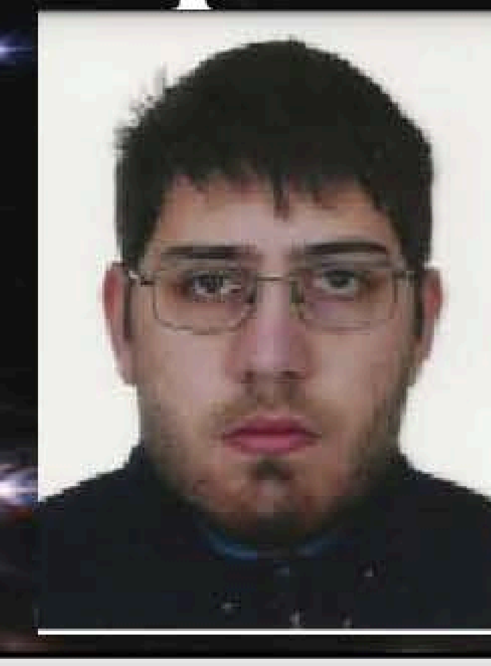
Spectral Classification of Gaia White Dwarfs within 500 pc using a Random Forest Algorithm

García-Zamora, E. M.[1], Torres, S.[1,2], Rebassa-Mansergas, A.[1,2]

[1] Departament de Física, Universitat Politècnica de Catalunya. [2]. Institut d'Estudis Espacials de Catalunya.



Introduction

White dwarfs (WD) are the most common stellar remnant in the Universe. Gaia Data Release 3 (June 2022) has provided low resolution spectra for around 100,000 WD candidates (Gentile-Fusillo et al., 2021).

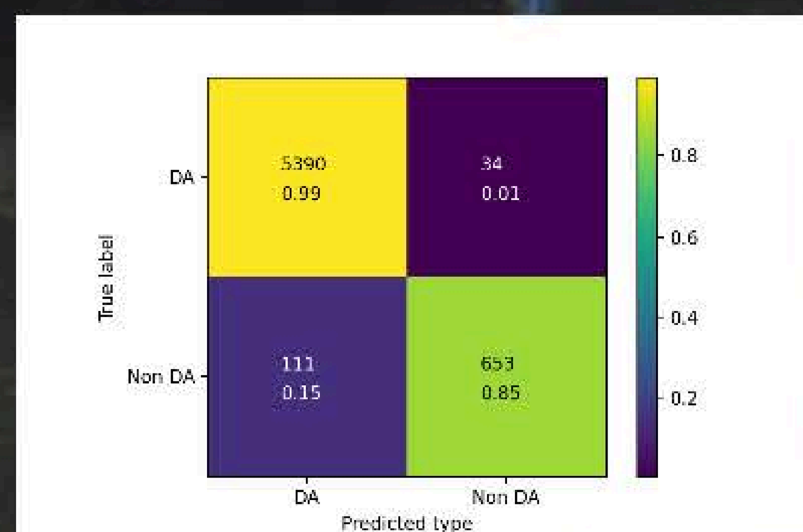
Accurate spectral classification is capital for population analysis, since WD stellar parameters, such as effective temperature, gravity or luminosity are obtained from spectral fitting to atmospheric models. However, with such a high quantity of available spectra, spectral classification by human inspection becomes extremely time consuming. Therefore, we resorted to Machine Learning techniques, namely, a Random Forest algorithm. This algorithm has not only already been used in WD population analysis with reliable results [1][2]; but it has also been used to spectrally classify 9446 WDs from a 100-pc volumen-complete WD simple [3]. We now continue this work by expanding this classification to 500 pc, covering almost every WD with Gaia mean spectra.

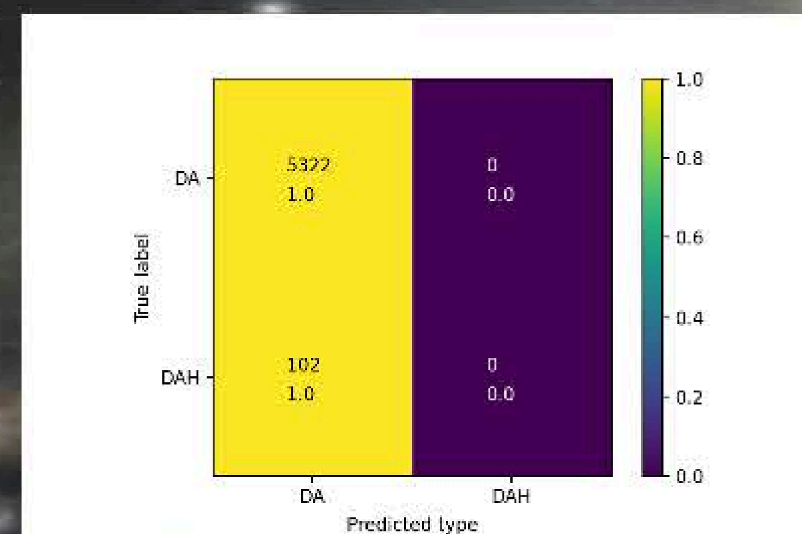
Method

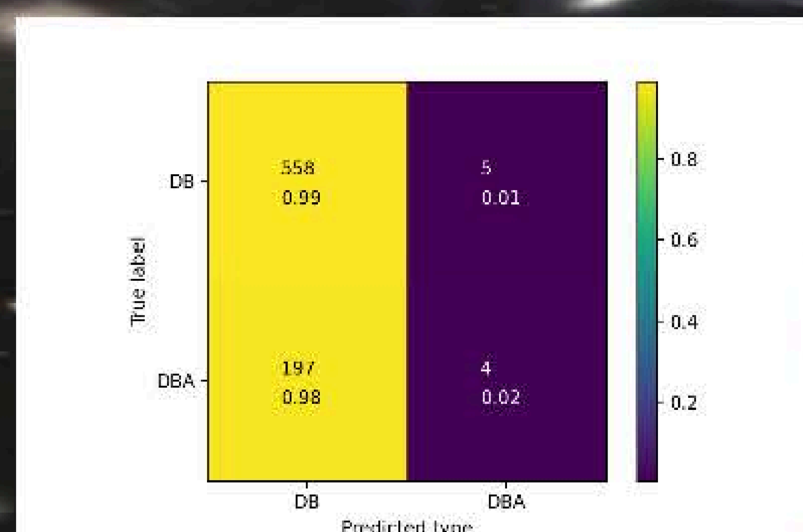
Gaia spectra (both Blue and Red Photometer) are provided as a linear combination of 55 Hermite functions. These 110 spectral coefficients are used as input parameters for the Random Forest;spectral type of classified WDs are also used for algorithm training.

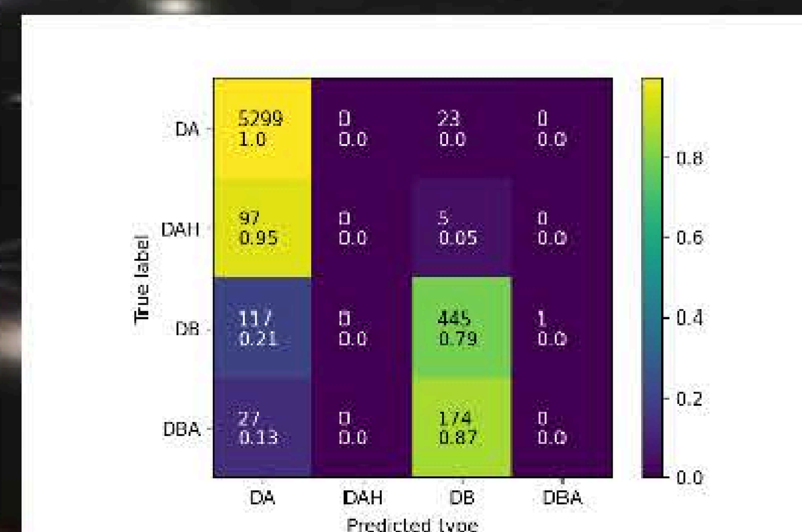
Validation tests

Four validation tests were performed on a set of 6188 WDs with assigned spectral types (DA, DAH, DB and DBA). The first distinguished DA from non DAs (5424 objects), the second classified DAs into DA and DAH, the third one classified non-DAs into DB and DBA (764 objects), and the last test simultaneously classified all objecs into their types.







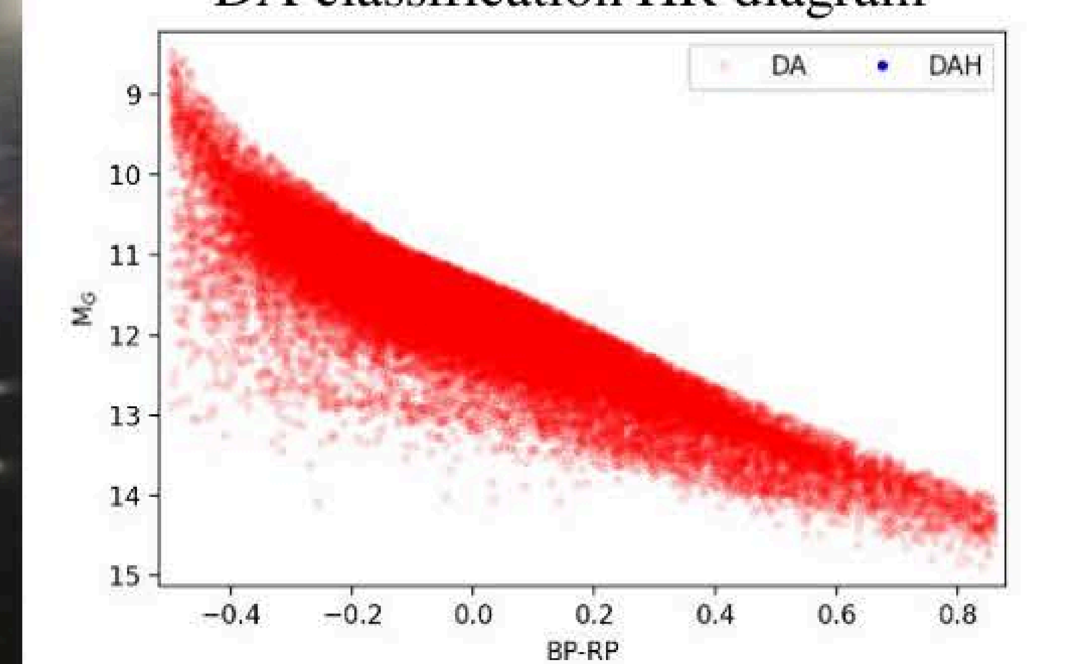


Metric	Validation test	DA	Non-DA	DAH	DB	DBA
Recall	Test 1	0.99	0.85	-	-	-
Recall	Test 2	1	-	0	-	-
Recall	Test 3	-	-	-	0.99	0.02
Recall	Test 4	0.996	-	0	0.79	0
Precision	Test 1	0.98	0.95	-	-	-
Precision	Test 2	0.98	-	0	-	-
Precision	Test 3	-	-	-	0.74	0.44
Precision	Test 4	0.96	-	0	0.69	0

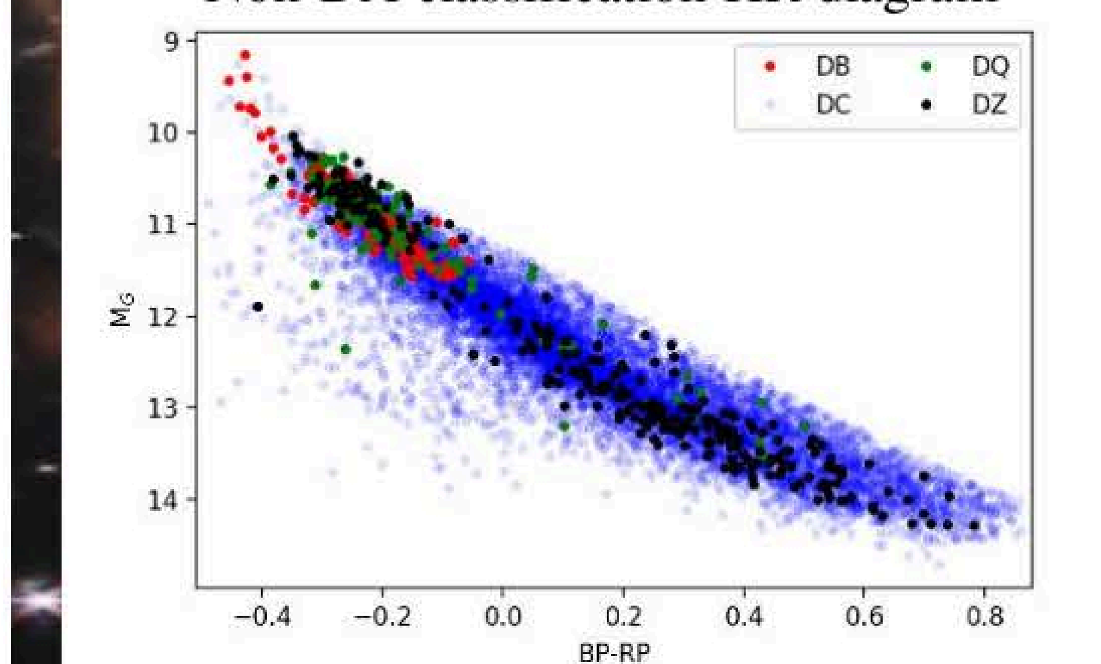
Preliminary results

Our algorithm was applied to a WD sample that had been classified into DA (hydrogen rich atmosphere) and non-DA (hydrogen-deficient atmosphere) [4]. DAs were classified into pure DA and magnetic DAH; non-DAs were classified into DB (He I features), DC (featureless spectrum), DQ (carbon features) and DZ (metallic features). 37581 DAs and 11225 non-DAs were classified into their spectral types. Training sets comprised 1993 already classified DAs and 912 classified non-DAs, respectively.

DA classification HR diagram



Non-DA classification HR diagram



Conclusions

Random Forest is able to distinguish between DAs and non DAs with very good recall (>85%) and excellent precisión (>95%).

Spectral subtypes DAH and DBA appear not to be recognized by the algorithm; noisy spectra and low spectral resolution could explain these results.

References

[1] Torres, S., Cantero, C., Rebassa-Mansergas, A., et al. 2019, MNRAS, 485, 5573

[2] Echeverry, D., Torres, S., Rebassa-Mansergas, A., & Ferrer-Burjachs, A. 2022, A&A, 667, A144

[3] García-Zamora, E. M., Torres, S., & Rebassa-Mansergas, A. 2023, arXiv e-prints, arXiv:2308.07090

[4]Torres, S., Cruz, P., Murillo-Ojeda, R., et al. 2023 [arXiv:2307.13629]

enriquegarciazamora@gmail.com