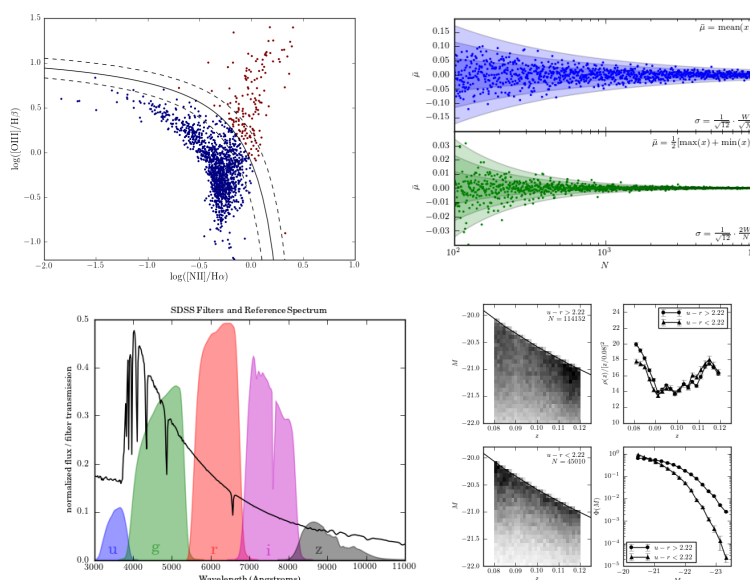



[Home](#)
[User Guide](#)
[Book Figures](#)
[Examples Plots](#)

AstroML: Machine Learning and Data Mining for Astronomy



AstroML is a Python module for machine learning and data mining built on [numpy](#), [scipy](#), [scikit-learn](#), [matplotlib](#), and [astropy](#), and distributed under the 3-clause BSD license. It contains a growing library of statistical and machine learning routines for analyzing astronomical data in Python, loaders for several open astronomical datasets, and a large suite of examples of analyzing and visualizing astronomical datasets.

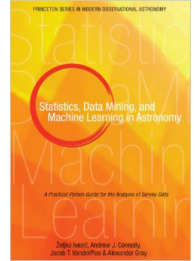
Downloads

- Released Versions: [Python Package Index](#)
- Bleeding-edge Source: [github](#)

The goal of astroML is to provide a community repository for fast Python implementations of common tools and routines used for statistical data analysis in astronomy and astrophysics, to provide a uniform and easy-to-use interface to freely available astronomical datasets. We hope this package will be useful to researchers and students of astronomy. If you have an example you'd like to share, we are happy to accept a contribution via a GitHub [Pull Request](#): the code repository can be found at <http://github.com/astroML/astroML>.

Textbook

The astroML project was started in 2012 to accompany the book **Statistics, Data Mining, and Machine Learning in Astronomy** by Zeljko Ivezic, Andrew Connolly, Jacob VanderPlas, and Alex Gray, published by [Princeton University Press](#). The table of contents is available [here\(pdf\)](#), or you can preview or purchase the book on [Amazon](#).



Did you find a mistake or typo in the book? We maintain an up-to-date [listing of errata](#) in the text which you can view on GitHub. If you find a mistake which is not yet noted on that page, please let us know via email or GitHub pull request!

User Guide

1. Introduction

- 1.1. Philosophy
- 1.2. Textbook

2. Installation of astroML

- 2.1. Important Links
- 2.2. Installation
- 2.3. Dependencies

3. Data Sets

- 3.1. Data Set Cache Location
- 3.2. SDSS Data
- 3.3. Combined Surveys
- 3.4. Time Domain Data
- 3.5. WMAP temperature map

4. Unsupervised Learning: Density Estimation

- 4.1. Bayesian Blocks: Histograms the Right Way
- 4.2. Extreme Deconvolution
- 4.3. Kernel Density Estimation
- 4.4. Nearest Neighbors Density Estimation

5. Unsupervised Learning: Clustering

6. Two-point Correlation Functions

7. Unsupervised Learning: Dimensionality Reduction

8. Supervised Learning: Regression

9. Supervised Learning: Classification

10. Time Series Analysis

11. Class reference

- 11.1. Plotting Functions: `astroML.plotting`
- 11.2. Density Estimation & Histograms: `astroML.density_estimation`
- 11.3. Linear Regression & Fitting: `astroML.linear_model`
- 11.4. Loading of Datasets: `astroML.datasets`
- 11.5. Time Series Analysis: `astroML.time_series`
- 11.6. Statistical Functions: `astroML.stats`
- 11.7. Dimensionality Reduction: `astroML.dimensionality`
- 11.8. Correlation Functions: `astroML.correlation`
- 11.9. Filters: `astroML.filters`
- 11.10. Fourier and Wavelet Transforms: `astroML.fourier`
- 11.11. Luminosity Functions: `astroML.lumfunc`
- 11.12. Classification: `astroML.classification`
- 11.13. Resampling: `astroML.resample`
- 11.14. Addon Functions `astroML_addons`

12. Other Resources

Example Plot Galleries

General astroML Examples

- Data Processing Algorithms
- Data set Examples
- Machine Learning and Data Modeling

Textbook Figures

- Figure Contents
- Getting Started/Frequently Asked Questions

Paper Figures

- CIDU 2012: Introduction to astroML

Development

1. Obtaining the Source Code

2. Contribution

- 2.1. General
- 2.2. Tests
- 2.3. Documentation and Examples
- 2.4. Add-on code

3. Authors

- 3.1. Package Author
- 3.2. Code Contribution

Citing astroML

If you make use of any of these datasets, tools, or examples in a scientific publication, please consider citing astroML. You may reference the following paper:

- [Introduction to astroML: Machine learning for astrophysics](#), Vanderplas *et al*, proc. of CIDU, pp. 47-54, 2012.

Recipient of the best paper award for CIDU 2012

Bibtex entry:

```
@INPROCEEDINGS{astroML,
  author={{Vanderplas}, J.T. and {Connolly}, A.J.
    and {Ivezi{'c}}, {\v Z}. and {Gray}, A.},
  booktitle={Conference on Intelligent Data Understanding (CIDU)},
  title={Introduction to astroML: Machine learning for astrophysics},
  month={oct.},
  pages={47 -54},
  doi={10.1109/CIDU.2012.6382200},
  year={2012}
}
```

You may also reference the accompanying textbook:

- [Statistics, Data Mining, and Machine Learning for Astronomy](#), Ivezić *et al*, 2014

Bibtex entry:

```
@BOOK{astroMLText,
  title={Statistics, Data Mining and Machine Learning in Astronomy},
  author={{Ivezi{'c}}, {\v Z}. and {Connolly}, A.J.
    and {Vanderplas}, J.T. and {Gray}, A.},
  publisher={Princeton University Press},
  location={Princeton, NJ},
  year={2014}
}
```

