ENAA 31

Detection of high-redshift Radio Galaxies using Machine Learning models

Rodrigo Carvajal, J. Afonso, I. Matute, S. Amarantidis, D. Barbosa

Instituto de Astrofísica e Ciências do Espaço, Universidade de Lisboa (IA); Departamento de Física, Faculdade de Ciências, Universidade de Lisboa (FCUL)

rcarvajal@oal.ul.pt

Introduction

More than **200 Active Galactic Nuclei (AGN) have been observed at z ≥ 6** [1], few of them with radio, contrary to what simulations predict [2].

It is expected that **radio emission can be detected from early AGN** --called Radio galaxies (RG)--, although its characteristics are still quite indeterminate [3].

Participation of IA in two projects that use SKA Precursors (EMU, with ASKAP, and MIGHTEE with MeerKAT) creates a remarkable opportunity to apply our research in both projects and, eventually, in SKA.

We want to use the power of Machine Learning (ML) techniques to 1) **Understand properties that make AGN Radio Galaxies** and 2) Generate a **catalogue of z Radio Galaxy candidates** which can be observed in the future.

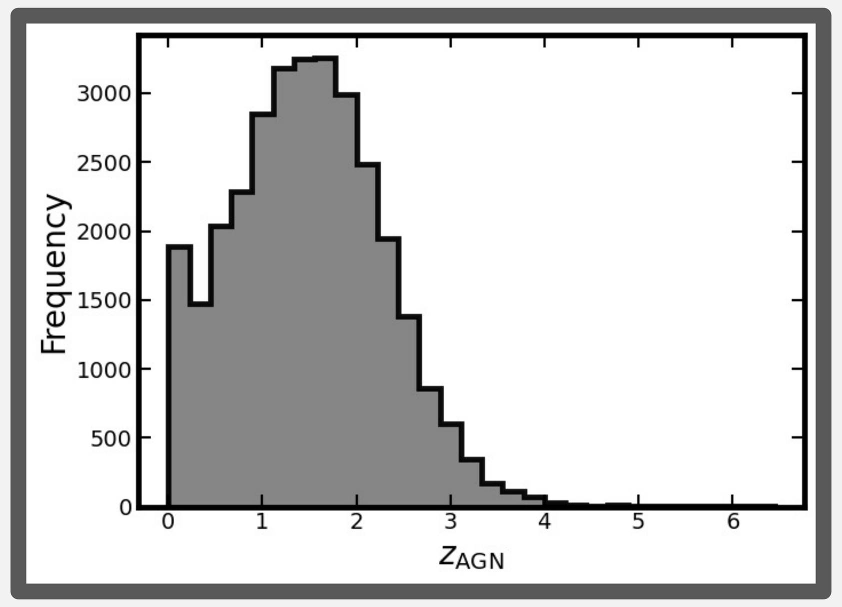
Data and Methods

To generate ML models, we used **photometry from around six million sources in the HETDEX Spring Field** [4] in several wavelengths --Radio, optical, NIR, UV, X-ray--. Close to **30,000 of them are classified as AGN**. Extra features were calculated from some features. The Stripe 82 Field [5] was used as validation set, with close to 350,000 objects. Around 3,000 of them are flagged as AGN.

Three ML models were trained separately. To detect AGN, to determine if these AGN might have radio emission (RGs), and, to predict their z value. For the classification steps (AGN and radio detection), a probability threshold of 75% has ben established for a positive output.

Australian Square Kilometre Array Pathfinder

MIGHTEE: The MeerKAT International GHz Tiered Extragalactic Exploration.



Distribution of redshift values in the training sample

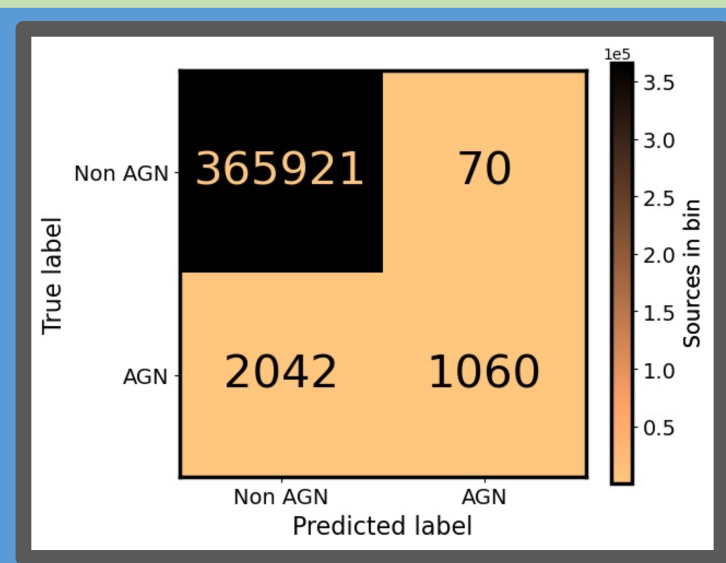
Results and Summary

In the Stripe 82 area, **1,130 objects have been predicted to be AGN, and 266 of them, as having radio detection**. For predicted AGN which have a previous z measurement, their **mean estimated deviation is 0.091**, in line with other redshift estimation techniques. From the Shapley analysis [6], we have recovered known correlations among features and our targets and. Also, we have obtained trends that have not been studied thoroughly in previous studies.

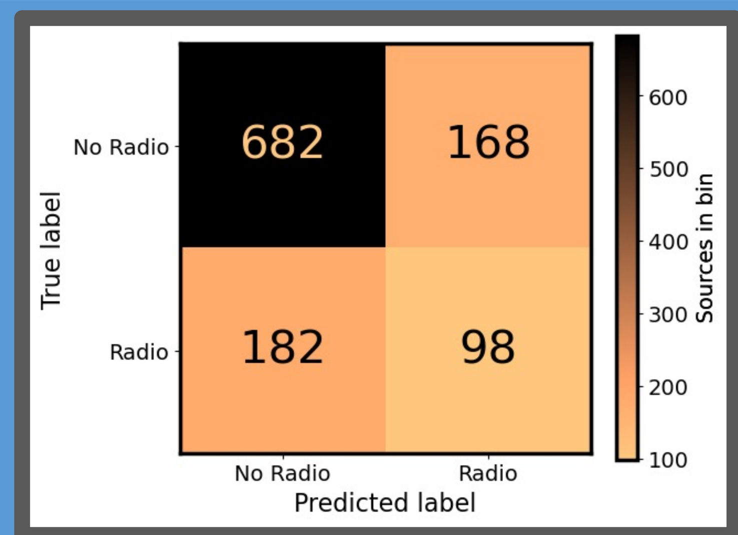
Using already-available photometric data, we can derive predictions for the detection of high-redshift Radio Galaxies.

Our models can be, smoothly, transferred among different areas in the Sky. Final results are not strongly affected by the change in positions.

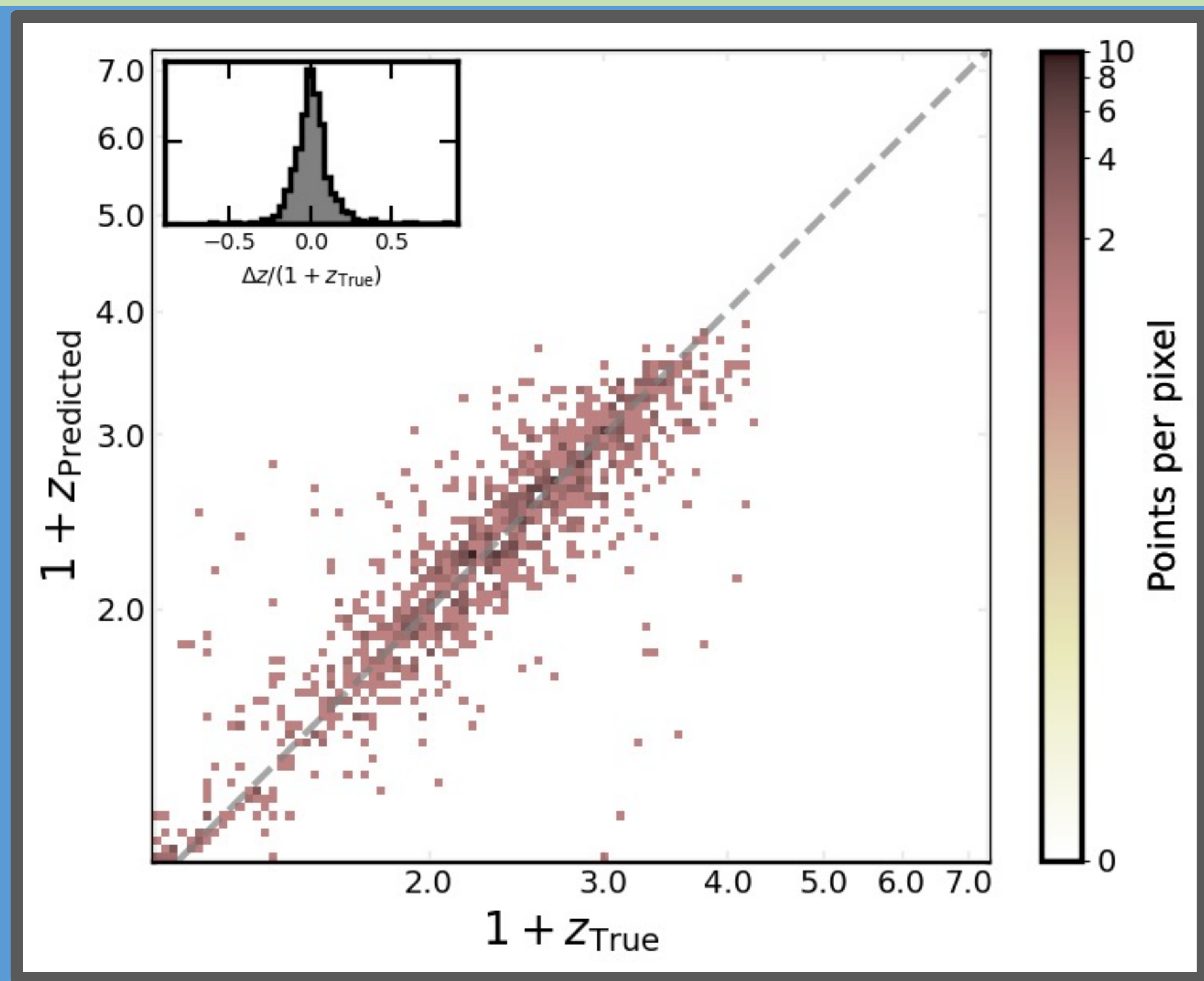
Rather than using them as black boxes, we are able to derive correlations and trends between features and the target properties.



Confusion matrix for AGN prediction in validation sample



Confusion matrix for radio prediction in validation sample



True vs. Predicted redshift for validation sample

Acknowledgements:

This work has been supported by the Fundação para a Ciência e a Tecnologia (FCT) through the Fellowship PD/BD/150455/2019 (PhD::SPACE Doctoral Network PD/00040/2012) and POCH/FSE (EC), and through research grants PTDC/FIS-AST/29245/2017, UID/FIS/04434/2019, UIDB/04434/2020 and UIDP/04434/2020.

References:

[1] Inayoshi, K., Visbal, E. & Haiman, Z., 2020, ARAA, 58:1

[2] Afonso, J. et al., 2015, PoS(AASKA14) 071

[3] Amarantidis, S. et al., 2019, MNRAS, 485, 2694

[4] Hill, G. J. et al., 2008, Panor. Views Galaxy Form. Evol., 399, 115

[5] Hodge, J. A. et al., 2011, AJ, 142, 3

[6] Shapley, L. S., 1953, Contrib. to the Theory of Games (AM-28), Vol. II

FCT Fundação para a Ciência e a Tecnologia | PD F FCT PhD PROGRAMMES | POCH | PORTUGAL 2020 | UNIÃO EUROPEIA Fundo Social Europeu