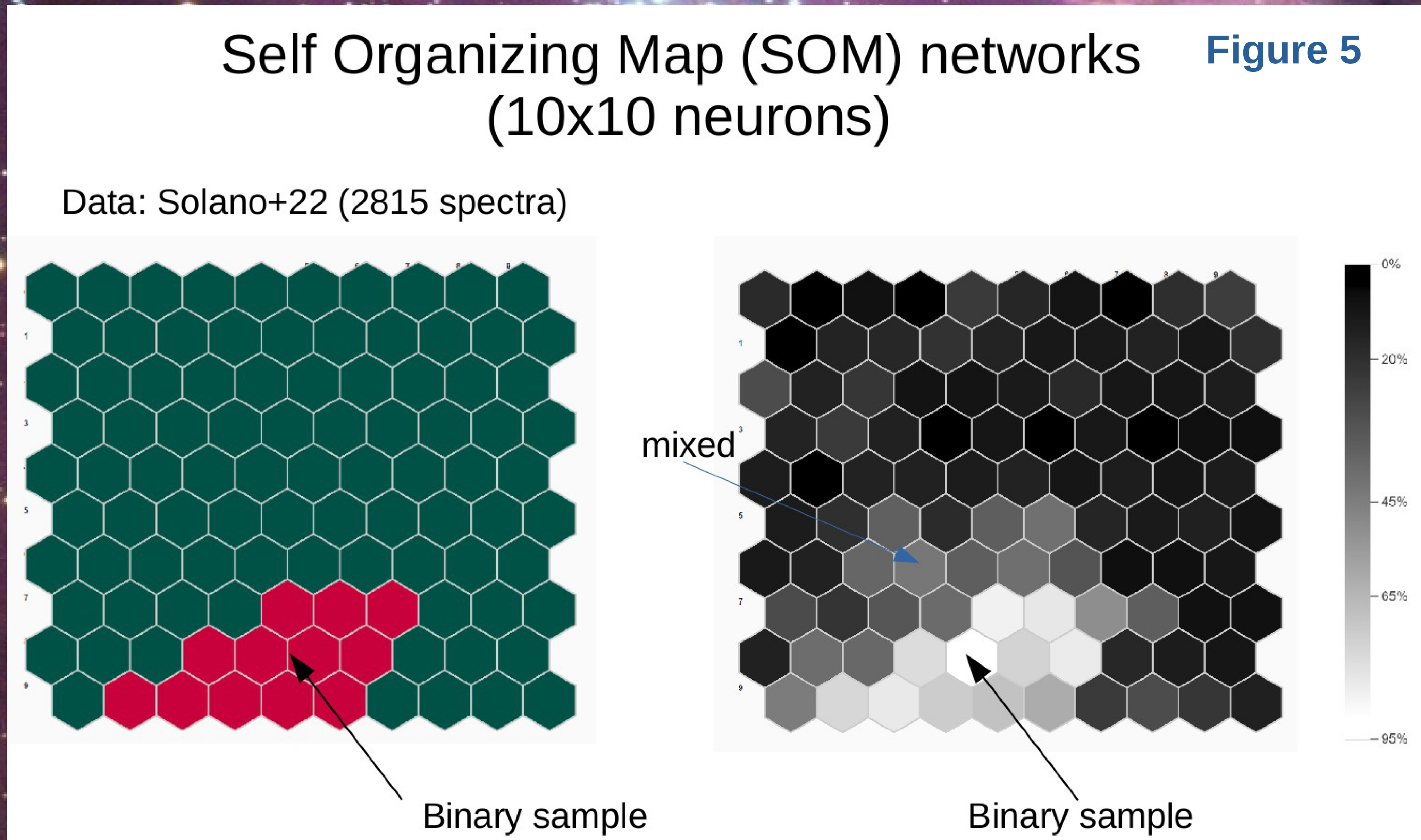

Figures 2 and 3

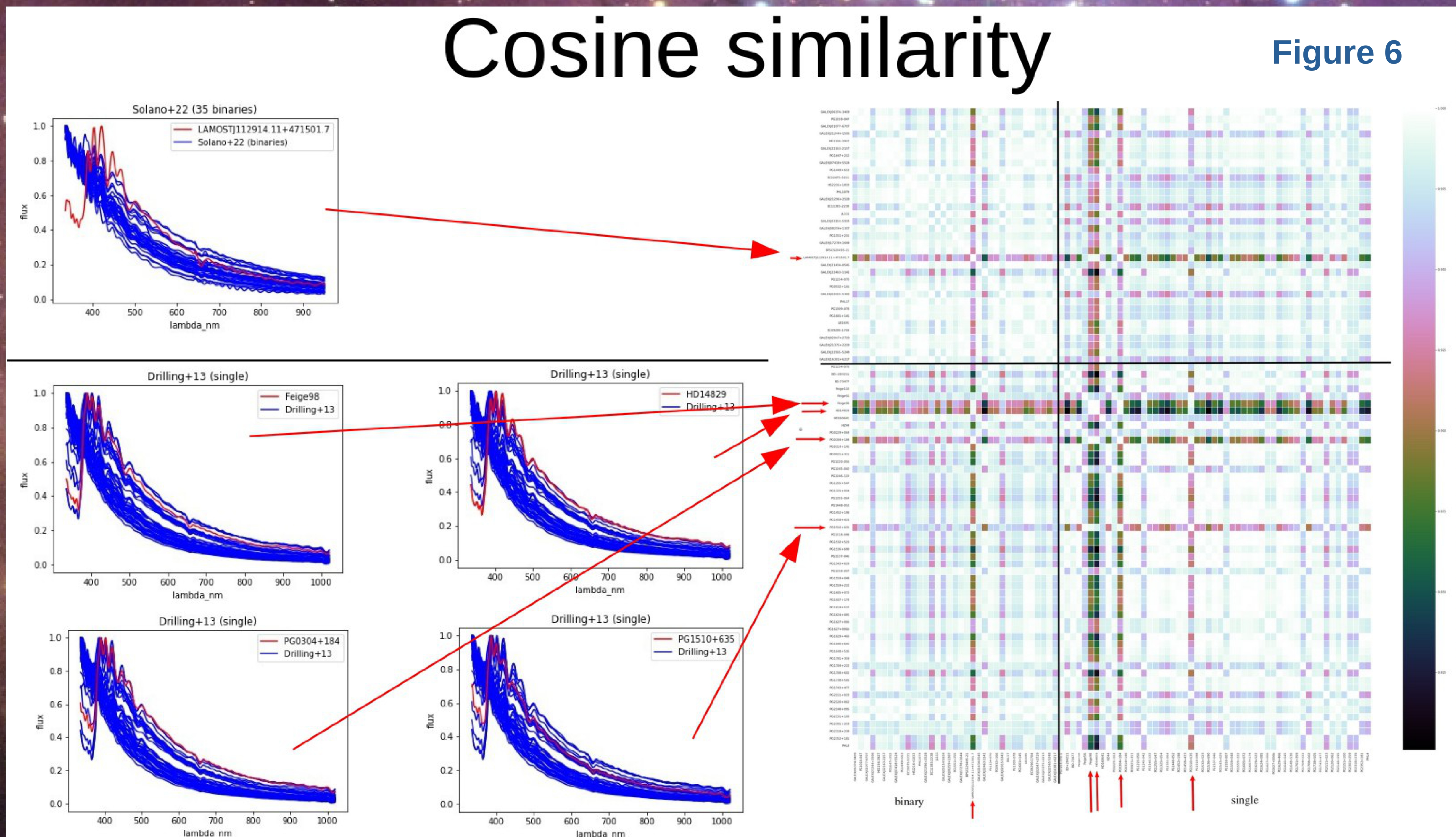
**Introduction**:

This study focuses on analyzing and classifying hot subdwarf stars (hot sds) utilizing Gaia DR3 data and artificial intelligence (AI) techniques. Hot sds are compact blue evolved objects, burning He in their cores surrounded by a thiny H envelope. In the H-R Diagram they are located by the blue end of the Horizontal Branch, near the so-called Extended Horizontal Branch. Besides, most models agree on a quite probable common envelope binary evolution scenario in the Red Giant phase, because it is virtually impossible for a single RG to lose so much of its total mass by its own. However, the actual current binarity rate for this class of objects is a yet unsolved, but key, question in this field. Collaborators from the Galician Group for the Gaia satellite (GGG), University of A Coruña (UDC), University of Vigo (UVIGO), Spanish Virtual Observatory (SVO), and Vilnius University (VU) work together. The aim is to develop a novel classification method for hot subdwarfs in binary systems, employing supervised machine learning (ML) techniques and comparing the results with Self Organized Maps (SOM) and Virtual Observatory (VO) techniques. This approach enhances our understanding of hot subdwarf evolution and binary systems using advanced AI methodologies.

**Methodology**:

In trying to shed light on the problem, our methodology employs Gaia DR3 data and AI techniques, including Self Organized Maps (SOM) (Álvarez et al., 2022), and software tools accessible to the collaboration group, to build up on previous hot sd uncovering via Virtual Observatory techniques (Oreiro et al., 2011; Pérez-Fernández et al., 2016). We have now worked with hot sd catalogues by Stephan Geier in 2020 (with about 6000 confirmed objetcs) and Geier et al. (2019 – about 40000 candidates). We selected training samples from Solano et al. (2022) and Drilling et al. (2013). The method was applied to a larger catalog of over 39,000 blue candidates from Geier et al. (2019) (**Figure 1**). Techniques such as supervised ML and probabilistic color-magnitude approaches were utilized to classify binary systems. At this scope, we compute a single-binary frontier (**Figures 2 and 3**) and probability of belonging to each class (**Figure 4**) using the Support Vector Machines (SVMs) analysis (Boser et al. 1992) implemented using the scikit-learn package (Pedregosa et al. 2011). We defined a training set based on the sample of Solano et al. (2022). For a SOM classification, we worked with GDR3 BP/RP spectra of 2815 objects out of the Geier (2020) catalogue, selecting a well-defined training subsample of 35 binaries from Solano et al. (2022) and 53 single stars from Drilling et al. (2013). The result is plotted in **Figure 5**, with the aid of the Gaia Utility for the Analysis of self-organizing maps (GUASOM). For comparison, we also used the 'cosine similarity' (**Figure 6**) implemented using the scikit-learn package (Pedregosa et al. 2011) to check the spectra in the same 35 binaries + 53 single hot sds SOM trainig sample previosuly defined.



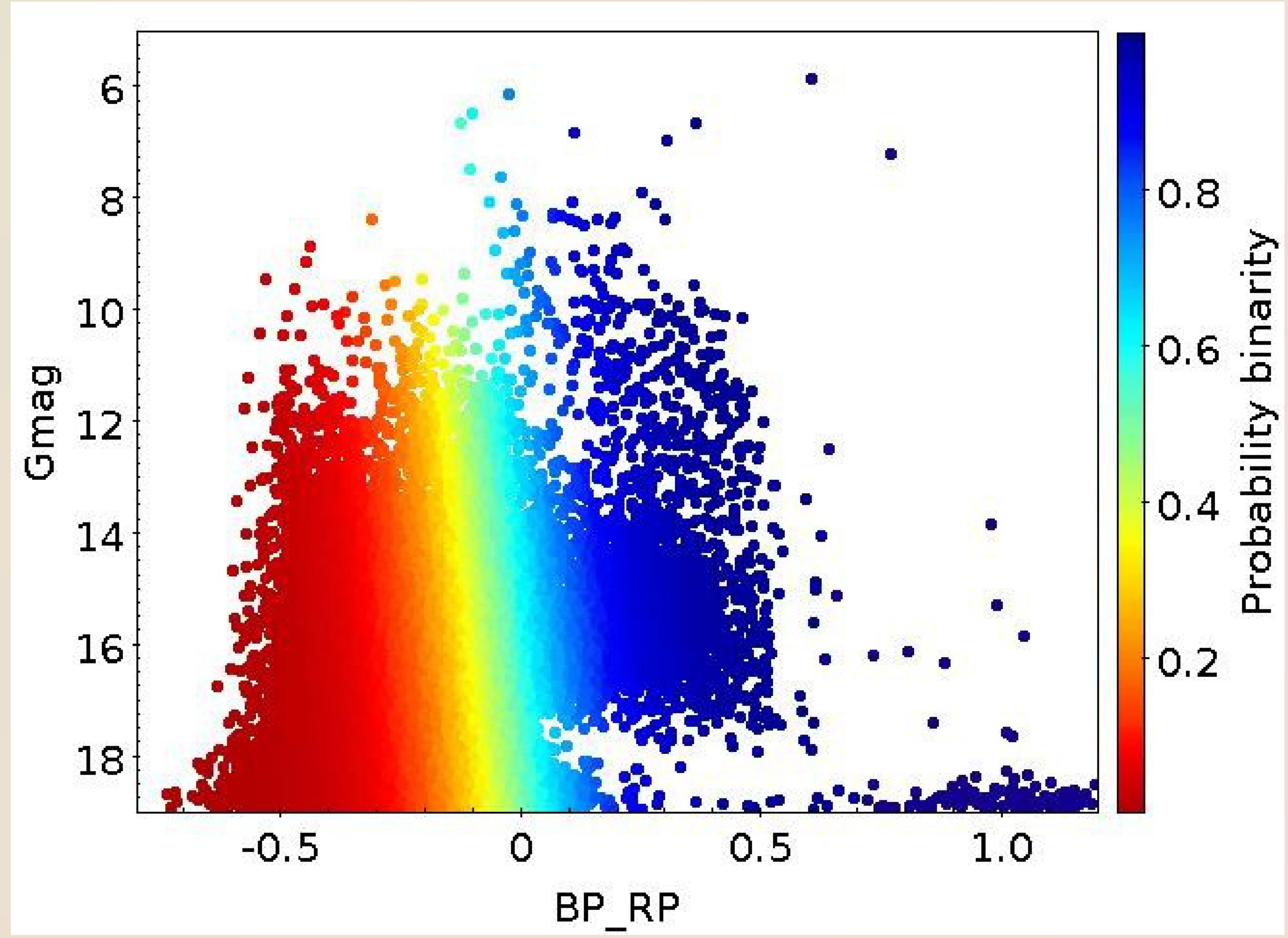




Figure 4

**Results:**

The collaborative effort led to the development of a novel classification method for hot subdwarfs in binary systems using supervised ML techniques. Preliminary results from the application of this method to the sample of binaries by Solano et al. (2022) show promising outcomes. The research team successfully treated Gaia DR3 BP/RP spectra, updated data parameters with Gaia DR3, and tested AI and supervised learning methods for classification. We were also able to identify possibly misclassified spectra using the 'cosine similarity' technique.
For example, one of the stars ("LAMOSTJ112914.11+471501.7", source id: 785814333240812544) from Solano's subsample+22 of 35 stars with spectrum (out of 42) that were classified as binary, both the 'cosine similarity' technique as the SOM classifies it as a single.

**References**:

Álvarez, M. A. ; Dafonte, C.; Manteiga, M.; et al. (2022). NC&A 34.
Boser, B. E., Guyon, I. M., & Vapnik, V. N. 1992, in Proceedings of the Fifth Annual Workshop on Computational Learning Theory, COLT '92, 144–152.
Drilling, J. S. ; Jeffery, C. S.; Heber, U. et al. (2013). A&A 551, A31.
Geier, S. ; Raddi, R.; Gentile Fusillo, N. P. ; Marsh, T. R. (2019). A&A, 621, A38.
Pedregosa, F., Varoquaux, G., Gramfort, A., et al. 2011, Journal of Machine Learning Research, 12, 2825
Solano, E. ; Ulla, A. ; Pérez-Fernández, E. et al. (2022). MNRAS 514, 3.

**Acknowledgments:**

CVV and AUM acknowledge the MW-Gaia COST Action CA18104.
This work is part of a STSM of the presenter at the University of Vigo (supervisor: Ana Ulla).

MW Gaia