Point Source Classification on Astronomical Photometric Images Using Artificial Neural Network

R. Yoyponsan, U. Sawangwit, S. Komonjinda
Outline

- Introduction
- Data
- Supervise Learning
  - Object Detection & Feature Extraction
  - Artificial Neural Network
- Analysis
  - K-Fold Crossvalidation
  - Confusion Metrics and Scores
- Next
- Acknowledgement
- References
GRavitational-wave OpticAl TRANSient Observer
University of Warwick – Monash University
Armagh Observatory – University of Sheffield – University of Leicester – NARIT
http://www.goto-observatory.org
This is Big Data

Phase 2:

48 Mpx camera per scope

1 frame per telescope every 2 minutes

3 frames per pointing

3600 sq. deg per night

2×10^5 sources per pointing

20,000 sq. deg observable sky

100 pointings per night

Whole sky every week

2×10^7 sources per night

10^8 sources updated each week

And all data needs to be processed close to real-time!
Data

- Simulated - SkyMaker (courtesy of Dr. James Mallaney, U. of Sheffields)
- Imitate real image
- Contains point and extended sources randomly distributed
8057 objects
1752 objects of class 0 (extended sources, others)
6305 objects of class 1 (point sources)
Refered to catalog of the simulated image
Produced from 1 image as of now
Features/ Attributes

- attributes/features = values in the data fields describing the properties of each object (Ball & Brunner 2009)
- Different classes have different properties

Point Source               Extended Source               Flare
Detection and Feature Extraction

- Source Extractor (Bertin & Arnouts 1996)
- Robust, fast, support batches analysis
- Influenced by DAOPHOT (Stetson 1987)
- Determine centroid using threshold above certain background level, and in this work we use 2sigma
Feature Extraction

- Rely on source extractor to extract some feature, i.e., semi-major / minor axis, ellipticity, instrumental magnitude.

- Astropy model fitting of 10x10 region around the centroid to models, such as; 2d airry disk profiles, Moffat profiles, gaussian profile, sersic profile.
Artificial Neural Network

- Imitate biological neural network
- Universal function approximator
- Backward propagation and gradient descent algorithm
K-fold Cross validation

- Split sample to k subset pick one as a test set
- Combine the rest as training set
- Alternate the test set until complete k iteration
- Ensemble the prediction, select the final result by voting
## Result

### Confusion Metrics

|       | Class 0 | Class 1 |
|-------|---------|---------|
| Class 0 | 674     | 1078    |
| Class 1 | 841     | 5464    |
## Result

| Scoring    | Calculation | Score Value |
|------------|-------------|-------------|
| Accuracy   | \( \frac{t_p + t_n}{t_p + t_n + f_p + f_n} \) | 0.7618220 |
| Recall     | \( \frac{t_p}{t_p + f_n} \) | 0.8666137 |
| Precision  | \( \frac{t_p}{t_p + f_p} \) | 0.8352182 |
● Perform feature selection
● Estimator tuning
● Prepare to add more sample
● Look for problem with data?
●
Acknowledgement

- We would like to express our gratitude for Dr. James Mullaney, U. of Sheffield and Dr. Krzysztof Ulaczyk, U. of Warwick for providing SkyMaker, a simulated data generator, and for the idea on feature extractions.

- I, the speaker, would like to thank NARIT and Ministry of Science and Technology for providing an opportunity to participate in the project.
N. M. Ball and R. J. Brunner. *Int. J. of Mod. Phys. D*, 19:1049–1106, 2010.
M. P. Hobson P. Graff, F. Feroz and A. Lasenby. *Not. R. Astron. Soc.*, 441:1741–1759.
A. Gramfort V. Michel B. Thirion O. Grisel M. Blondel P. Prettenhofer R. Weiss V. Dubourg J. Vanderplas
A. Passos D. Cournapeau M. Brucher M. Perrot F. Pedregosa, G. Varoquaux and E. Duchesnay. *J. Mach.
Learn. Research*, 12:2825–2830, 2011.
G. Orr and K. Muller Y. LeCun, L. Bottou. *Efficient BackProp*. Springer, 1998.
E. Bertin and S. Arnout. *Astron. Astrophys. Sup.*, 17:393–404, 1996.
P. B. Stetson. *Publ. Astron. Soc. Pac.*, 99:191–222, 1987.
J. S. Bloom and J. W. Richards. *Data Mining and Machine Learning in Time-Domain Discovery and
Classification*, pages 89–112. 2012.
D. Poznanski J. S. Bloom J. Rice S. Negahban H. Brink, J. W. Richards and M. Wainwright. *Not. R. Astron.
Soc.*, 435:1047–1060, 2013.
J. A. Fischer R. J. Foley R. R. Gupta R. Kessler A. G. Kim R. C. Nichol P. E. Nugent A. Papadopoulos M.
Sako M. Smith M. Sullivan R. C. Thomas W. Wester R. C. Wolf F. B. Abdalla M. Banerji A. Benoit-Lévy
E. Bertin D. Brooks A. Carnero F. J. Caster L. N. da Costa R. Covarrubias D. L. DePoy S. Desai H. T.
Diehl P. Doel T. F. Eifler A. Fausti Neto D. A. Finley B. Flaugher P. Fosalba J. Frieman D. Gerdes D.
Gruen R. A. Gruendl D. James K. Kuehn N. Kuropatkin O. Lahav T. S. Li M. A. G. Maia M. Makler M.
March J. L. Marshall P. Martini K. W. Merritt R. Miquel B. Nord R. Ogo A. A. Plazas A. K. Romer A.
Roodman E. Sanchez V. Scarpine M. Schubnell I. Sevilla-Noarbe R. C. Smith M. Soares-Santos F. Sobreira
E. Suchyta M. E. C. Swanson G. Tarle J. Thaler A. R. Walker D. A. Goldstein, C. B. D’rea. Automated
transient identification in the dark energy survey. *Astrophys.J.*, 150:82, 2015.