Steven Boada, Ph.D

Contact Information

(615) 200–0119
stevenboada@gmail.com

github.com/boada
linkedin.com/in/theboada

Skills

Machine Learning: Linear Models, Decision Trees, SVM, Clustering, Deep Learning, Feature Engineering
Statistical Methods: Hypothesis testing, error analysis, Monte Carlo methods, maximum likelihood
Software and Computing: Python (e.g. Scikit-learn, Numpy, Scipy, Pandas, Matplotlib, PyTorch), mySQL, ANSI C, Linux Command Line Environments, GPGPU, and HPC applications
Leadership: Experience organizing and leading workshops and collaboration meetings; Eagle Scout.

Professional Experience

Insight Data Science, New York, New York USA
Fellow   **January, 2020 – Present**
- Helped optimize the way NYC health inspectors perform restaurant inspections in order to reduce the time critical health violations remain unaddressed.
- Trained a random forest in Python to prioritize NYC restaurant inspections based on environmental variables and their past inspection histories and provided the results to NYC through an easy to use API.
- Resulted in NYC inspectors identifying ∼2.5% more violations in the first half of an inspection window, leading to critical violations being discovered up to 7 days earlier than by the current approach implemented by NYC.

Dept. of Physics and Astronomy, Rutgers University, New Brunswick, New Jersey USA
Postdoctoral Research Associate   **September, 2016 – 2020**
- Designed and built parallelized pipelines to process and analyze TBs of astronomical imaging; producing calibrated, standardized data catalogs and rigorous results leading to 2 peer reviewed publications and several hundred hours of telescope time.
- Project managed and 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.
- Contributed to open source, astronomy-focused, Python projects through bug fixes and feature additions: see photometrypipeline, astLib, and easyGalaxy on GitHub as examples.

Dept. of Physics and Astronomy, Texas A&M University, College Station, Texas USA
Ph.D Candidate   **August, 2010 – 2016**
- Demonstrated that measurements from a planned large observation campaign could be improved by up to a factor of 3 over traditional statistical methods through the use of machine learning.
- Implemented these machine learning methods and produced reliable 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, astrophysical journals and at international conferences.

Data Projects

Using Imaging to Infer Galaxy Properties
- Predicted galaxy chemical composition with ∼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 ∼150,000 images from the Sloan Digital Sky Survey.
- Project start to publication: 4 months (typically ∼1.5 years). See: github.com/boada/galaxy-cnns.

Predicting Tournament Performance in Warmachine
- Created an Elo based model to forecast the results of upcoming tournaments and identify potential upsets.
- Integrated predictions into a local community ranking system and forecasted ∼1800 tournament game results of the popular tabletop game using Python (e.g., Pandas).

Education

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

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