Machine Learning Applications in Astronomy

Umaa Rebbapragada, Ph.D.
Machine Learning and Instrument Autonomy Group

Big Data Task Force
November 1, 2017

Jet Propulsion Laboratory
California Institute of Technology

Machine Learning (ML) for Astronomy

- Enabling Science
- Transient Science
- Catalog Science
- New Techniques
- Supporting ML in Astronomy

JPL ML-Astronomy Collaborations

Current and Past

- Palomar Transient Factory, intermediate Palomar Transient Factory, Zwicky Transient Facility
- The Very Long Baseline Array (VLBA) Fast Radio Transients Experiment (V-FASTR)
- Variables and Slow Transients (VAST) survey, part of Australian Square Kilometre Array Pathfinder (ASKAP)
- RealFast project at the Very Large Array (VLA) radio telescope
- MIT Lincoln Lab collaboration on Space Surveillance Telescope (SST) data
- Dark Energy Survey

Enabling Science

Big Telescopes, Big Science, Big Data

Large Synoptic Sky Telescope (LSST)
15 TB/night

Square Kilometre Array
160 TB/second

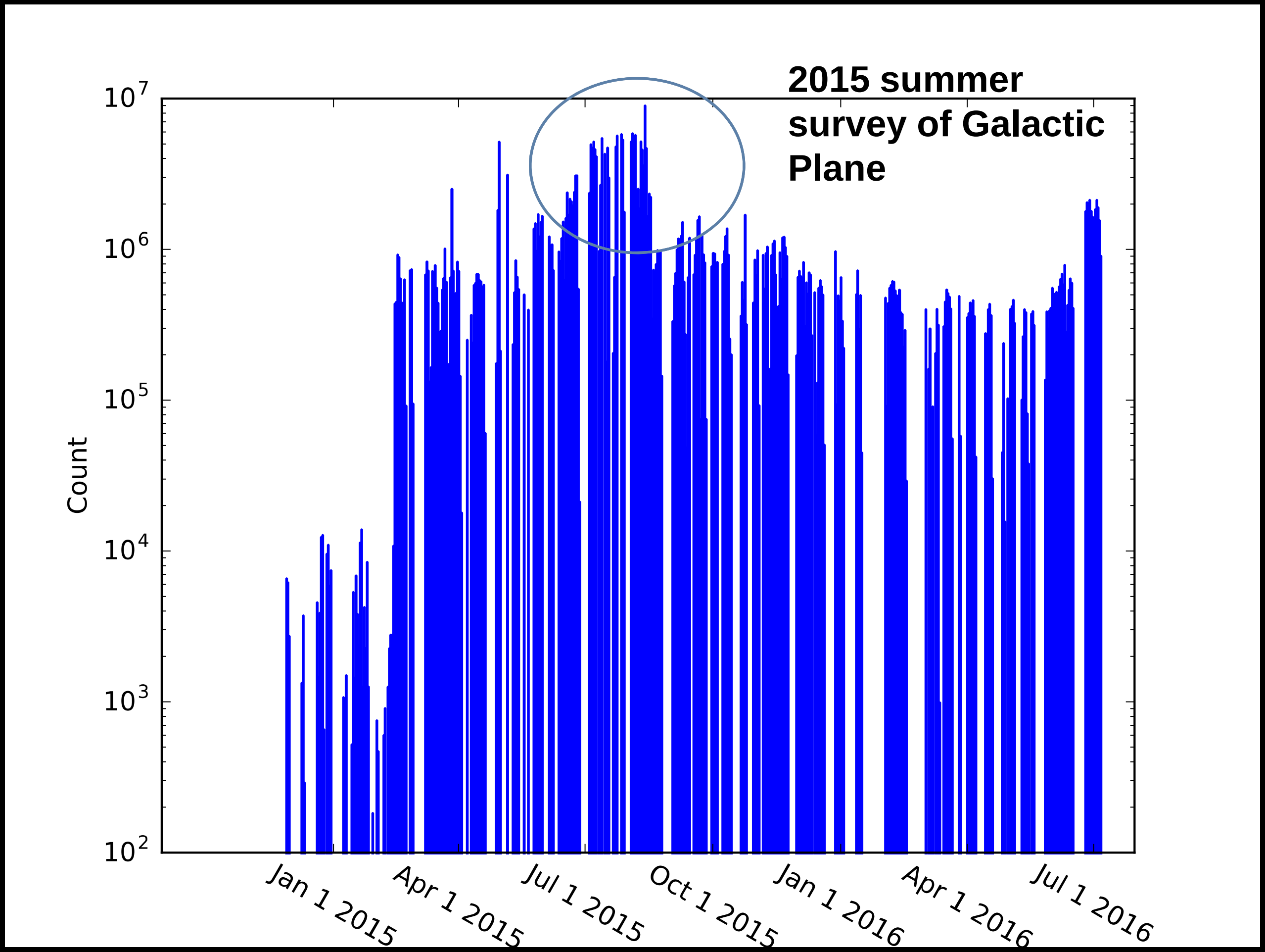
James Webb Space Telescope (JWST)

Wide-field Infrared Survey Telescope (WFIRST)

Transiting Exoplanet Survey Satellite (TESS)

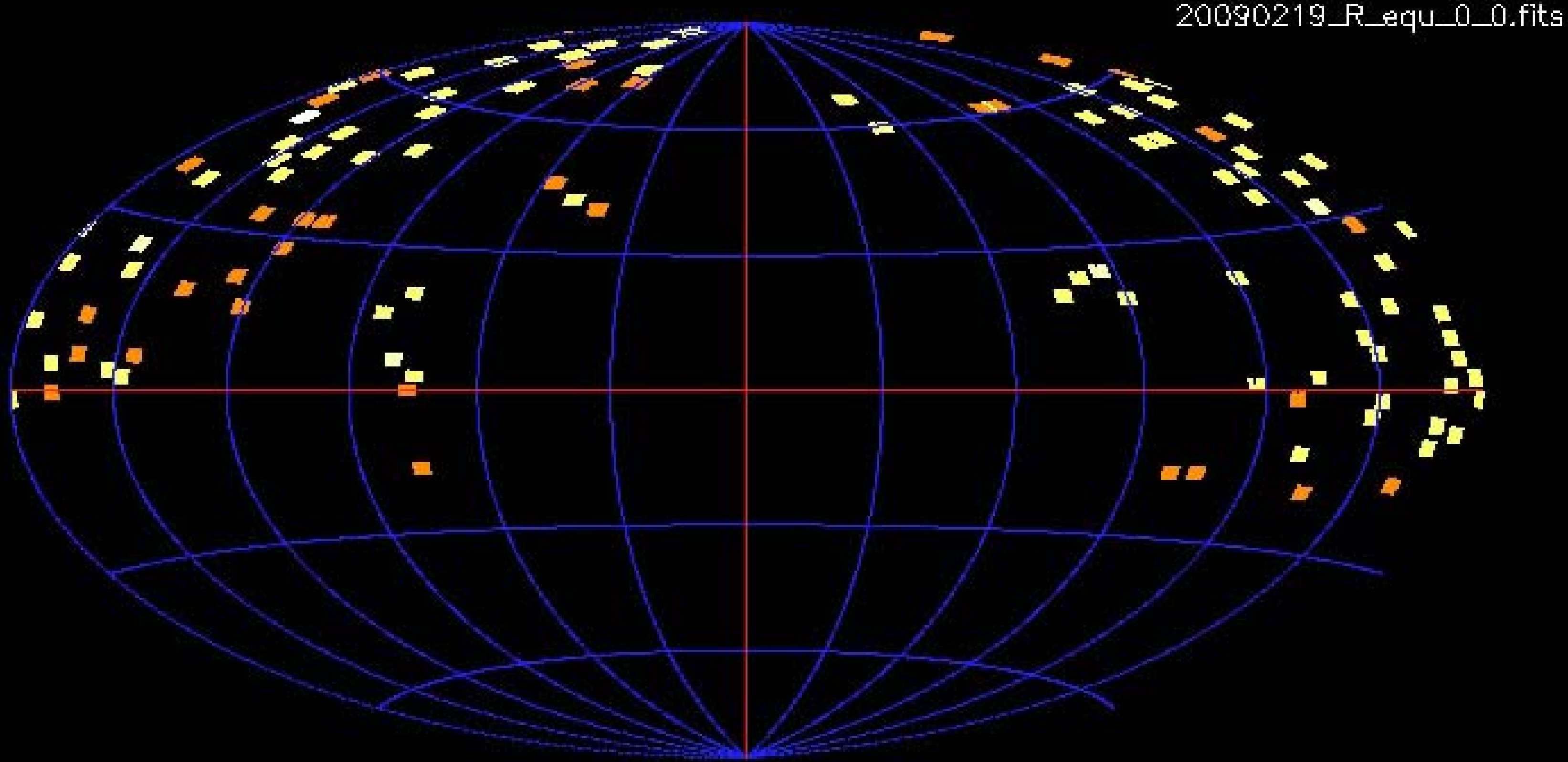
Millions of Detections per Night…in 2015

intermediate Palomar Transient Factory 2015-16

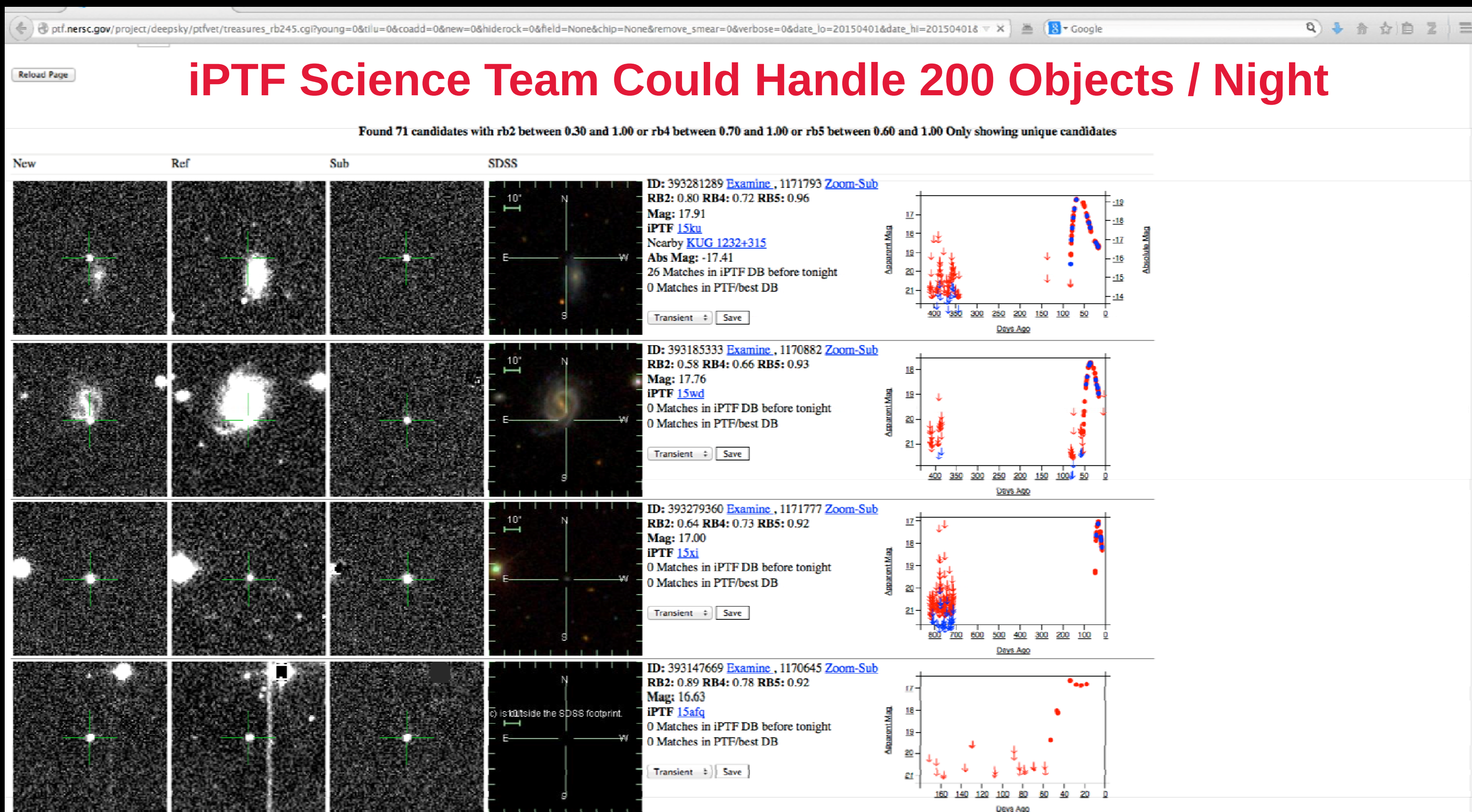


Huge Catalogs from All-Sky Surveys

Observations of millions of objects, including spectra and lightcurves

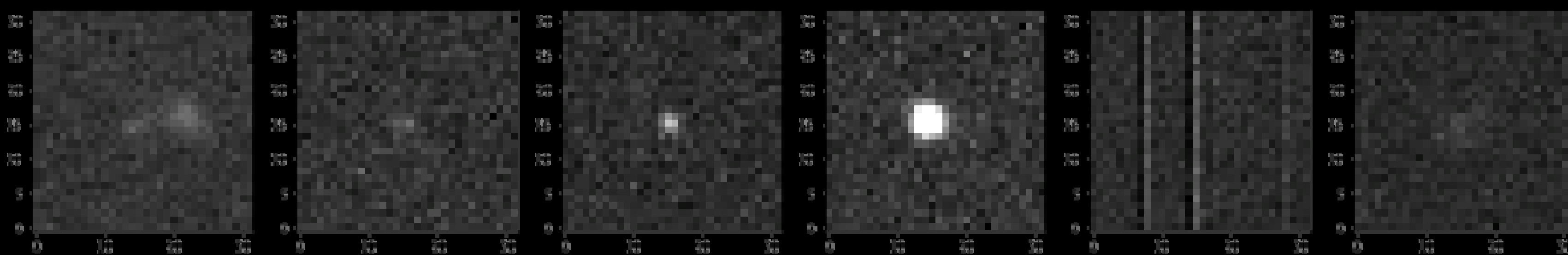


Science Team Size is Constraining Factor

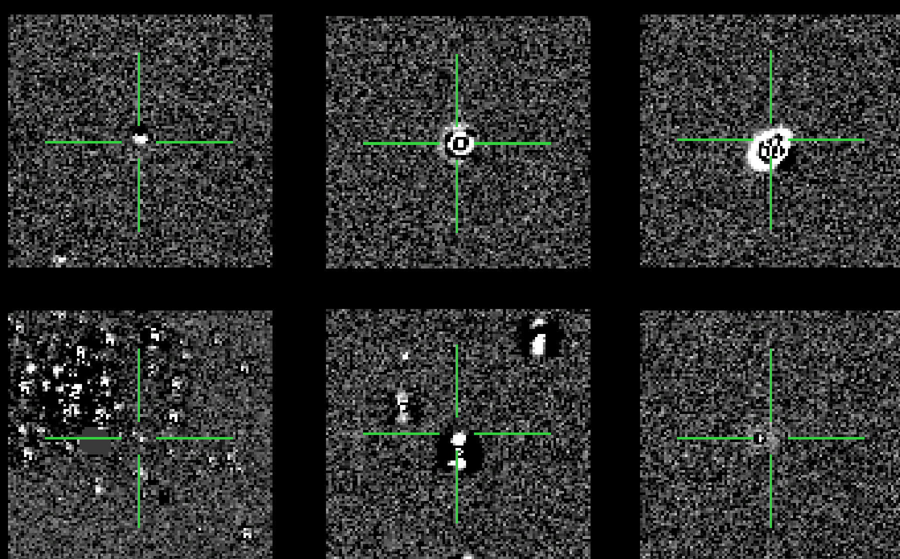


Machine Learning Classifiers Filter False Detections

Are these candidates optical or pipeline artifacts?



Artifacts of Source Extraction



Artifacts of
Image Differencing

Transient Science

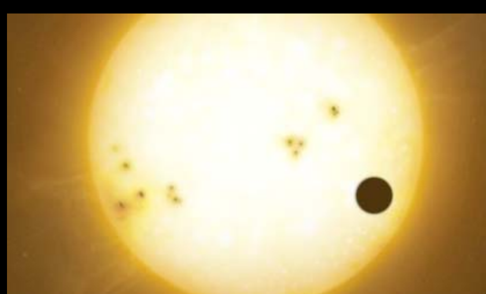
Astronomical Transients

- Explosive events, very short duration

Gamma Ray Burst
(milliseconds to hours)

Supernova
(weeks to months)

- Transience from our observational perspective



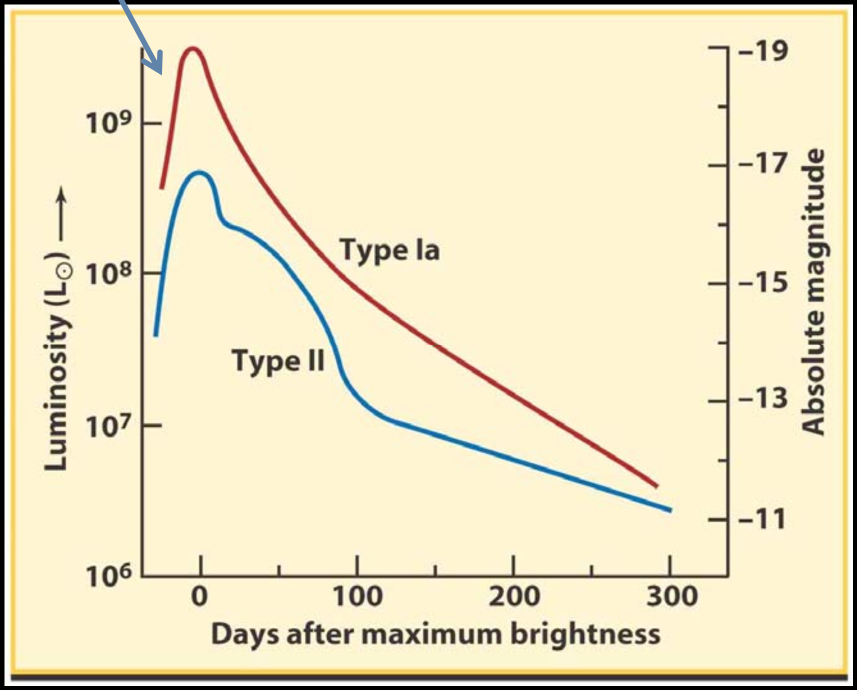
Planetary transit

Asteroid

Transient Science Requires Real-time Algorithms

- Find all scientifically-interesting observations
- Filter all irrelevant observations
- Goal: trigger follow-up of science-rich targets

Ideally, find it here

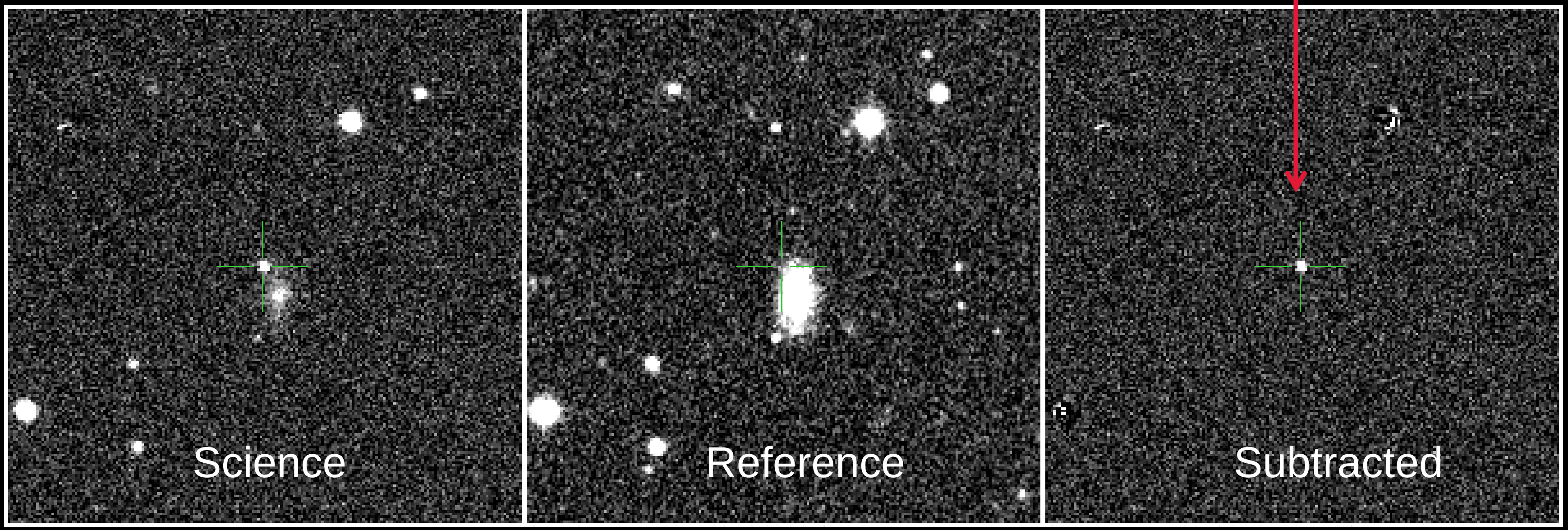


Real-time Filtering

Optical Astronomy Example from iPTF

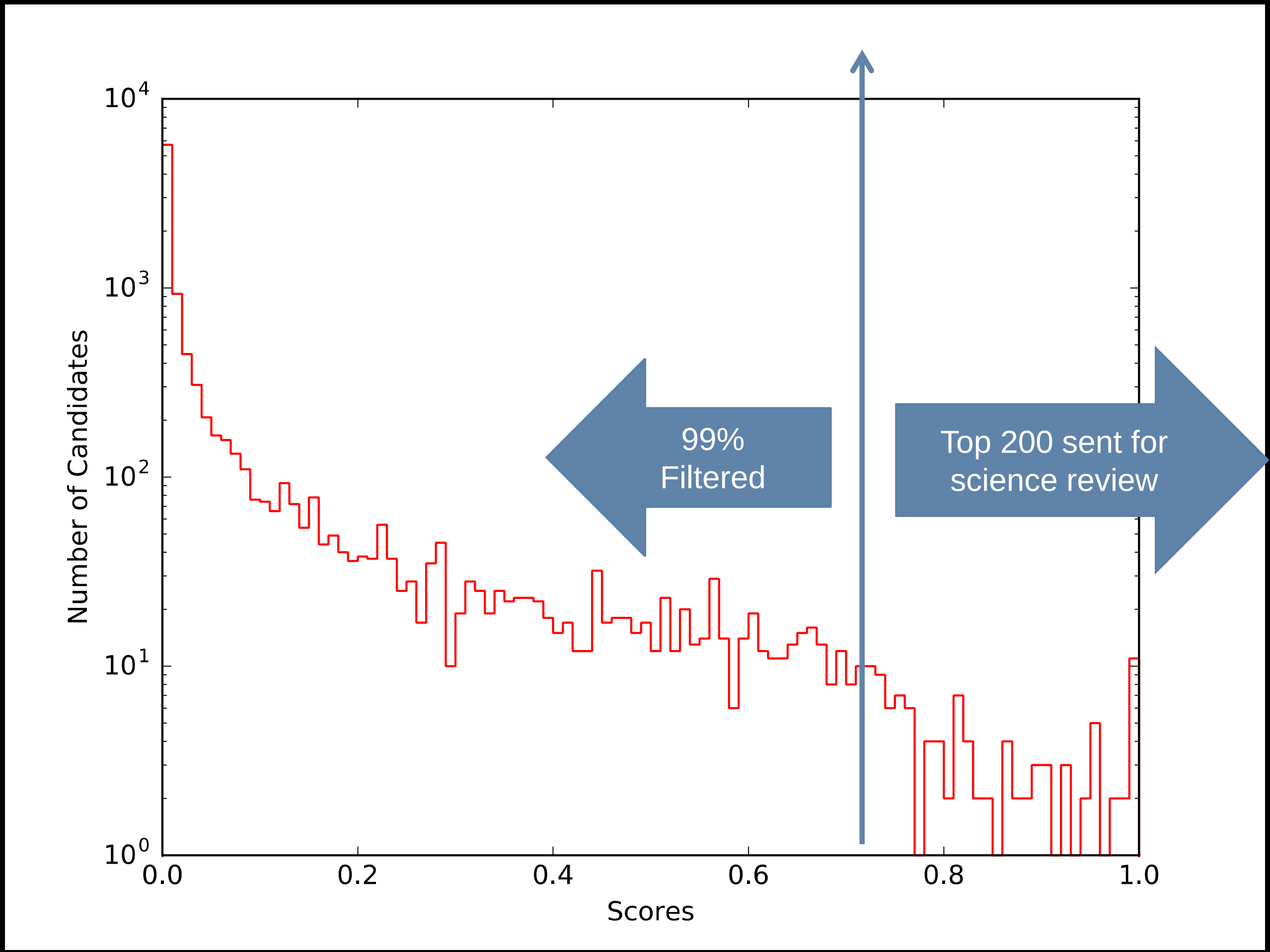
- Automated classifier scores candidates from 0 (bogus) to 1 (real)

Is this candidate real or bogus?



Real-time Filtering

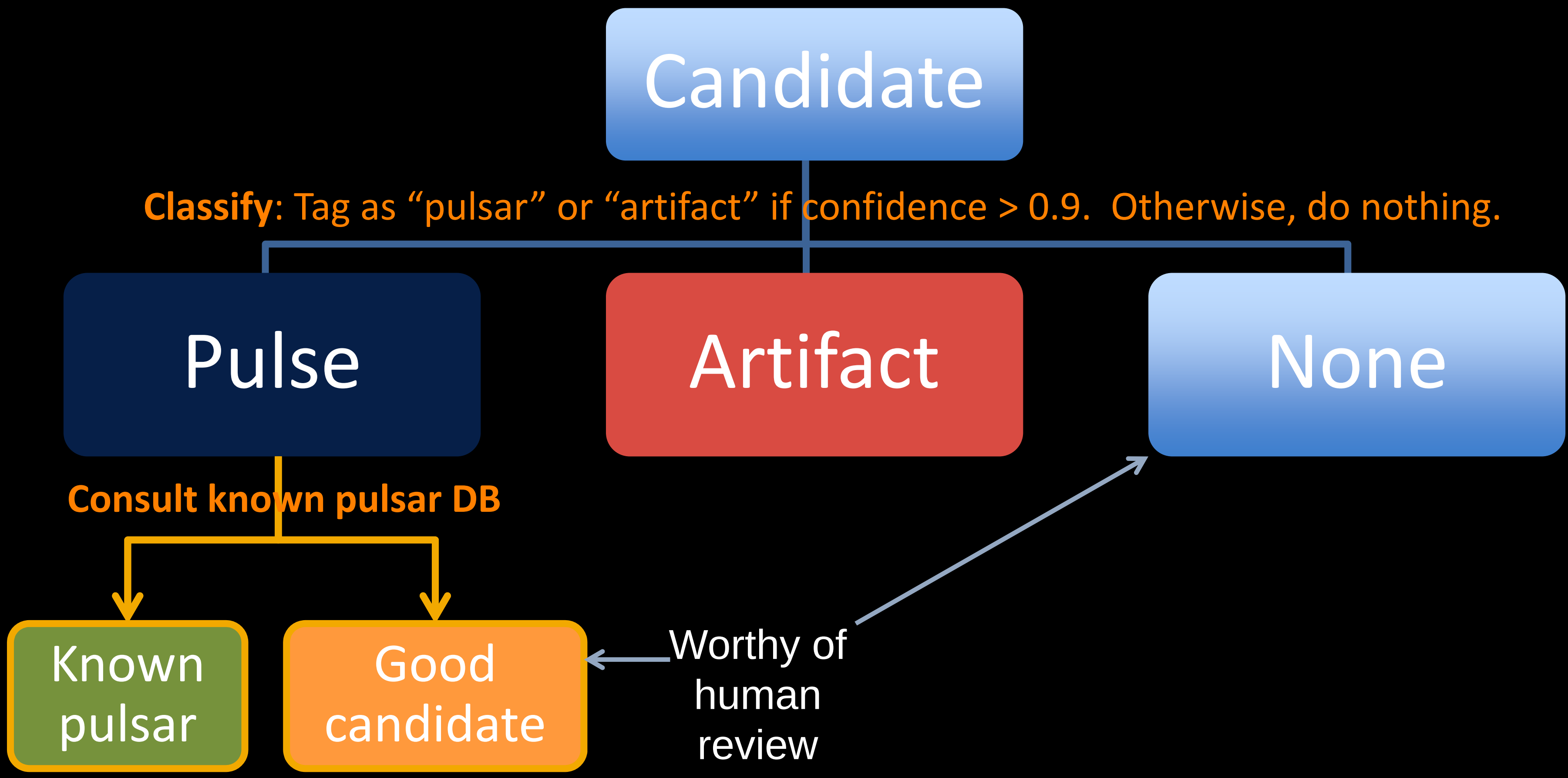
Optical Astronomy Example from iPTF



Set decision threshold that filters 99%

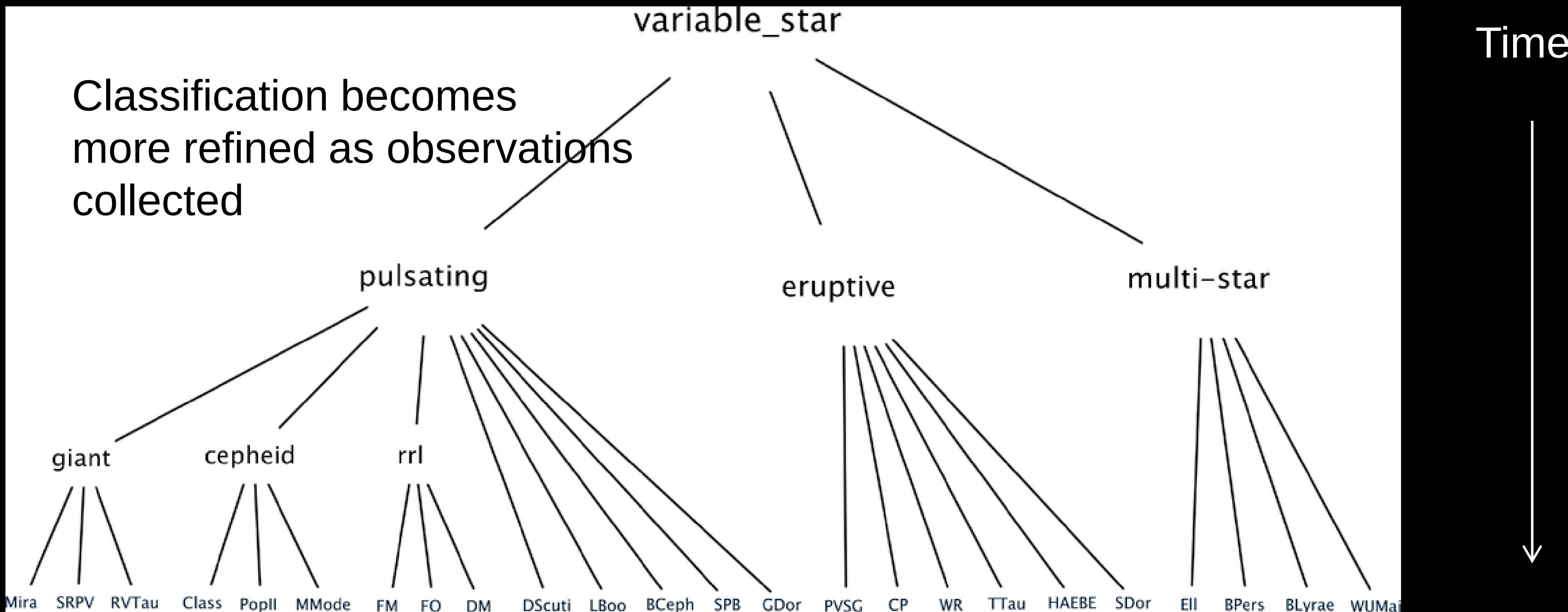
Real-time Filtering

Radio Astronomy Example from V-FASTR / VLBA



Real-time Classification

- Modern “data brokers” perform real-time classification of objects



Richards, et al. 2011, ApJ

- Science users contribute filters to downselect the information they want

Catalog Science

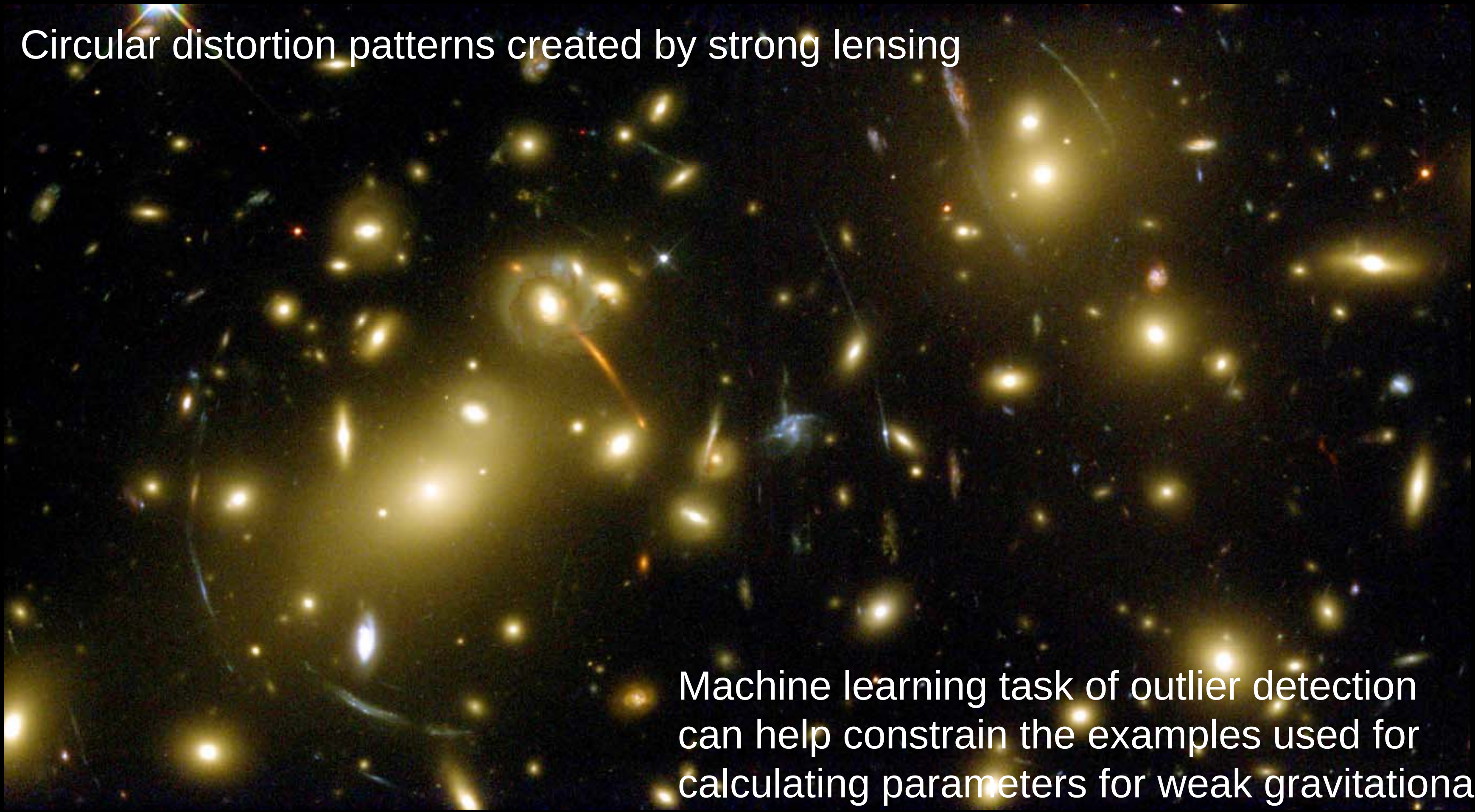
ML Applications

- Stellar classification
- Star / galaxy separation
- Identification of planetary transits (exoplanets)
- NEO identification / tracking
- Estimating cosmological parameters

Mining Astronomical Archives

Weak Lensing Archives

Circular distortion patterns created by strong lensing

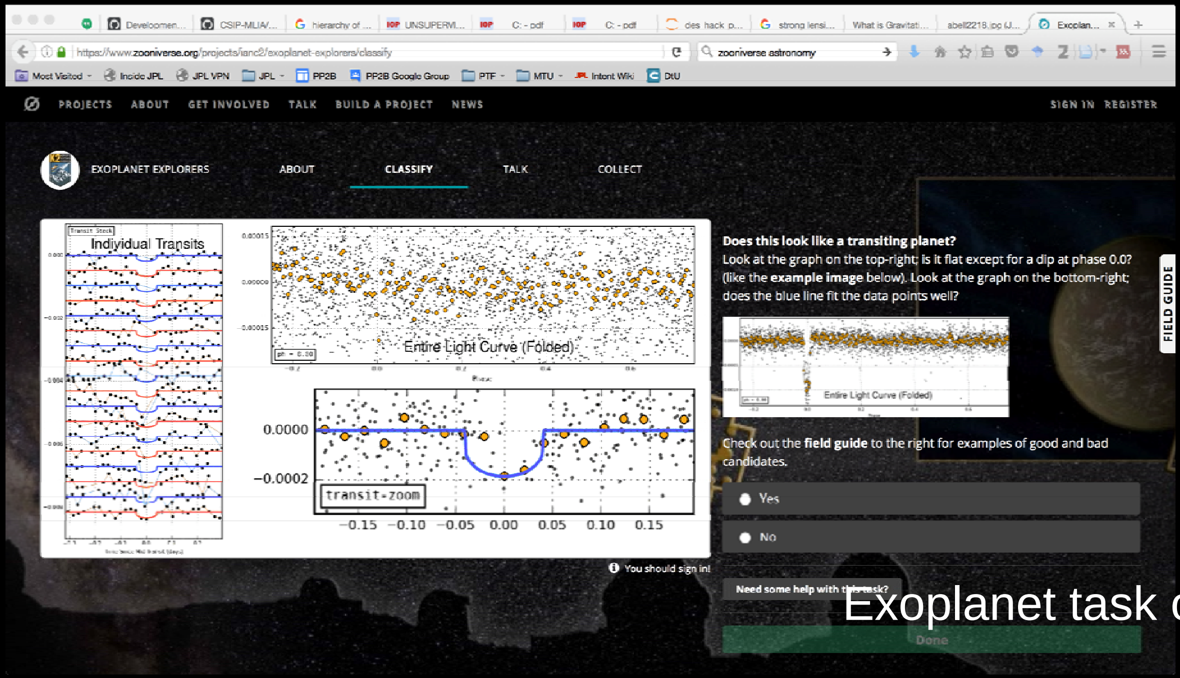


Machine learning task of outlier detection can help constrain the examples used for calculating parameters for weak gravitational lensing.

New Techniques

Crowdsourcing + ML

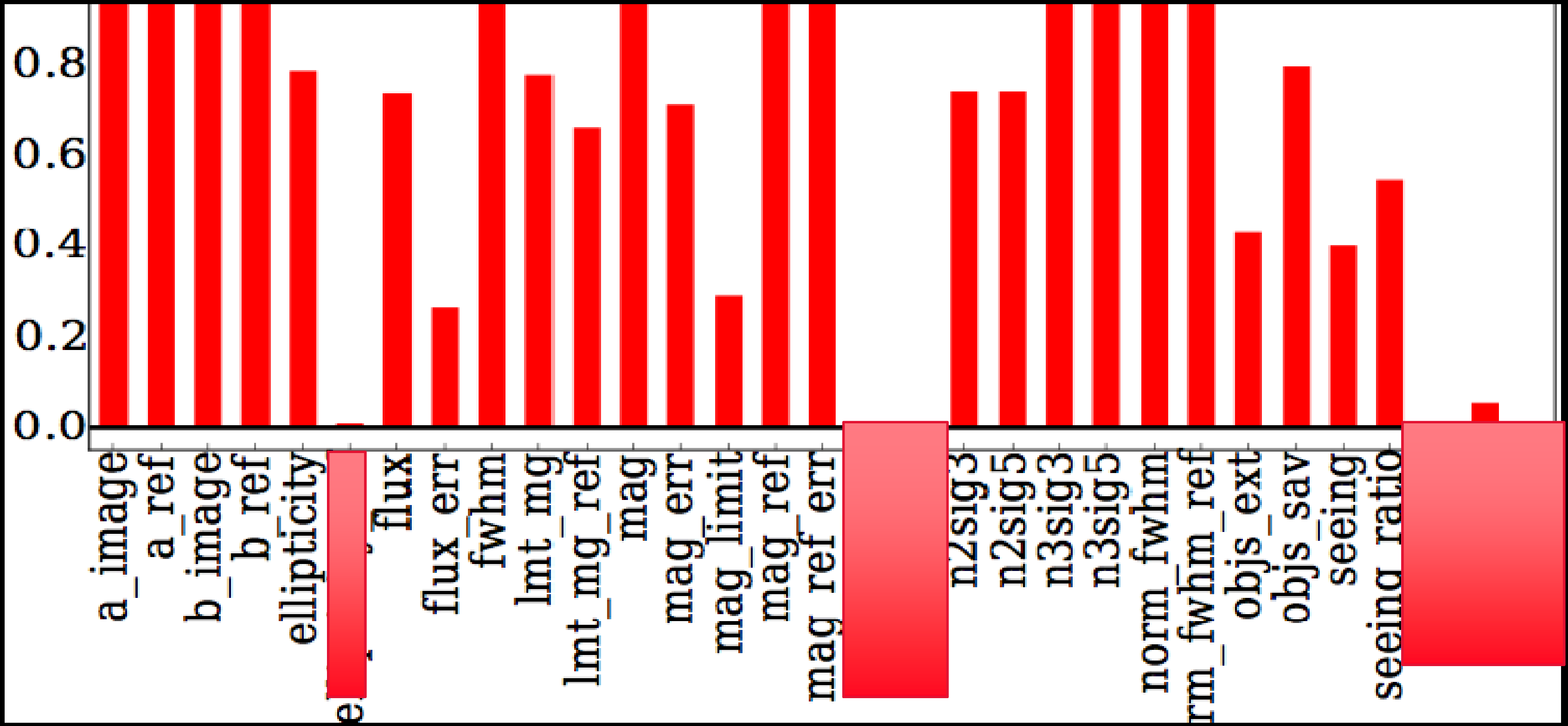
- Crowdsourcing platforms offer interfaces that integrate with ML algorithms and build training data quickly
- Allow interaction with scientists, general public



Exoplanet task on Zooinverse

Responding to Instrument and Survey Changes

- Example of how a pipeline upgrade changed data characteristics at iPTF

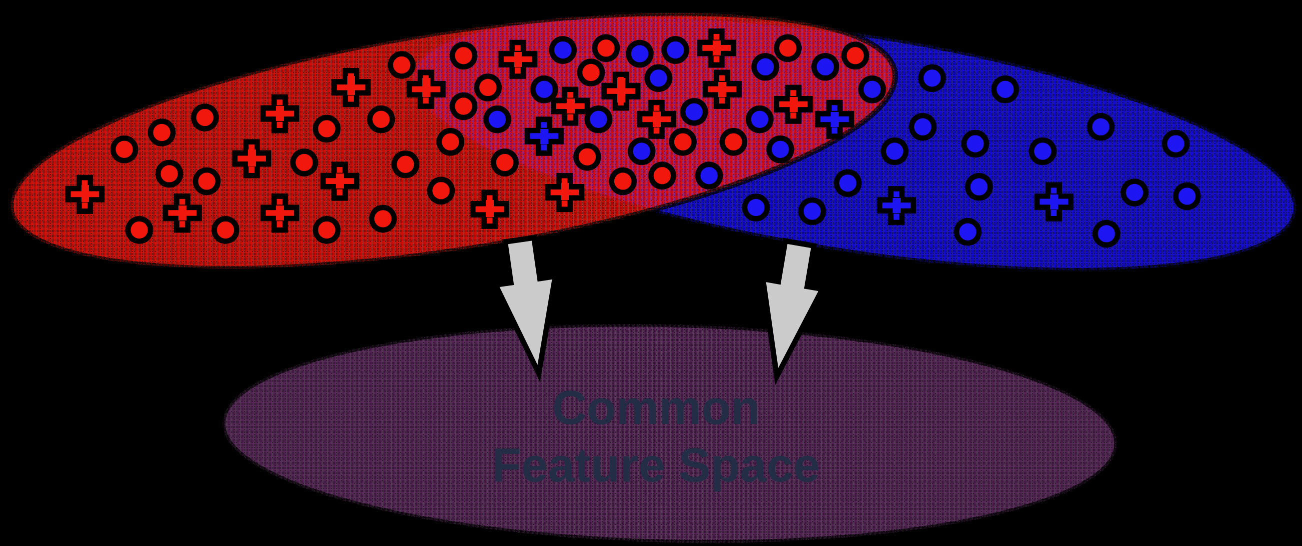


Domain Adaptation

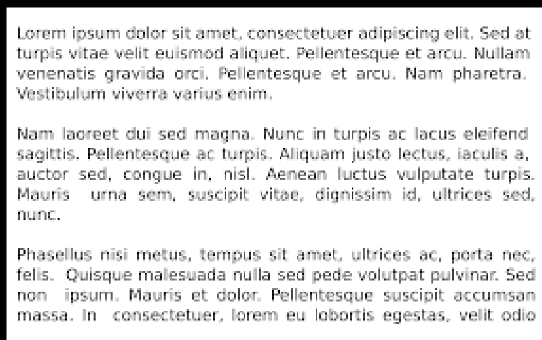
- Computes mapping between source and target data sources that share common science goals
- Continuous data record between old and new missions

Old Instrument, pre-upgrade

New Instrument, post-upgrade



Deep Learning


Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Sed at turpis vitae velit euismod aliquet. Pellentesque et arcu. Nullam venenatis gravida orci. Pellentesque et arcu. Nam pharetra. Vestibulum viverra varius enim.

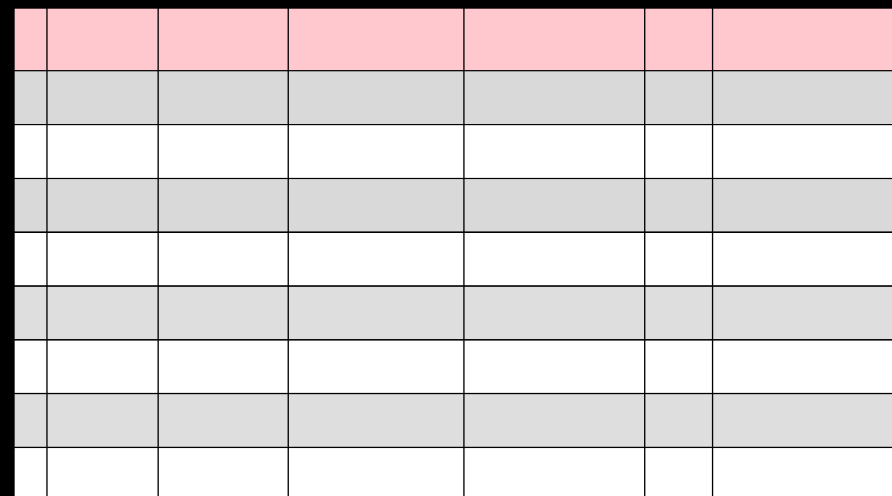
Nam laoreet dui sed magna. Nunc in turpis ac lacus eleifend sagittis. Pellentesque ac turpis. Aliquam justo lectus, iaculis a, auctor sed, congue in, nisl. Aenean luctus vulputate turpis. Mauris urna sem, suscipit vitae, dignissim id, ultrices sed, nunc.

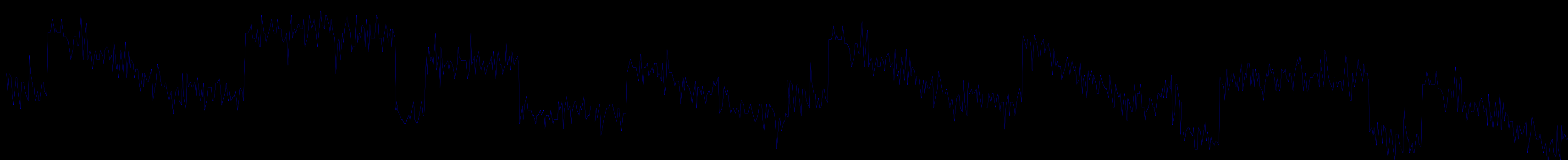
Phasellus nisi metus, tempus sit amet, ultrices ac, porta nec, felis. Quisque malesuada nulla sed pede volutpat pulvinar. Sed non ipsum. Mauris et dolor. Pellentesque suscipit accumsan massa. In consectetuer, lorem eu lobortis egestas, velit odio

Bag of Words
TFIDF

Pixel values,
SIFT, HoG,
histograms of visual words

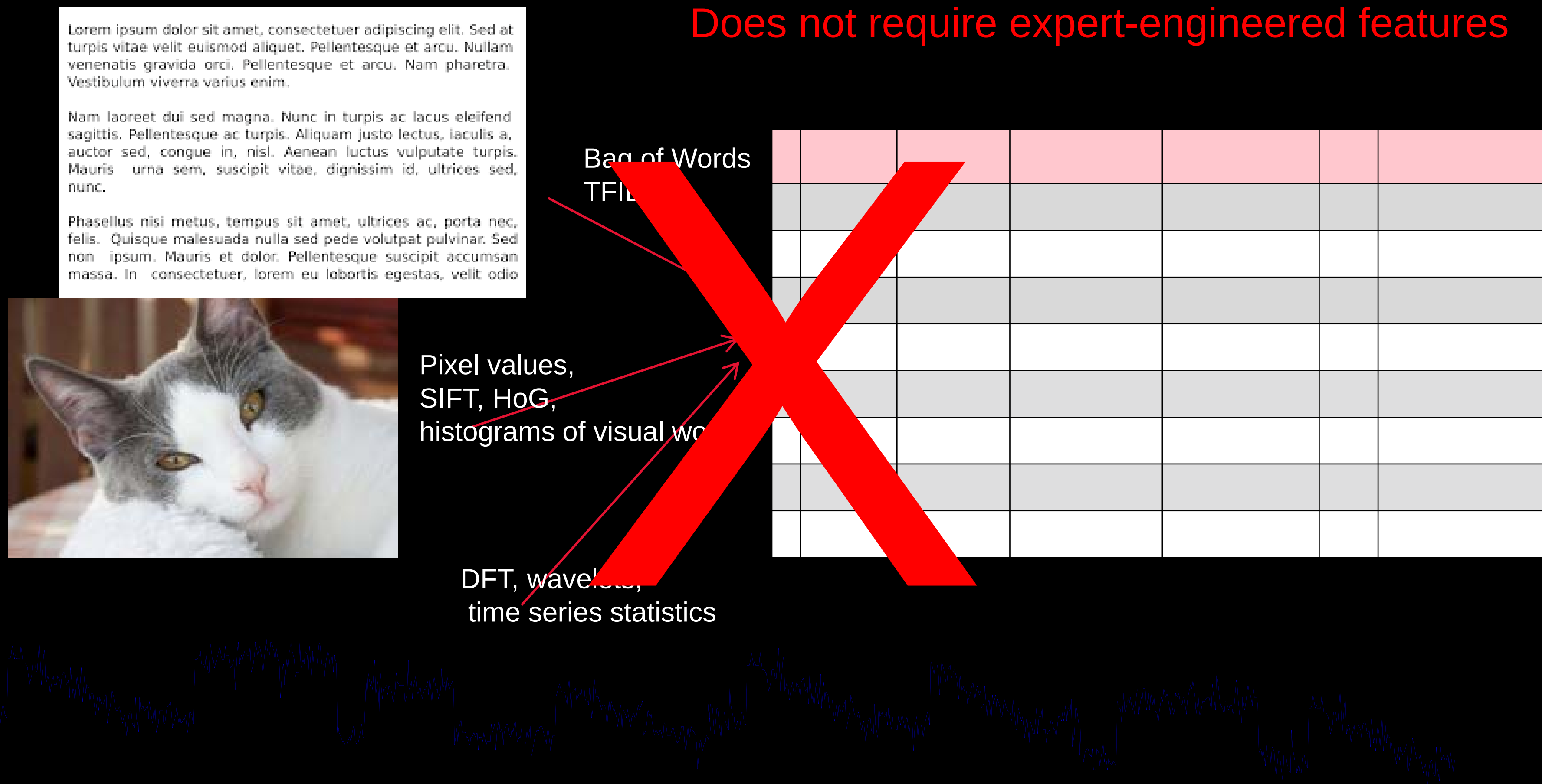
DFT, wavelets,
time series statistics





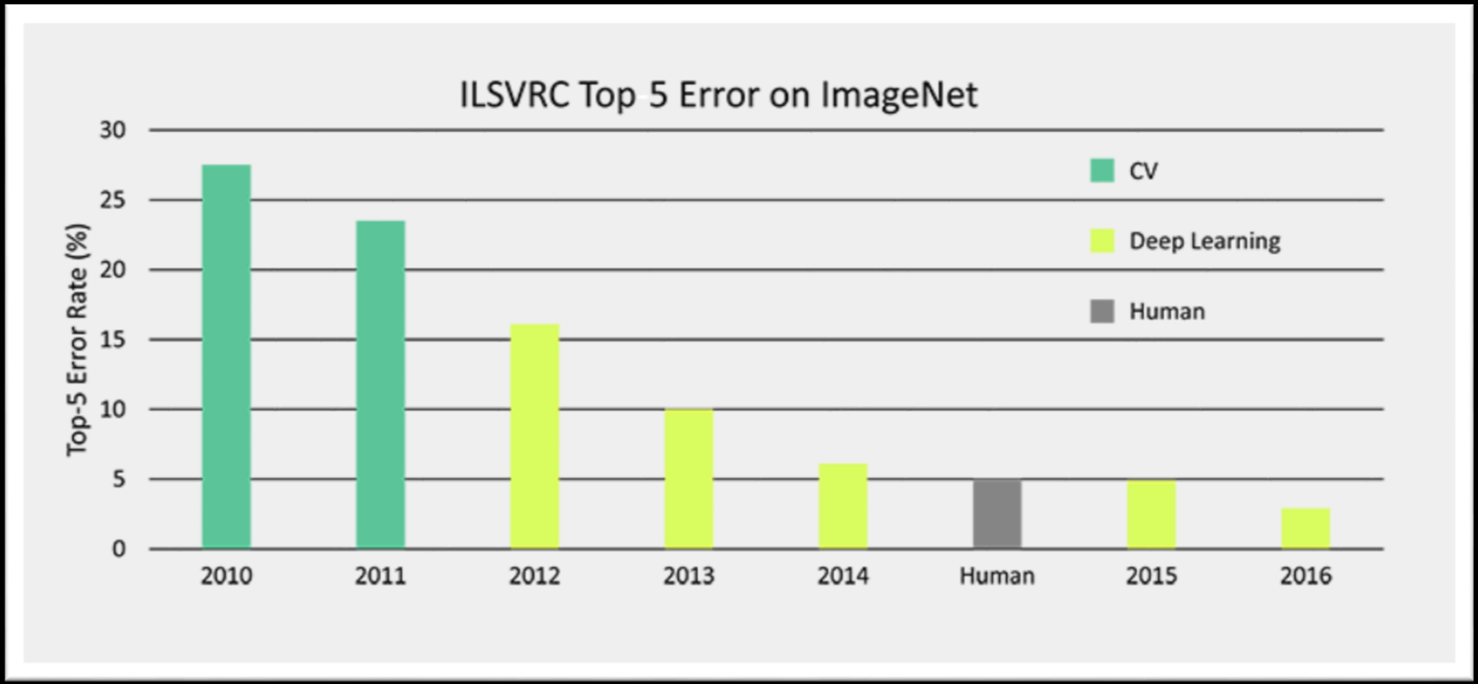
Deep Learning

Does not require expert-engineered features



Deep Learning

ImageNet Large Scale Visual Recognition Challenge (ILSVRC)



https://www.dsiac.org/resources/journals/dsiac/winter-2017-volume-4-number-1/real-time-situ-intelligent-video-analytics

Supporting ML in Astronomy

Challenges and Opportunities

- Programmatic
  - Space telescope managers don’t know we exist
  - ML experts not involved in data pipeline requirements and design
- Financial
  - Few ROSES opportunities that support ML work in astronomy
  - Limited or no budget for high level data products produced by ML
- Cultural
  - “It’s just software.”
  - “My post-doc can do it.”

Jet Propulsion Laboratory
California Institute of Technology

jpl.nasa.gov