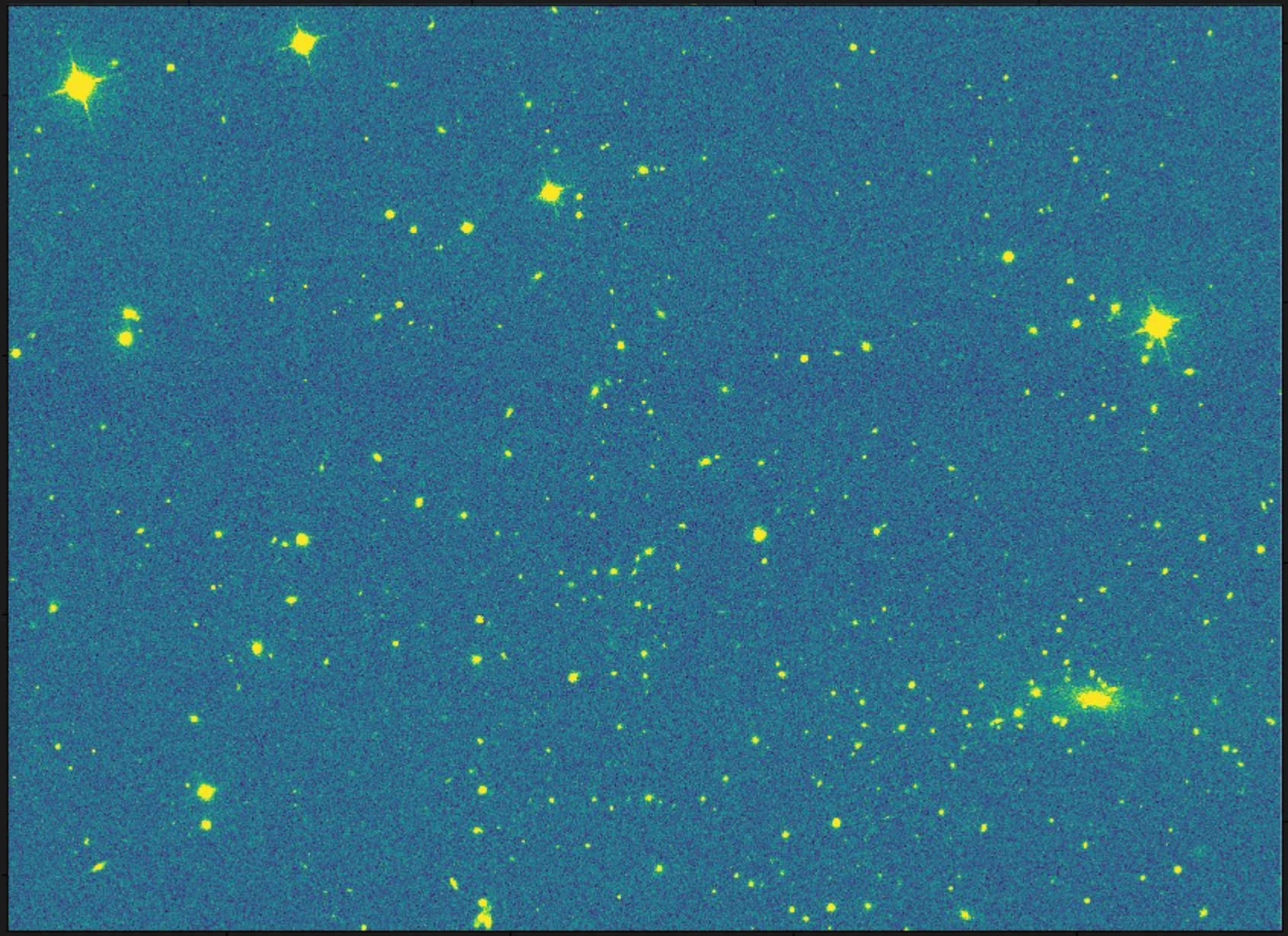
Stars,

Identifying Spirals and Elliptical Galaxies in SDSS using Machine Learning

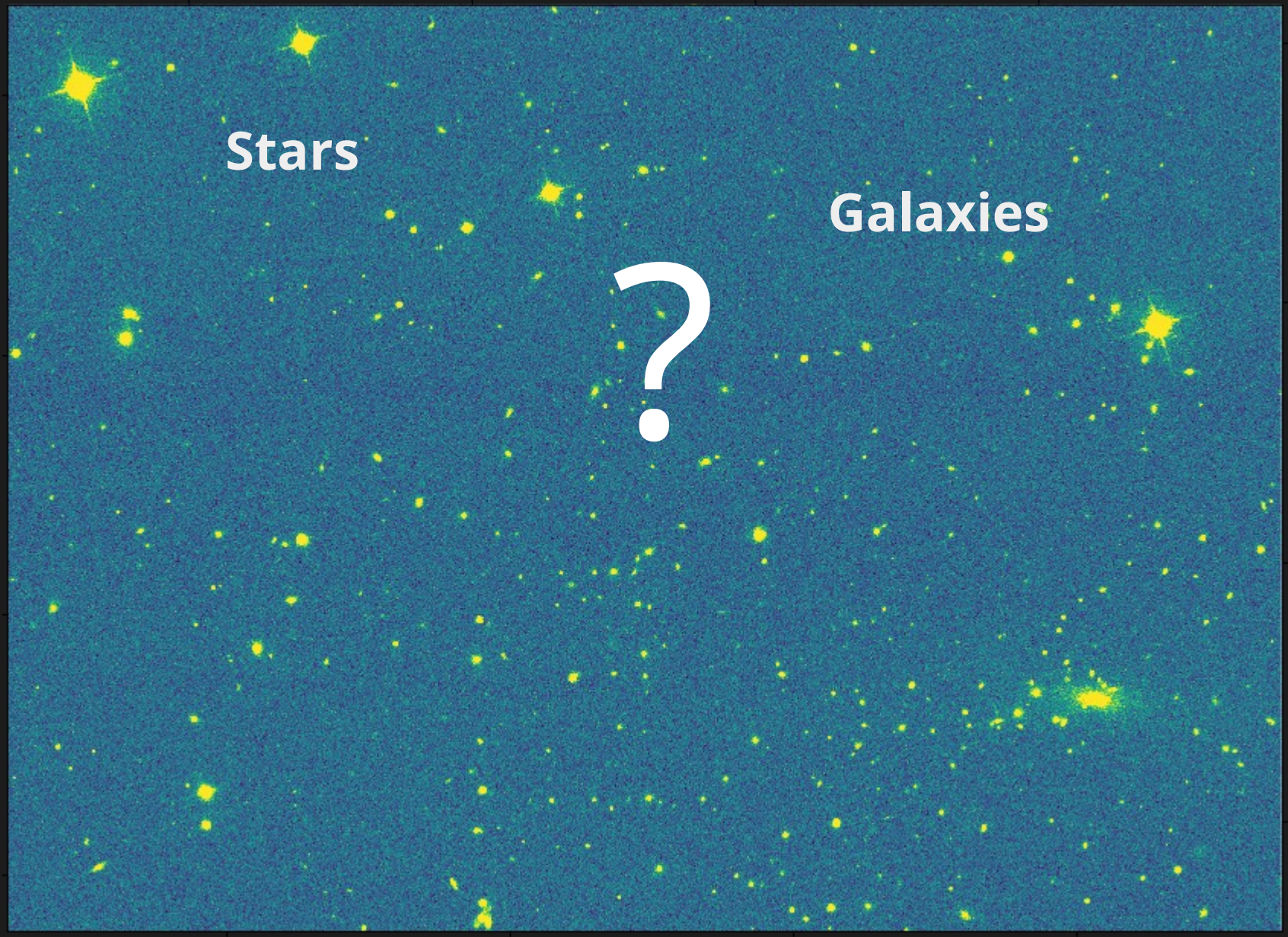
A *and-now-for-something-completely-different* Talk

Michael Mommert | Lowell Observatory

Motivation



Motivation



Motivation

Approach

- **Goal:** distinguish between stars/ellipticals/spirals from only a few photometric properties

Approach

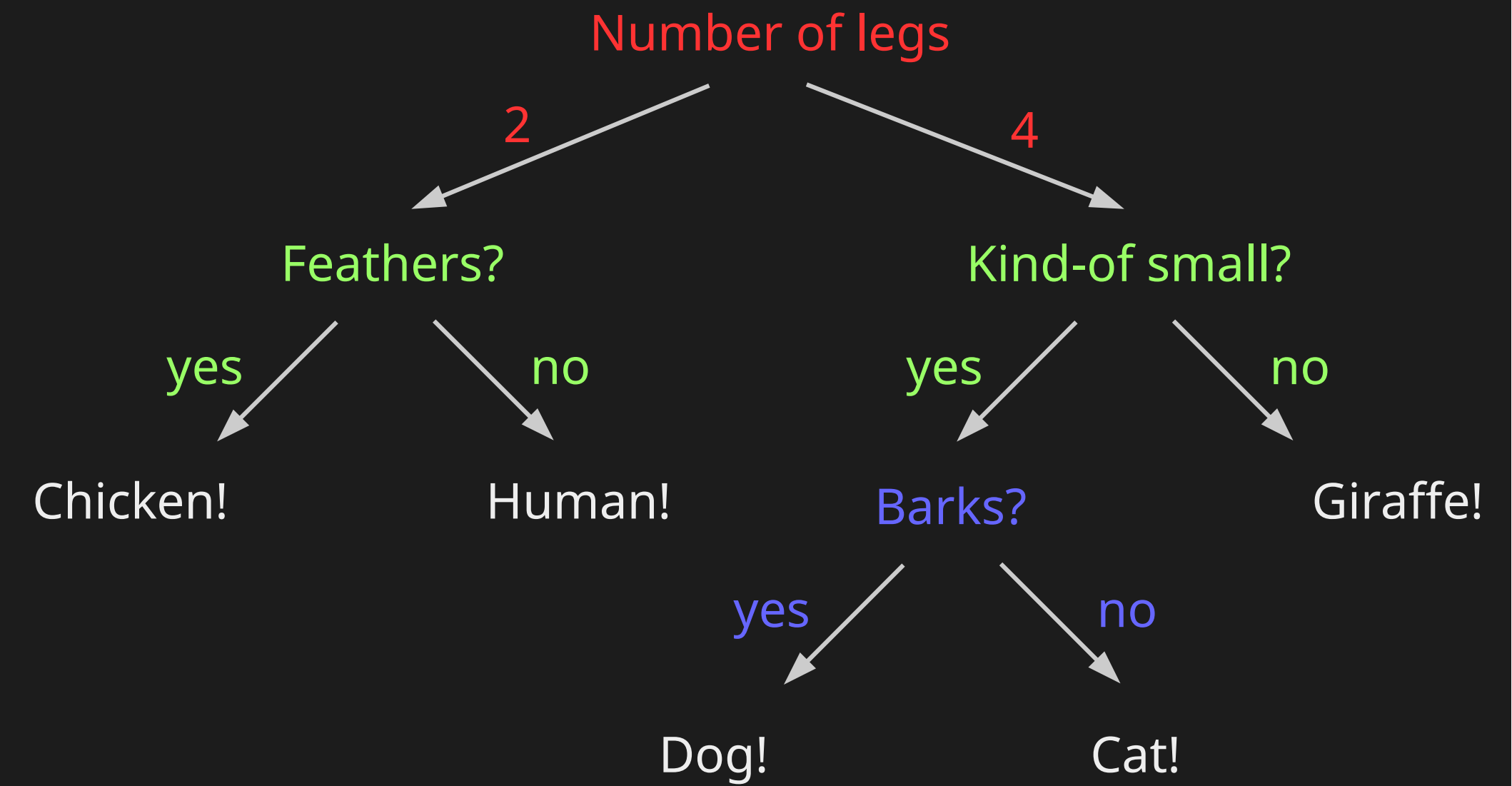
- **Goal:** distinguish between stars/ellipticals/spirals from only a few photometric properties

- **Training data**:
- 250k spiral/elliptical classifications from galaxyzoo.org
- + 180k stars
- = 430k sets of photometric properties from SDSS:
  - flavors: PSF, Petrosian, de Vaucouleurs, exponential
  - metrics: magnitudes, radii, axis ratios

Approach

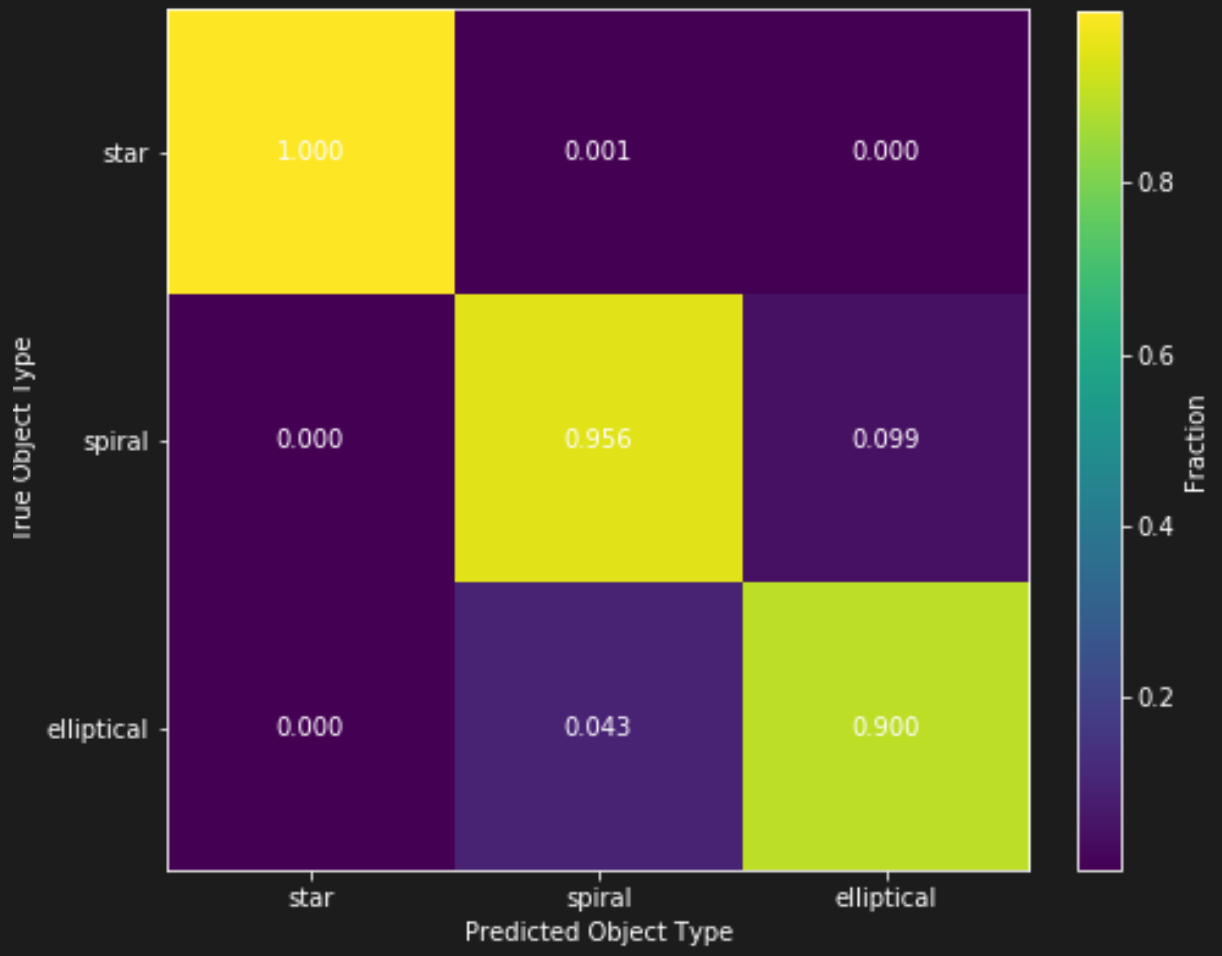
- **Goal:** distinguish between stars/ellipticals/spirals from only a few photometric properties
- **Training data**: 430k photometric data sets from SDSS
- **Method**:
  - Tried different supervised learning techniques
  - Cross validation across entire training set
  - **Winner**: a simple **Decision Tree**

What's a Decision Tree?

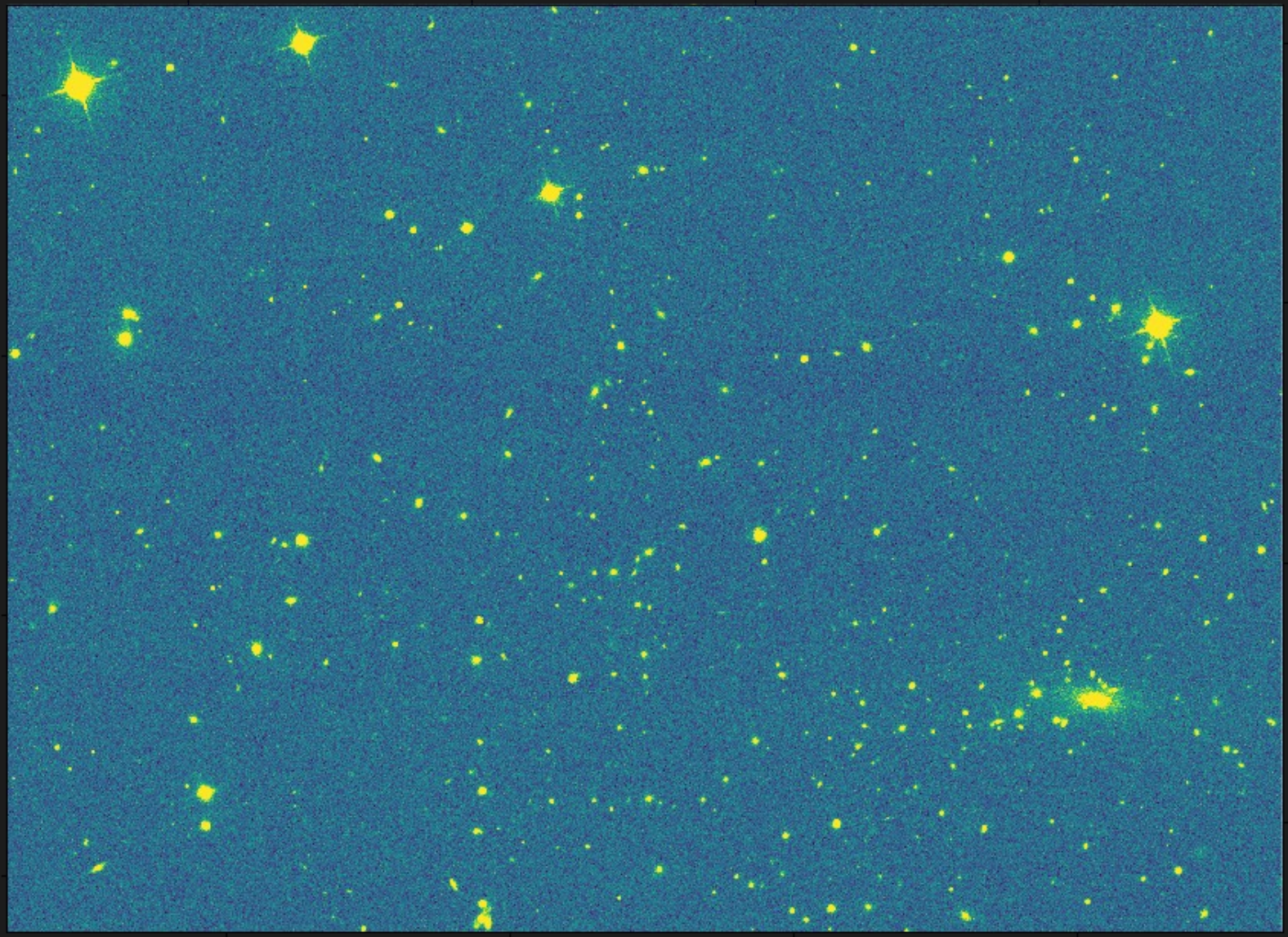


Results

- 96.4% overall accuracy – but this is misleading
- Confusion matrix:
  - Stars: 100%
  - Spirals: 95.6%
  - Elliptical: 90%
- Confusion among galaxies
- Stars are unambiguous

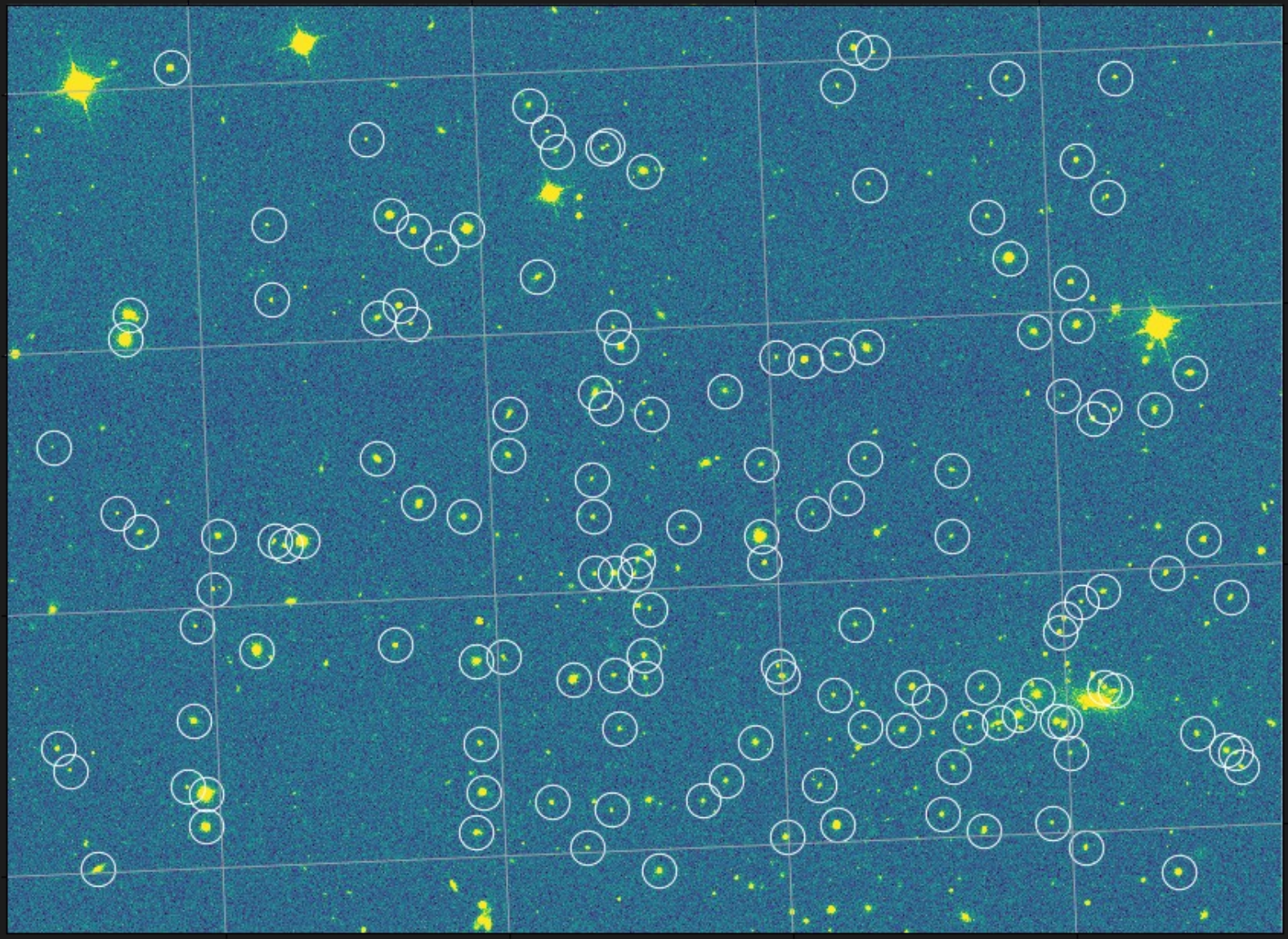


Example: Abell 1631



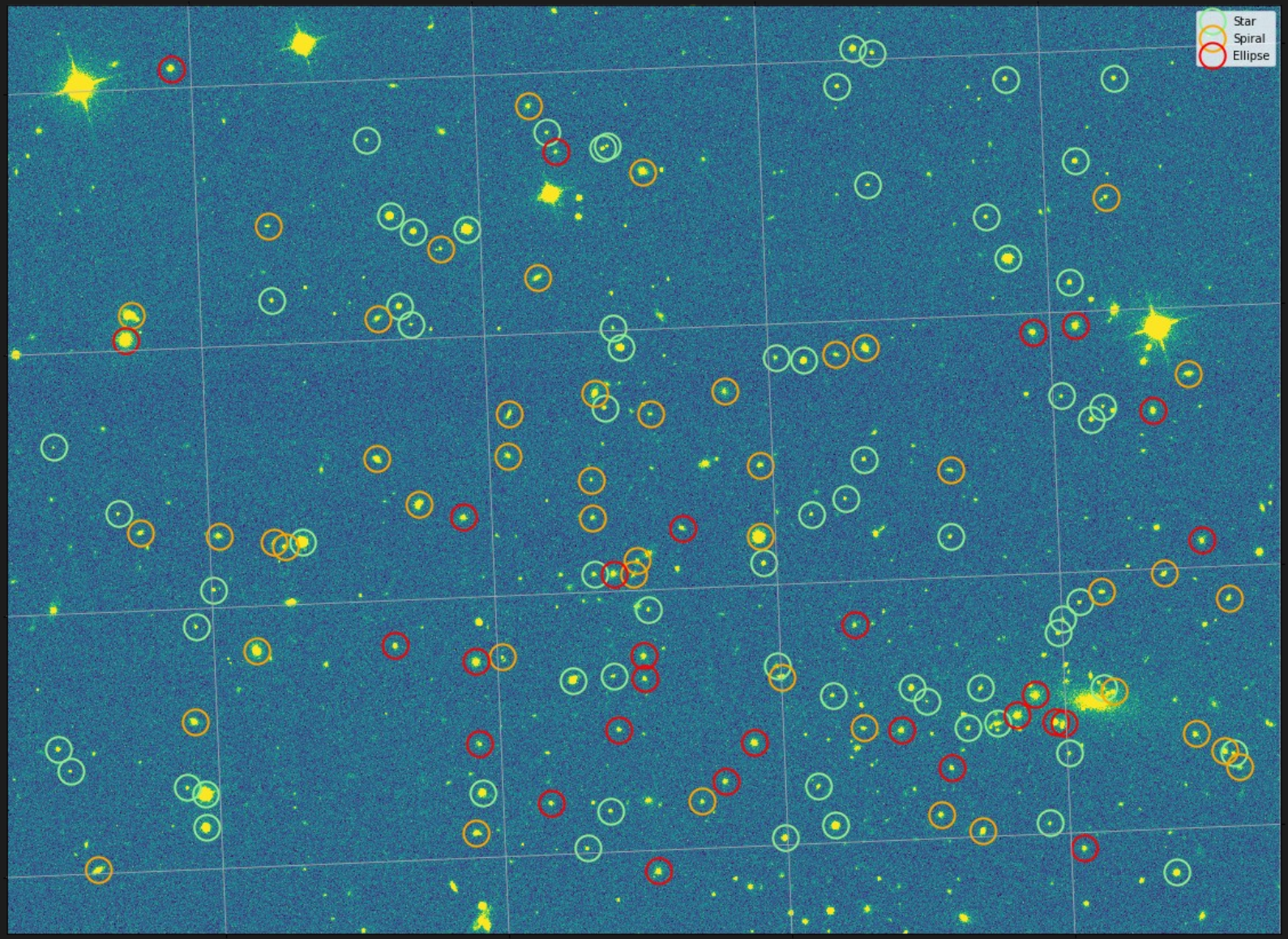
Retrieve SDSS image and photometric properties

Example: Abell 1631



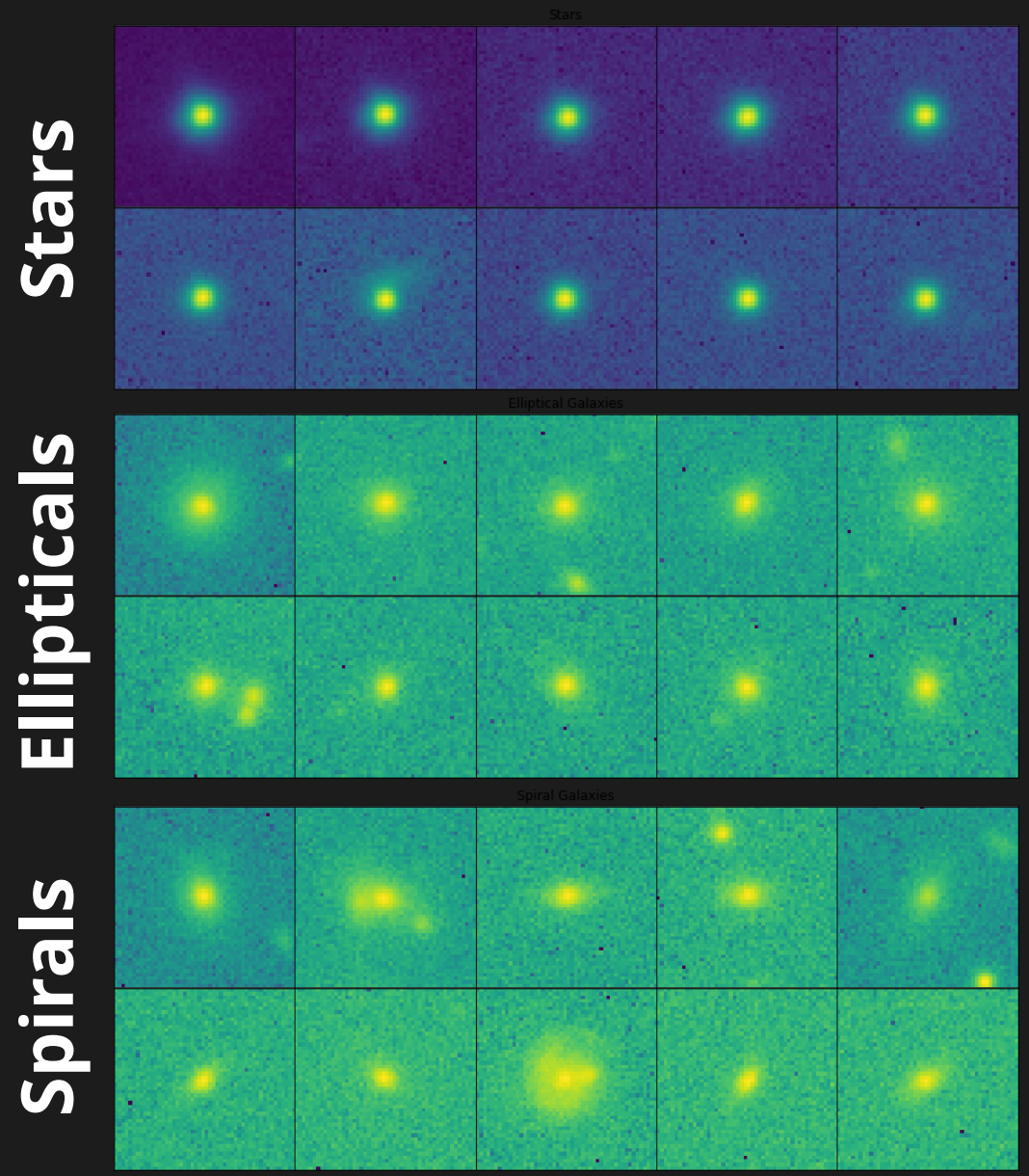
Rejecting unreliable detections and faint sources

Example: Abell 1631



Predict target type based on learned properties

Example: Abell 1631



Conclusions

- Simple toy model
- Room for improvements
- Don’t be afraid of machine learning!
- Notebook available: github.com/mommermi/sdss_stars_galaxies