Steven Boada, Ph.D

Contact Information

Cranford, New Jersey, USA
(615) 200-0119
stevenboada@gmail.com
`http://boada.github.io`

Profile

Collaborative, scientific thinker passionate about discovering and communicating nuanced insight from complicated data. Strong programming and analytical background working with large, heterogeneous, and often noisy datasets.

Education

Texas A&M University, College Station, Texas USA
Ph.D, Physics (Astronomy focus), August, 2016
- Dissertation Title: "Measuring the Scatter in the Cluster Optical Richness–Mass Relation with Machine Learning"

The University of Tennessee, Knoxville, Tennessee USA
M.S., Physics (Astronomy focus), August, 2009
- Thesis Title: "An Automated Approach to the Study and Classification of Colliding and Interacting Galaxies"

B.S., Physics, May, 2007

Technical Skills

Machine Learning: Regression (Linear, Random Forests), classification (RF, SVM), feature engineering, optimization, deep learning
Statistical Methods: Hypothesis testing and confidence intervals, error analysis, image analysis, Monte Carlo methods (e.g., emcee)
Software and Computing: Python (e.g. Scikit-learn, Numpy, Scipy, Pandas, Matplotlib, fast.ai), SQL, ANSI C, Linux, Microsoft Excel, GPGPU, and HPC (100k+ core) applications

Data Projects

Using Imaging to Predict Galaxy Spectroscopic Properties
- Leveraged Convolution Neural Networks, trained on GPUs, to analyze $\sim 150,000$ images from the Sloan Digital Sky Survey. See `https://github.com/boada/galaxy-cnns`.
- Predicted spectroscopic properties with $\sim 5\%$ error from psuedo-three color imaging.

Predicting Tournament Performance in Warmachine
- Wrangled tournament results of a popular tabletop game using Python (e.g., Pandas).
- Created an Elo based model to forecast the results of an upcoming tournament.
- Explored tournament entries for insights into and to identify potential problematic future matches.

Professional Experience

Rutgers University, New Brunswick, New Jersey USA
Postdoctoral Research Associate **September, 2016 - Present**
- Designed and built a pipeline to analyzed TBs of astronomical imaging; producing calibrated, standardized data catalogs.
- Led weekly collaboration meetings with senior scientists and supervised student research.
- Contributed to open source projects including: PHOTOMETRYPIPELINE, ASTLIB, and EASYGALAXY.

Texas A&M University, College Station, Texas USA
Ph.D Candidate **August, 2010 - 2016**
- Conducted original research of a forthcoming astronomical survey and showed that (simulated) results could be improved by implementing machine learning techniques (e.g., random forest regression) when compared to traditional analysis methods. Implemented these ML methods and produced improved results in a pilot survey of the real sky and under real-world conditions.
- Collaborated with group members both in person, and through collaborative tools (e.g., GitHub, SVN).
- Presented scientific results in high-impact, peer reviewed journals and at international conferences.

The University of Tennessee, Knoxville, Tennessee USA
Master's Candidate **August, 2007 - 2009**
- Conducted original research at the National Center for Computational Science, part of Oak Ridge National Laboratory, using HPC ($\sim$ 100k cores) scientific simulations.
- Optimized simulation initial conditions using a genetic algorithm based search.
- Implemented a computer vision algorithm to automatically analyze hundreds of GBs of simulation results.