RICHARD KNOCHE

Greater New York City
(703)-801-4456 • raknoche@gmail.com
www.dealingdata.net • www.linkedin.com/in/richardknoche

EDUCATION

University of Maryland, College Park Aug 2011 - Dec 2016
Doctor of Philosophy in Physics
Dissertation: Signal Corrections and Calibrations in the LUX Dark Matter Detector

James Madison University Aug 2007 - May 2011
Bachelor of Science in Physics, magna cum laude

EXPERIENCE

Insight Data Science *Data Science Fellow* Jan 2017 - Present

- Consulted for AptDeco.com to quantify the subjective quality of half a million user-uploaded images.
- Implemented 60 hand-crafted image features in a Python-based AdaBoost random forest classifier to automatically identify high-quality and low-quality images with 73% and 68% accuracy, respectively.
- The classifier will improve editor efficiency by 40%, and save an estimated $45,000 annually.

University of Maryland *Graduate Research Assistant* May 2011 - Dec 2016

- Worked with an international collaboration to produce upper limits on dark matter interaction rates.
- Developed analysis techniques to remove over 99.9% of background signals from our data. These techniques have been adopted by experiments in the USA, Asia, and Europe.
- Implemented maximum likelihood estimation in Python and MATLAB to produce signal corrections and improve our detector's sensitivity by an order of magnitude.
- Automated the extraction of hundreds of data features from 650 TB of data. Maintained these features in a MySQL database that was queried by 200 collaborators.

NASA GSFC *Research Assistant* June 2010 - Sep 2010

- Performed exploratory and statistical analysis of data from the Swift Burst Alert Telescope (BAT) to search for hard X-ray emissions around the on-set time of supernovae.

PERSONAL PROJECTS

DealingData.net *Data Science Blog* Jul 2016 - Present

- Predicted the top fantasy football performers of each season with a support vector machine ($R^2 = 0.72$). Used K-means clustering to automatically extract tiers of players to target during a fantasy draft.
- Implemented a Naive Bayes sentiment analysis on Pokemon Go tweets from each state. Used the results to measure the dominance of each team across the country. Visualized the results in an interactive map.
- Presented in-depth tutorials on various topics, including time series analysis, statistical analysis of correlated variables, Bayesian statistics, specific APIs, web scraping, and web hosting.

TECHNICAL STRENGTHS

Computer Languages	Python (Sklearn, SciPy, Pandas, Gensim, Flask), MATLAB, HTML
Quantitative Skills	Physics, Calculus, Profile Likelihood, Error Analysis, Monte Carlo
Machine Learning	MLE, Linear Regression, Naive Bayes, Random Forest, Logistic Regression, TF-IDF, Association Rules, K-means
Tools	SVN, Git, MySQL