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.

Technical Skills

Machine Learning: Regression (linear, logistic), Random Forests, SVM, Clustering, Feature Engineering, Optimization, Deep Learning
Statistical Methods: Hypothesis testing and confidence intervals, error analysis, image analysis, Monte Carlo methods (e.g., emcee), maximum likelihood
Software and Computing: Python (e.g. Scikit-learn, Numpy, Scipy, Pandas, Matplotlib, fast.ai), mySQL, ANSI C, Linux Command Line Environments, Microsoft Excel, GPGPU, and HPC (100k+ core) applications

Data Projects

Using Imaging to Infer Galaxy Properties

- Predicted galaxy chemical composition with $\sim 5\%$ error from pseudo-three color imaging, a result better than other current, similar efforts in the literature.
- Leveraged Convolution Neural Networks, trained on GPUs, to analyze $\sim 150,000$ images from the Sloan Digital Sky Survey.
- Project start to publication, 4 months. See ArXiv:1810.12913

Predicting Tournament Performance in Warmachine

- Created an Elo based model to forecast the results of an upcoming tournament and to identify potential future upsets.
- Wrangled ~ 1800 tournament game results of a popular tabletop game using Python (e.g., Pandas).

Professional Experience

Dept. of Physics and Astronomy, Rutgers University, New Brunswick, New Jersey USA
Postdoctoral Research Associate **September, 2016 - Present**

- Designed and built massive, parallelized, Python pipelines to process and analyze TBs of astronomical imaging; producing calibrated, standardized data catalogs and rigorous results.
- Coordinated a team of 4, including both senior scientists and graduate students, to perform quality control tasks; deliver science products; and produce peer-reviewed publications. See: ArXiv:1808.06378
- Contributed to open source Python projects including: PHOTOMETRYPIPELINE, ASTLIB, and EASYGALAXY.

Texas A&M University, College Station, Texas USA
Ph.D Candidate **August, 2010 - 2016**

- Proved simulated results for an upcoming astronomical survey could be improved, by a factor of ~ 3, over in-house statistical methods by using Random Forest regression. 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**

- Implemented a C-based pipeline to process hundreds of GBs of simulation results. Including a computer vision algorithm to automatically analyze and compare results to expected targets.
- Optimized simulation parameters using a genetic algorithm based search utilizing HPC (100k+ core) systems at the National Center for Computational Science, part of Oak Ridge National Laboratory

Education

Texas A&M University, College Station, Texas
- Ph.D, Physics (astronomy focus), August, 2016

The University of Tennessee, Knoxville, Tennessee
- M.S., Physics (astronomy focus), August, 2009
- B.S., Physics, May, 2007