

Scott Clark

scott@scottclark.io

@DrScottClark

June 30, 2016

www.scottclark.io

Not looking for new job opportunities

Education

- **Cornell University** Ithaca, NY
Ph.D. Applied Mathematics, M.S. Computer Science 2008 - 2012
 - Department of Energy Computational Science Graduate Fellow (Full Scholarship, 4 years)
 - Machine Learning, Data Science, Data Mining in bioinformatics and global optimization
- **Oregon State University** Corvallis, OR
B.Sc. Mathematics, B.Sc. Computational Physics, B.Sc. Physics 2004 - 2008
 - Graduated Magna Cum Laude with minors in Actuarial Sciences and Mathematical Sciences

Research and Work Experience

- **SigOpt Inc - <https://sigopt.com>** San Francisco, CA
Co-founder and CEO November 2014 - current
 - Helping lead a world class team to eliminate expensive trial and error from every experts workflow. Using cutting edge optimization behind a simple API to help tune machine learning models and build better products in a variety of fields.
- **Yelp Inc** San Francisco, CA
Data Mining Engineer and Lead on Ad Targeting Team July 2012 - December 2014
 - **Optimization:** Co-developed and led team for MOE: the Metric Optimization Engine (github.com/Yelp/MOE, an **open source** optimization framework), found significant gains in different metrics across the organization using Bayesian Global Optimization algorithms.
 - **Targeting:** Implemented multi-armed bandit strategies for ad selection, sole targeting engineer on mobile app ads rollout, developed new location-based targeting algorithms, advised and helped develop other machine learning and math based targeting projects.
 - **Recruiting:** Created, implemented, and directed yelp.com/dataset_challenge, gave tech talks across the country, led events, gave hundreds of technical interviews, closed candidates.
- **Bloomberg LP** New York, NY
Financial Software Development Intern Summer 2011
 - Implemented statistical models to perform forward and backward portfolio analysis
- **DOE Joint Genome Institute (Lawrence Berkeley National Lab)** Walnut Creek, CA
Researcher in Analysis Group under Dr. Zhong Wang Summer 2010

- Used machine learning to mine TBs of genome data efficiently using novel likelihood function

- **Los Alamos National Laboratory** Los Alamos, NM
Researcher in Metagenomics Group under Dr. Nick Hengartner *Summer 2009*

- Used statistical models to discover sequence alignments using parallel algorithms on GPUs

- **Oregon State University** Corvallis, OR
Research Assistant under Prof. Malgorzata Peszynska and Prof. Rubin Landau *2005-2008*

- Finite element analysis with uncertainty and web-based teaching in **Java**

- **Max Planck Institute for the Physics of Complex Systems** Dresden, Germany
NSF REU Research Assistant under Prof. Steven Tomsovic *Summer 2007*

- Research on extreme value statistics in **MATLAB** and **FORTRAN**

- **University of California: Davis** Davis, CA
NSF REU Research Assistant under Prof. Daniel Cox *Summer 2006*

- Computational biophysics research as applied to protein folding in **Java**

Writing and Awards

- **2016 Forbes 30 Under 30:** Enterprise Tech. <http://onforb.es/10ILpBZ>
- **Department of Energy Computational Science Graduate Fellow:** Four year full fellowship. ~20 awarded nationally per year. Won the Communicating Science award (bit.ly/VbcTZK).
- **SigOpt Blog:** Posts talking about using SigOpt to optimize everything (blog.sigopt.com).
- **Yelp Blog:** Wrote several posts announcing the open sourcing of MOE, the Yelp Dataset Challenge and more. bit.ly/1x73xdr, bit.ly/1oCCZvv, bit.ly/1s0sEBS, bit.ly/1p1X7Hk
- **Press:** WSJ: on.wsj.com/Va0vqQ, Cornell: bit.ly/1oB2dzm, DIEXIS: bit.ly/1oofb14

Skills

- **Numerical Analysis and Computer Science:** Machine Learning, Data Mining, Optimization, Computational Science, Artificial Intelligence, Linear Algebra, Monte Carlo Methods, ODEs, PDEs, Iterative Methods, Parallel Programming, Distributed Systems, Data Structures
- **Tech Stack:** Python, numerical libraries, linux, git, vim
- **Public Speaking:** I've given several hundred technical talks to audiences at machine learning conferences, Fortune 500 boards, and beyond.
- Exploring and implementing ideas. Give me an API/dataset and a problem and I will figure it out.

Selected Open Source Projects

- **SigOpt Examples** (github.com/sigopt/sigopt-examples) Python
Examples of using SigOpt to tune ML algorithms. 2014 - current
 - Examples of using SigOpt to tune everything from sklearn to beating Vegas and beyond.
- **MOE: Metric Optimization Engine** (github.com/Yelp/MOE) Python, C++, CUDA
A global, black box optimization engine for real world metric optimization 2010 - 2015
 - Implemented throughout Yelp, optimizing ad metrics. 2nd most popular open source project.
 - Talk: bit.ly/1p1YZA2, Slides: slidesha.re/1z0r0Jy, Blog: bit.ly/1x73xdr
 - Presented to executives, universities, conferences and companies around the country.
- **ALE: Assembly Likelihood Estimator** (github.com/sc932/ALE) C, Python
Probabilistic evaluation of genome assemblies 2010 - 2013
 - Uses statistical function to score and rank genome assemblies, published in Bioinformatics