

# Aaron McMillan Fraenkel

9500 Gilman Dr MS 0555, SDSC 214E, La Jolla, CA 92093

☎ 610-207-2298 | ✉ [afraenkel@ucsd.com](mailto:afraenkel@ucsd.com) | 📧 [afraenkel](#) | 📄 [aaron-fraenkel-53723095](#) | <http://afraenkel.github.io/>

## Experience

---

### Halicioglu Data Science Institute, UCSD

TENURE TRACK ASSISTANT TEACHING PROFESSOR

- Chair, Data Science Program Committee (major/minor)

San Diego, CA

Aug. 2018 - PRESENT

### Computer Science and Engineering, UCSD

TENURE TRACK ASSISTANT TEACHING PROFESSOR

San Diego, CA

Aug. 2018 - Aug. 2019

### Amazon.com

SENIOR MACHINE LEARNING SCIENTIST III

- Machine learning technical lead for bot detection.
- Model design and development for account takeover.
- Lead research initiatives and intern projects.

San Diego, CA

Nov. 2016 - Aug. 2018

### ID Analytics / LifeLock

SENIOR DATA SCIENTIST

- Design and initial implementation of new consumer products (full stack).
- Develop, train, and maintain fraud models.
- Research projects: ML algorithm development and new technologies.

San Diego, CA

Sep. 2014 - Nov. 2016

### Boston College

VISITING ASSISTANT PROFESSOR OF MATHEMATICS

Boston, MA

Aug. 2012 - Aug. 2014

### Penn State University

CHOWLA RESEARCH ASSISTANT PROFESSOR

State College, PA

Aug. 2011 - Aug. 2012

### University of Mass, Boston

LECTURER (GRADUATE SCHOOL)

Boston, MA

Jan. 2011 - Jun. 2011

### University of California, Berkeley

GRADUATE STUDENT INSTRUCTOR (GRADUATE SCHOOL)

Berkeley, CA

Aug. 2004 - Aug. 2011

## Education

---

### University of California, Berkeley

PH.D. IN MATHEMATICS

- Dissertation: *On Embedding Singular Poisson Spaces*

Berkeley, CA

2011

### University of California, Berkeley

B.S. IN MATHEMATICS

Berkeley, CA

2004

### MIT

VISITING SCHOLAR

Cambridge, MA

2008

### UPMC, Universite Paris VI

VISITING SCHOLAR

Paris, Fr

2009

## Research

---

My research interests span applied machine learning across a variety of fields, particularly understanding the robustness of techniques applied to heterogeneous data-types without a natural choice of metric. I'm especially interested in exploring these concepts in profiling and understanding behaviors on the internet, in the areas of security and abusive conduct. I'm also active in Data Science pedagogy and course development. See the curriculum page of my website for details.

# Teaching Activities

---

## UCSD

San Diego, CA

AS AN ASSISTANT TEACHING PROFESSOR (\* MEANS COURSE DEVELOPED FROM SCRATCH)

2018-PRESENT

- Capstone Methodology: Data Science Software Development (2020, 2021)\*
- Capstone Seminar: Conflict on Online Platforms (2020, 2021)\*
- Capstone Seminar: Anomalies in Heterogeneous Graphs (2020, 2021)\*
- Capstone Seminar: Predictive Policing and Fair Policing (WI2020, SP2020)\*
- Practice and Applications of Data Science (WI2019, SP2019, FA2910)\*
- Teaching Methods in Data Science (WI2019, FA2019)
- Principles of Data Science (FA2018)
- Workshop in Data Science (FA2018)\*

## Boston College

Chestnut Hill, MA

AS A VISITING ASSISTANT PROFESSOR

2012-2014

- Probability theory: an Introduction to Random Variables using R (S2013, F2013)
- Complex Analysis (F2012, S2013)
- Calculus I (F2012)
- Calculus II (F2013)
- Introduction to Abstract Mathematics (S2014)

## Penn State University

State College, PA

AS A RESEARCH ASSISTANT PROFESSOR

2011-2012

- Multi-Variable Calculus (F2011, course coordinator)

## University of Mass, Boston

Boston, MA

AS A LECTURER (GRADUATE STUDENT)

2011

- Calculus II, S2011

## UC Berkeley

Berkeley, CA

AS A GRADUATE STUDENT INSTRUCTOR

2004-2010

- courses taught: Real Analysis (Su2010), Multi-Variable Calculus (Su2005)
- recitations led: Real Analysis, Linear Algebra, Calculus I,II, III

# Mentoring & Advising

---

## Undergraduate Research (Funded)

HDSI, UCSD

ADVISOR

2018-2020

- Non-parametric Graph Sampling
- **babypandas**: An Opinionated Pandas API for New Coders.
- Zillow ZTrax: Organizing and Querying Large Housing Data.
- Zillow ZTrax: Understanding Housing Transactions.
- Scaling the Estimation of the Rate of Recession of Glaciers.
- Data Compression via MMD.
- Reassessment of Credit Risk Modeling with Macroeconomic Factors.
- Twitter Bots and Propoganda amid Hong Kong Protests.
- Malware Detection via Graph Learning. (20 projects)
- Issues of Fairness in Algorithmic Policing. (4 projects)
- Quantifying conflict on Wikipedia (6 projects)

## Advising: Amazon Summer Intern Program

Amazon.com

ADVISOR

2017

- Developed summer project and mentored intern in machine learning.
- Project resulted in a production product / job offer.

## Advising: Undergraduate Honors Projects

Penn State University

ADVISOR

2011-2012

- Undergraduate honors project ("Applications of gaussian curvature," Fall 2011).
- Undergraduate independent study ("Differential Topology," Spring 2012).

# Honors & Awards

---

- 2012 **Recipient**, Recognized faculty for influencing students in their transition to college
- 2010 **Recipient**, RTG NSF Graduate Student Fellowship
- 2009 **Recipient**, Outstanding Graduate Student Instructor Award
- 2004 **Recipient**, Percy Lionel Davis Award for Excellence in Scholarship in Mathematics

Penn State

UC Berkeley

UC Berkeley

UC Berkeley

## Publications & Preprints

---

- 2018 **Bot Detection and Machine Learning**, Amazon Machine Learning Conference (internal)  
2013 **Extensions of Poisson Structures on Singular Hypersurfaces**, arXiv:1310.6083  
2011 **On Embedding Singular Poisson Spaces**, arXiv:1108.2207

## Presentation

---

- June 2020 **Panel: California Data Science Strategy**, Data Science Pedagogy Workshop *UC Berkeley*  
June 2020 **Panel: Human Contexts and Ethics**, Data Science Pedagogy Workshop *UC Berkeley*  
June 2019 **Data8 meets Messy Data**, Data Science Pedagogy Workshop *UC Berkeley*  
Dec. 2018 **Adversarial Machine Learning in Production**, AI Seminar, CSE UCSD *UCSD*  
Apr. 2018 **Assessing, Detecting, and Mitigating Bots in the Absence of Ground-Truth.**, Special Lecture Series, CSE UCSD *UCSD*  
Dec. 2017 **Machine Learning in Bot Detection**, Amazon Machine Learning Lecture Series *San Francisco, CA*  
Apr. 2017 **Bots and Account Takeover**, Amazon Science Fair *Seattle, WA*  
Jul. 2013 **20 years of the Poisson embedding problem**, Symposium in honor of Alan Weinstein, Institute Henri Poincare *Paris, FR*  
Jul. 2013 **The Poisson embedding problem**, Geomtry Seminar, Instituto de Ciencias Matematicas *Madrid, ES*  
Oct. 2011 **Embedding singular Poisson varieties**, “Gone Fishing” Poisson Geometry Conference, Washington University *STL, MO*  
May. 2011 **Embedding singular Poisson varieties**, Northern California Symplectic Geometry Seminar *Stanford, CA*  
April. 2011 **Embedding singular Poisson varieties**, Department Colloquium, UMass Boston *Boston, MA*  
Oct. 2010 **A Levi-type decomposition of singular Poisson varieties**, Geometry and Physics Seminar, Penn State University *State College, PA*

## Committees

---

- 2019- **Chair**, Undergraduate Program *HDSI, UCSD*  
2019- **Member**, Infrastructure Committee *HDSI, UCSD*  
2019- **Member**, Committee on Diversity, Equity, and Inclusion (DEI) *HDSI, UCSD*  
2019- **Member, Chair**, HDSI Search Committee(s) *HDSI, UCSD*  
2018-2019 **Advisor**, Industry Advisor Program *HDSI, UCSD*  
2018 **Member**, M.S. Admission Committe *CSE, UCSD*  
2017 **Reviewer, security and bot detection**, Amazon Research Awards *Amazon.com*  
2017 **Referee**, Amazon Machine Learning Conference *Amazon.com*

## Skills

---

- Programming Languages** Python, golang, bash, SQL, Scala, R, javascript,  $\LaTeX$   
**Systems** unix, osx, hdfs, aws  
**Tools, Libraries, Platforms**
- ml libraries: spark-ml, scikit-learn, tensorflow, keras
  - productivity: emacs, git, jupyter, d3.js, react.js
  - big data: hadoop ecosystem, docker, kubernetes