Adaptive Learning for SN photometric classification

EWASS – Prague, June/2017

Emille E. O. Ishida

Laboratoire de Physique de Clermont - Université Clermont-Auvergne
Clermont Ferrand, France
Spectroscopy
X
Photometry
Spectroscopy x Photometry

Redshift – from host galaxy

Supernovae Typing

J C Wheeler and R P Harkness, 1990
How are spectroscopic sets constructed?
How are spectroscopic sets constructed?

Take spectra for learning and determine everything else
As a consequence ...

Training set is a highly biased subset of the target set.
The problem with text-book ML: Representativeness

Performance in validation set does not guarantee results in target set
Representativeness

Redshift – from host galaxy

SDSS spec
SDSS photo

Beck, Lin, Ishida et al., 2017

Supernovae Typing

SNPCC data, Kessler et al., 2010 – plot by Ishida
Pro-active approach to the construction of spectroscopic samples
Active Learning

**Step 1**
Standard Classification

**Step 2**
Identify problematic as well as representative elements

**Step 3**
Make a query

**Step 4**
Classify

Class 1
Class 2
Simple logistic regression + AL

- Efficiency
- Purity
- Form
- % Correct Classifications

Graphs showing the performance metrics for different sampling methods:
- Active Learning
- Random Sampling
- Passive
Alternative approach

Landmark selection
Alternative approach

Landmark selection
Alternative approach

Landmark selection
Alternative approach

Landmark selection
Alternative approach

Landmark selection + Active Learning
Alternative approach

Landmark selection + Active Learning
Sometimes,

There will be a group without a minimum necessary number of labels...
Sometimes,

There will be a group without a minimum necessary number of labels…

Ask!
Automated Supernova Ia Classification Using Adaptive Learning Techniques

Kinjal Dhar Gupta*, Renuka Pampana*, Ricardo Vilalta*, Emille E. O. Ishida†, Rafael S. de Souza†

supervised classification

Dhar Gupta et al. (incl. Ishida), 2016 IEEE Symposium in Computational Intelligence, Greece
But SN fade away...
But SN fade away...

1. Use simulations to identify the ideal spectroscopic sample:
   - redshift range
   - errors
   - colors
But SN fade away...

1. Use simulations to identify the ideal spectroscopic sample:
   - redshift range
   - errors
   - colors

2. Use partial light curves (pre-max)
But SN fade away...

1. Use simulations to identify the ideal spectroscopic sample:
   - redshift range
   - errors
   - colors

2. Use partial light curves (pre-max)
https://github.com/COINtoolbox
The COIN Residence Program - CRP

Annual meetings

Conference

Workshop

Hackathon
The COIN Residence Program - CRP

Annual meetings

Conference

Workshop

Hackathon
The COIN Residence Program - CRP

A non-profit start-up?

Annual meetings

John Johnson/HBO

The COIN Residence Program - CRP

A non-profit start-up?

Annual meetings

John Johnson/HBO

CRP #2, UK, 2015

CRP #3, Budapest, 2016

COIN Residence Program #4

20-27 August 2017
Clermont Ferrand, France
Extra slides
Logistic regression

ROC curve for kPCA (area = 0.83)

ROC curve for Full lc (area = 0.66)
**Background:** Active Learning in Astronomy

ACTIVE LEARNING TO OVERCOME SAMPLE SELECTION BIAS: APPLICATION TO PHOTOMETRIC VARIABLE STAR CLASSIFICATION
