Applications of ML in Astronomy

Hamza BA-MOHAMMED
EAIC Deputy President, Training Head

Sara EL-KALLOUTI
Data Head

ENSIAS AI Club

November 16, 2022

Summary

What is astronomy?

Astronomy (from Ancient Greek $\alpha\sigma\tau\rho o\nu o\mu i\alpha$ (astronomía) 'science that studies the laws of the stars') is a natural science that studies celestial objects and phenomena.
It uses mathematics, physics, and chemistry in order to explain their **origin** and **evolution**.

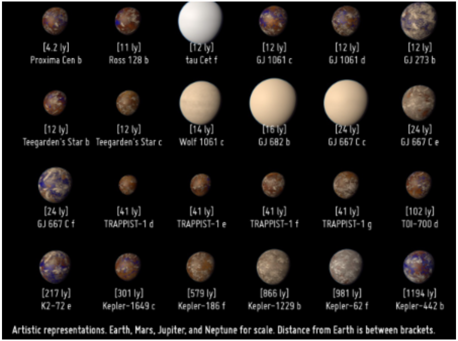
Sub-fields of astronomy

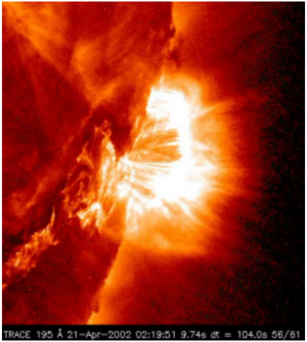
- Astrobiology
- Astrochemistry
- Astrometry
- Astrophysics
- Cosmology
- Planetary geology
- ...

What are astronomers dealing with nowadays?

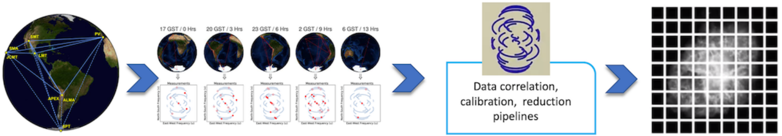
- **Exosolar planetary systems:**
  - Can we find other planetary systems that look like the Solar System?
  - Can we find Earth-like planets in another's star's habitable zone?
  - How do planetary systems form and evolve?
  - Can we find evidence for life elsewhere?
- **Violent events:**
  - How and why do some solar active regions produces huge flares?
  - could flares of other further stars produce damage to us on earth?

What are astronomers dealing with nowadays?





first ever black hole image reconstruction



Event Horizon Telescope — Petabytes of data (measurements) — Scientific, Astrophysics & Computing Algorithms — Image analysis

https://github.com/achael/eht-imaging/

Imaging, analysis and simulation software for radio interferometry

Software: DiFX (Deller et al. 2011), CALC, PolConvert (Martí-Vidal et al. 2016), HOPS (Whitney et al. 2004), CASA (McMullin et al. 2007), AIPS (Greisen 2003), ParselTongue (Kettenis et al. 2006), GNU Parallel (Tange 2011), GILDAS, eht-imaging (Chael et al. 2016, 2018), Numpy (van der Walt et al. 2011), Scipy (Jones et al. 2001), Pandas (McKinney 2010), Astropy (The Astropy Collaboration et al. 2013, 2018), Jupyter (Kluyver et al. 2016), Matplotlib (Hunter 2007).

NumPy was one of the software used!

M87 – the first image of a black hole

Decision Trees for Automated Identification of Cosmic-Ray Hits in *Hubble Space Telescope* Images

STEVEN SALZBERG[1]
Department of Computer Science, Johns Hopkins University, Baltimore, Maryland 21218
Electronic mail: salzberg@cs.jhu.edu

RUPALI CHANDAR
Department of Physics and Astronomy, Johns Hopkins University, Baltimore, Maryland 21218
Electronic mail: rupali@jhufos.pha.jhu.edu

HOLLAND FORD
Department of Physics and Astronomy, Johns Hopkins University, Baltimore, Maryland 21218
Electronic mail: ford@stsci.edu

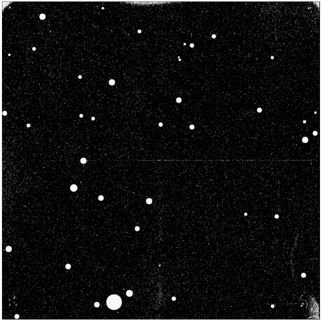
SREERAMA K. MURTHY
Department of Computer Science, Johns Hopkins University, Baltimore, Maryland 21218
Electronic mail: murthy@cs.jhu.edu

RICHARD WHITE
Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, Maryland 21218
Electronic mail: rlw@stsci.edu

Received 1994 April 18; accepted 1994 November 30

ABSTRACT. We have developed several algorithms for classifying objects in astronomical images. These algorithms have been used to label stars, galaxies, cosmic rays, plate defects, and other types of objects in sky surveys and other image databases. Our primary goal has been to develop techniques that classify with high accuracy, in order to ensure that celestial objects are not stored in the wrong catalogs. In addition, classification time must be fast due to the large number of classifications and to future needs for on-line classification systems. This paper reports on our results from using decision-tree classifiers to identify cosmic-ray hits in *Hubble Space Telescope* images. This method produces classifiers with over 95% accuracy using data from a single, unpaired image. Our experiments indicate that this accuracy will get even higher if methods for eliminating background noise improve.

Hubble telescope and one of its images



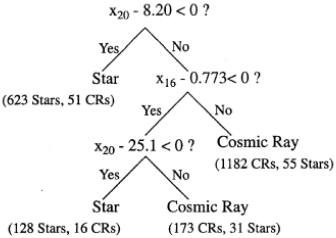


FIG. 4—Small, accurate tree produced by OC1-AP.

TABLE 1
Comparison of Decision-Tree Accuracies for WF1 Data

	Accuracy (%)			Tree
Algorithm	Overall	Stars	Cosmic Rays	Size
CART	93.7±1.4	90.2	95.7	6.6±2.3
C4.5	92.7±1.3	89.1	94.8	36.6±3.1
OC1	91.8±1.2	87.9	94.1	7.5±1.3
OC1-AP	92.9±0.5	88.6	95.4	7.3±1.4
5-NN	89.4	87.3	90.6	

- trees built stochastically: report sd over 10 runs
- model complexity reported as mean number of nodes

A&A 467, 1373–1387 (2007)
DOI: 10.1051/0004-6361:20077334

Astronomy & Astrophysics

Estimation of stellar atmospheric parameters from SDSS/SEGUE spectra

P. Re Fiorentin[1], C. A. L. Bailer-Jones[1], Y. S. Lee[2], T. C. Beers[2], T. Sivarani[2], R. Wilhelm[3], C. Allende Prieto[4], and J. E. Norris[5]

[1] Max Planck Institut für Astronomie, Königstuhl 17, 69117 Heidelberg, Germany
e-mail: fiorent@mpia.de
[2] Department of Physics & Astronomy, CSCE: Center for the Study of Cosmic Evolution, and JINA: Joint Institute for Nuclear Astrophysics, Michigan State University, East Lansing, MI 48824, USA
[3] Department of Physics, Texas Tech University, Lubbock, TX 79409, USA
[4] Department of Astronomy, University of Texas, Austin, TX 78712, USA
[5] Research School of Astronomy and Astrophysics, Australian National University, Weston, ACT 2611, Australia

Received 20 February 2007 / Accepted 8 March 2007

- supervised learning: spectra to stellar APs
- nonlinear, multidimensional regression (ANN)
- dimensionality reduction via PCA

MILKY WAY MAPPER

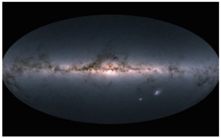


Image credit: ESA/Gaia/DPAC

A time-domain, optical+IR spectroscopic survey of Milky Way stars of all types.

Program Head

Jennifer Johnson
(The Ohio State University)

Explore MWM

LOCAL VOLUME MAPPER

Image credit: NASA, ESA/A. Nota

An optical, integral-field spectroscopic survey of the Milky Way and its neighbors.

Program Head

Niv Drory
(University of Texas at Austin)

Explore LVM

BLACK HOLE MAPPER

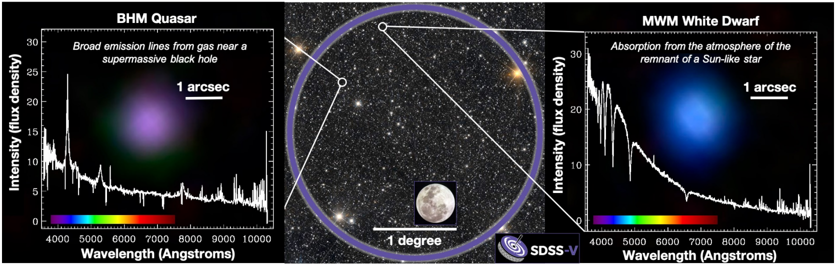
Image credit: ESO/M. Kornmesser

An optical time-domain spectroscopic survey of quasars and X-ray sources.

Program Head

Scott Anderson
(University of Washington)

Explore BHM


BHM Quasar
Broad emission lines from gas near a supermassive black hole
1 arcsec
Intensity (flux density)
Wavelength (Angstroms)
1 degree
SDSS-V
MWM White Dwarf
Absorption from the atmosphere of the remnant of a Sun-like star
1 arcsec
Intensity (flux density)
Wavelength (Angstroms)

Example topics

- Signal detection
- Object classification
- Stellar spectral parametrization
- Galaxy and quasar spectral parametrization
- (Galaxy) morphological classification
- Time series
- Spatial clustering
- Population analysis
- Class discovery
- Large scale structure

How we'll proceed?

- teams of 5 people: 2 data + 2 training (1A) + 1 member 2A
- 1 subject per team
- up to 20 teams
- the work will be mentored by the data and training heads mainly, and occasionally by other volunteers
- an eventual collaboration with a Moroccan astronomic team for scientific review of the astrophysics part (guess who!)
- parallel training sessions also on the engineer and researcher tools (Git and Github, LaTeX, ResearchGate, article redaction, etc..)

Webography

- NASA Slides
- Max Planck Institute Seminary about ML applications in Astronomy (2011)
- NumPy Documentation
- SDSS-V
- Sloan Extension for Galactic Understanding and Exploration (SEGUE)

Applications of ML in Astronomy

Hamza BA-MOHAMMED
EAIC Deputy President, Training Head

Sara EL-KALLOUTI
Data Head

ENSIAS AI Club

November 16, 2022