Astronomy & Astrophysics manuscript no. 0head_ML_3FGL
August 14, 2019

Machine learning methods for classification of unassociated *Fermi* LAT sources

A. Bhat [⋆1] and D. Malyshev [⋆⋆1]

Erlangen Centre for Astroparticle Physics, Erwin-Rommel-Str. 1, Erlangen, Germany

Received September 15, 1996; accepted March 16, 1997

ABSTRACT

Context. Classification of sources is one of the most important tasks in astronomy. Sources detected in one wavelength band, e.g., in gamma rays, may have several possible associations in other wavebands or there may be no plausible association candidates.
Aims. In this work, we take unassociated sources in the third *Fermi*-LAT point source catalog (3FGL) and suggest associations to known classes of gamma-ray sources using machine learning methods trained on associated sources in the 3FGL.
Methods. We use several machine learning methods to separate *Fermi*-LAT sources into two major classes: pulsars and active galactic nuclei (AGNs). We evaluate the dependence of results on meta-parameters of the ML methods, such as the depth of the tree in tree-based classification methods and the number of layers in neural networks. We test the performance of the methods with a test sample drawn from the associated sources in 3FGL. We compare the predictions with the preliminary forth *Fermi*-LAT catalog (4FGL).
Results. Summary of results

Key words. Methods: statistical – Catalogs

Contents

1. Introduction

2. Methods

The purpose of this section is to describe the ML methods that we use for the classification: random forests, boosted decision trees, (simple decision trees?), neural networks, logistic regression, support vector machine.

Methodology

1. Split the PS with known classification into learning and test samples.
2. Use the learning sample for training and for selection of features. In particular, continuous parameters, such as the thresholds in the decision trees or mixing matrices in neural networks, are determined from the leaning sample.
3. Meta-parameters, which encode the complexity of the methods, such as the depth of the decision trees, are determined from the best performance on the test sample.

2.1. Details of the analysis

In this subsection we give more details on the analysis:

1. Describe the features that we use for the analysis.
2. Describe the objective function for minimization (accuracy of classification on learning sample). Weighted objective function: give more weight to pulsars, since there are fewer of them in the catalog.

⋆ e-mail: aakash.bhat@fau.de
⋆⋆ e-mail: dmitry.mayshev@fau.de

3. Learning curve using all features? *Plot: classification accuracy using the total list of features for learning and test sample as a function of complexity parameter.*
4. Selection of the most important features. *Table: features vs algorithms. Columns: algorithms, rows: features, values: significance.*
5. Selection of meta-parameters. *Plot: classification results for the test sample using a subset of features.*
6. Train the final classifier. *Table: classification accuracy of the final classifiers for different algorithms using the test sample from 3FGL.*

Discuss the general features of the optimal algorithms: which features turn out to be important, what is the depth of the trees, the number of trees in random forests, the depth and number of internal nodes in the neural networks.

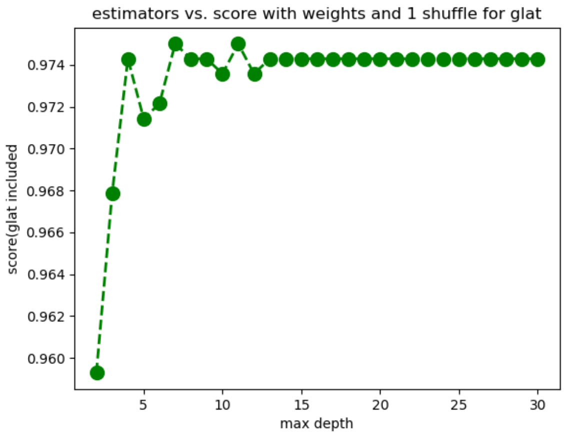

Fig. 1. Example of a figure for one column.

2.2. Comparison of the classification algorithms

Plot: classification domains for a pair of features (or different pairs of features, e.g., latitude vs index, index vs curvature, latitude vs variability).

Probabilistic classification? Result: probability for a source to belong to a particular class. Result of classification: table of sources with probabilities for different algorithms. Final probability: the probability for one of the algorithms (for the most precise one?) and uncertainties determined from the other algorithms.

Discuss a few examples where algorithms give different predictions (are these sources at the boundaries of the domains).

Discuss examples where algorithms misclassify sources from the test sample.

3. Prediction for unassociated sources in 3FGL and comparison with 4FGL

Apply the algorithms on unassociated sources in 3FGL.

Plot: add unassociated sources on the plots with domains for the best algorithm.

Create a table with sources which are more likely to be pulsars (select about 20 the most likely candidates). Compare the accuracy of the algorithm for the sources which have an associate now in the 4FGL catalog.

4. Comparison with 4FGL

5. Conclusions

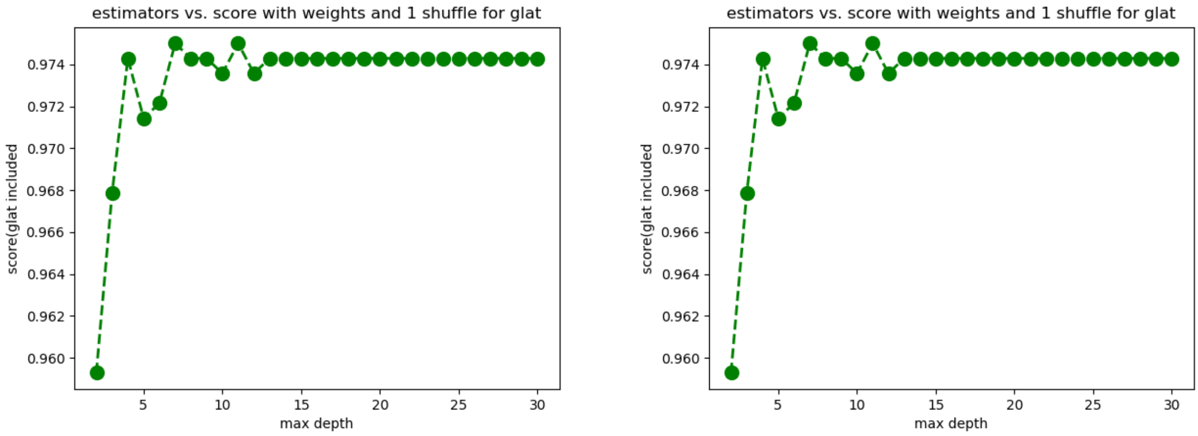

Fig. 2. Example of a figure for both columns.

Appendix A: Appendix

If we need one.