Aaron **Fraenkel**

Machine Learning Scientist · Machine Learning Engineer · Data Scientist

1-510-207-2298 | aaron.fraenkel@gmail.com | afraenkel.github.io | afraenkel | aaron-fraenkel-53723095

Experience

amplisal.io *San Diego, CA*

Full-Stack Data Scientist / Software Engineer (Co-Founder) *2020 - PRESENT*

- Developed a remote-work recommendation system for relocation suggestions.
- Developed models for labor-pool profiling using in-house discrete-choice surveys and adminstrative data sources.
- Architected data management systems for 100s of heterogeneous data sources.
- Designed and deployed survey collection and model execution using s3/aws-lambda/jQuery/d3.js.

Halicioglu Data Science Institute, UCSD *San Diego, CA*

Assistant Teaching Professor (Data Science; Computer Science) *2018 - PRESENT*

- Planned and implemented the start of a new undergraduate program in Data Science.
- As chair of the undergraduate program, guided the major through 10x growth in users and offerings.
- Developed sui generis courses on 'Fairness and Algorithmic Decision Making' and 'Software Development for Data Science.'
- Conducted research on geometric deep-learning and fairness in algorithmic systems.
- Designed software systems to scale scientific research.
- Developed infrastructure and services built on Docker and Kubernetes for the educational Data Science computing platform.

Amazon.com *San Diego, CA*

Senior Machine Learning Scientist *2016 - 2018*

- Machine learning technical lead for bot detection at/post sign-in.
- Model design, development, and deployment in policy abuse, fraud, and account integrity.
- Implemented a model training pipeline in Scala using Spark ml-pipelines.
- Lead initiatives in the areas of graph-learning and adversarial ML, adapting academic research to large-scale business solutions.
- Developed instrumentation for real-time evaluation of production models (Performance, Fairness, Anomalies).

ID Analytics / LifeLock *San Diego, CA*

Senior Data Scientist *2014 - 2016*

- Rapid prototyping and initial productionization of new ML-driven consumer products in Python and Golang on AWS.
- Designed and implemented full-stack internal tooling (model optimization; large-scale data munging; model explanability).
- Developed, trained, and maintained production fraud models.
- Implemented fairness auditing instrumentation into the modeling pipeline.

Education

UC Berkeley *Berkeley, CA*

Ph.D. Mathematics

- Representation Theory, Differential Geometry, Algebraic Topology

UC Berkeley *Berkeley, CA*

B.A. Mathematics

Skills

Programming Languages	Python, bash, SQL, + others used occasionally (Golang, Javascript, Scala, Java)
ML Fundamentals	prediction, clustering, kernel methods, regularization, NLP, anomaly detection, imputation models
Advanced ML Topics	geometric deep learning (graph NN, symmetry-preserving NN), graph learning, algorithmic fairness, adversarial ML
Modeling Libraries	PyData Stack, Sklearn, PyTorch, statsmodels
Model Serving	ML Pipelines (Sklearn; Spark), MLFlow
Distributed Computing	Spark, Hadoop (Hive, Pig), dask
Cloud Systems	Serverless Architectures; Distributed Workflows; Autoscaling and Containerization
Infrastructure and Tooling	Docker, Ansible, Git (Actions), Make
Web Frameworks	Sanic; MongoDB, SQLite, PostGreSQL, s3; jQuery, d3.js; CSS