Steve Herrin

CONTACT INFORMATION

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

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

EXPERIENCE

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

Senior Software Engineer

- Led engineering effort (supervising 1–2 engineers) to create Genotyping Services, leading to increased sales by allowing external researchers to easily run studies on a Django and Backbone JS webapp.
- Created, as part of a team, a web-based platform for internal and external researchers to access dynamically computed aggregate statistics on anonymized data stored in HBase for >1 million customers.
- Architected and implemented a Python library that provides a unified API for accessing customer data across MySQL, HBase, and other data stores.
- Built data pipelines with Luigi and Celery to run Python, C++, and R algorithms to impute, transform, and analyze TB of genetic data in both a scientific computing cluster and the AWS cloud.
- Developed, maintained, and monitored ETL processes for billions of genotypic and phenotypic records into MySQL and HBase tables.

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

- Created a PHP logbook webapp with a MySQL back-end for tracking work on the EXO-200 experiment.
- Maintained and improved (refactoring, adding tests) a real time LabVIEW-based control and monitoring system with over 600 channels.
- Developed batch data pipelines using Python, C++, and shell scripts to routinely measure detector characteristics, improving energy resolution by 25%.
- 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.
- Investigated and benchmarked many different machine learning classification algorithms for their power to discriminate signals of supersymmetry from backgrounds.

SKILLS

- **Languages:** Python, C++, Bash Shell Scripting, PHP, MATLAB/Octave, Mathematica (some experience), Java (some experience), JavaScript (some experience), R (some experience)
- **Tools:** AWS (Certified Solutions Architect – Associate Level), NumPy, SciPy, Celery, Luigi, Django, MySQL, PostgreSQL, Git, SVN, P4, LaTeX, HBase, Spark, Backbone JS
- **Other:** Machine learning, classification, regression, statistics, hypothesis testing, Monte Carlo simulations, time-series analysis

EDUCATION

Stanford University, Stanford, CA — June 2013

- Ph.D. (Physics)

Rice University, Houston, TX — May 2007

- B.S. (Physics)