Steve Herrin

Contact Information

steve.herrin@gmail.com
650-814-8865
San Jose, CA

www.github.com/steveherrin
www.linkedin.com/in/herrinsteve
medium.com/@steveherrin

Experience

23andMe, Mountain View, CA — January 2014 – present

Engineering Team Lead: Data Services and Applications Team

- Grew an engineering team from 2 engineers (including myself) working on one project to 6 engineers working on 2–3 projects
- Architected machine learning infrastructure and services, allowing training, evaluation, and easy promotion to production of models for dozens of users across multiple departments
- Created Genotyping Services, a webapp allowing external researchers to easily run genetic studies, increasing sales by over 2%
- Led team responsible for a web-based portal that allows internal and external researchers to dynamically query anonymized data stored in HBase for >3 million customers
- Migrated a 20 kLOC web application to the AWS cloud, upgrading the back-end from Python 2 to 3 and standardizing the front-end using Elm
- Architected and implemented a Python library that provides a unified API for accessing customer data across MySQL, HBase, and other data stores
- Built 3 generations of data pipelines with Celery, Luigi, and AWS Step Functions to run Python, C++, and R algorithms that impute, transform, and analyze TB of genetic data

Insight Data Science, Mountain View, CA — August 2013 – December 2013

Postdoctoral Fellow

- Developed Parksafely, a web app applying a heuristic algorithm to make parking recommendations, reducing bike theft risk by 40% while requiring only 150 ft more walking on average. Used Flask, PostgreSQL/PostGIS, and Javascript.

SLAC National Accelerator Lab, Menlo Park, CA — May 2008 – August 2013

Research Associate

- Applied machine learning and computer vision algorithms to improve detector energy resolution by 25%.
- Created a PHP logbook webapp with a MySQL backend for tracking work on the EXO-200 experiment.
- Developed batch data pipelines using Python, C++, and shell scripts to process TB of calibration data.
- Mentored 1–2 junior graduate students (at any given time) on lab, coding, and statistical technique.

Rice University, Houston, TX — May 2005 – May 2007
and **University of Washington**, Seattle, WA — June 2006 – August 2006

Undergraduate Research Assistant

- Implemented (in C++) and evaluated random forest and boosted decision tree algorithms that contributed to the discovery of single top quark production by Fermilab's D0 experiment.

Skills

Languages: Python, Elm, C++, SQL, Shell Scripting, MATLAB/Octave, Rust (some experience), JavaScript (some experience), R (some experience)

Tools: AWS, NumPy, SciPy, Celery, Luigi, Django, MySQL, PostgreSQL, Git, LaTeX, HBase, Spark

Other: Machine learning, classification, regression, statistics, hypothesis testing, Monte Carlo simulations, time-series analysis

Open Source

SpookyOTP: A lightweight Python implementation of TOTP/HOTP authentication

Education

Stanford University, Stanford, CA — June 2013
- Ph.D. (Physics)

Rice University, Houston, TX — May 2007
- B.S. (Physics)