

Scott Clark

scott@scottclark.io
@DrScottClark

July 3, 2016

scottclark.io

Not looking for new job opportunities

Education

Ph.D. Applied Mathematics, M.S. Computer Science 2008 – 2012
Cornell University Ithaca, NY

- Department of Energy Computational Science Graduate Fellow (Full Scholarship, 4 years)
- Machine Learning, Data Science, Data Mining in bioinformatics and global optimization

B.Sc. Mathematics, B.Sc. Computational Physics, B.Sc. Physics 2004 – 2008
Oregon State University Corvallis, OR

- Graduated Magna Cum Laude with minors in Actuarial Sciences and Mathematical Sciences

Research and Work Experience

Co-founder and CEO Nov 2014 – ...
SigOpt Inc. San Francisco, CA

- 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.

Data Mining Engineer and Lead on Ad Targeting Team Jul 2012 – Dec 2014
Yelp Inc. San Francisco, CA

- **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.

Financial Software Development Intern Summer 2011
Bloomberg L.P. New York, NY

- Implemented statistical models to perform forward and backward portfolio analysis

Researcher in Analysis Group under Dr. Zhong Wang Summer 2010
DOE Joint Genome Institute (Lawrence Berkeley National Lab) Walnut Creek, CA

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

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

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

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

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

NSF REU Research Assistant under Prof. Steven Tomsovic
Max Plank Institute for the Physics of Complex Systems

Summer 2007
Dresden, Germany

- Research on extreme value statistics in MATLAB and FORTRAN

NSF REU Research Assistant under Prof. Daniel Cox
University of California, Davis

Summer 2006
Davis, CA

- Computational biophysics research as applied to protein folding in Java

Writing and Awards

- **2016 Forbes 30 Under 30:** Enterprise Tech. 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

Examples of using SigOpt to tune ML algorithms
SigOpt Examples (github.com/sigopt/sigopt-examples)

2014 – ...
Python

- Examples of using SigOpt to tune everything from sklearn to beating Vegas and beyond.

A global, black box optimization engine for real world metric optimization

2010 – 2015

MOE: Metric Optimization Engine (github.com/Yelp/MOE)

Python, C++, CUDA

- 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.

Probabilistic evaluation of genome assemblies

2010 – 2013

ALE: Assembly Likelihood Estimator (github.com/sc932/ALE)

C, Python

- Uses statistical function to score and rank genome assemblies, published in Bioinformatics