Katelyn L. Arnemann

klarnemann@gmail.com ♦ (551) 227-0834 ♦ linkedin.com/in/klarnemann ♦ github.com/klarnemann

Work Experience

Data Scientist, *Unlearn.AI* (San Francisco, CA) 03/2021 to present

- **Led team** of data scientists and engineers focused on data processing for new **disease progression models**
- Collaborated to develop software in **python** to generalize and streamline data processing and EDA

Data Scientist, *Datacubed Health* (Brooklyn, NY) 12/2019 to 03/2021

- Conducted **statistical validation** of Linkt mobile app features, including surveys, cognitive tasks, and geofencing
- Performed **data wrangling** and **QA** for data transfers to clients conducting **clinical trials**
- Conducted ML analyses (e.g., **survival analysis**, **classification**) to predict behavior of study participants

Postdoctoral Fellow, *Rutgers University* (Newark, NJ) 07/2018 to 12/2019

- Developed a novel measure of brain connectivity by applying **factor analysis** (dimensionality reduction) to neuroimaging scans; results improved predictions of evoked brain activation patterns and behavior
- Contributed to open-source software *Brain Activity Flow Toolbox* for predictive models of brain activation patterns

Graduate Student Researcher, *University of California, Berkeley* (Berkeley, CA) 08/2012 to 05/2018

- Modeled spread of Alzheimer's pathology by developing a **directed graph model** from neuroimaging data
- Integrated multimodal neuroimaging data to predict the topology of Alzheimer's pathology using a **power-law model**, achieving state-of-the-art prediction of brain areas vulnerable to amyloid-β

Data Analyst, *U.S. Department of Veteran's Affairs* (Martinez, CA) 10/2010 to 08/2012

- Predicted individual differences in response to cognitive training by applying tools from **graph theory** (community detection) and a perceptron **neural network** to neuroimaging data for patients with traumatic brain injuries
- Contributed to open-source software *brainx* in **Python** for **community detection** for brain network analyses

Leadership & Projects

Data Science Fellow, *Insight* (New York, NY) 2019

- Built a web-app **recommender system** for little-known travel destinations incorporating **NLP** (topic analysis) and categorical features (MCA, k-means) of data scraped from wikipedia.org and stateparks.com

Founder, Neuroscience Data Mining Group, *University of California, Berkeley* (Berkeley, CA) 2014 to 2017

- Founded and organized peer-based group to disseminate **computational** and **machine learning** techniques

Participant, CDIPS Data Science Workshop, *University of California, Berkeley* (Berkeley, CA) 2017

- Collaborated to build a Wikipedia page **recommender system** using **NLP** (topic analysis) and validated results using **graph theory** (link-distance)

Graduate Student Instructor, *University of California, Berkeley* (Berkeley, CA) 2015

- Led cooperative graduate course "Statistics for Neuroscience" on **statistics/ML** using **R** and **Matlab**

Education

Ph.D. in Neuroscience, *University of California, Berkeley* (Berkeley, CA)

B.A. in Cognitive Science and Philosophy, *Case Western Reserve University* (Cleveland, OH)

Skills

Languages: Python, R, Matlab

Tools: github, SQL, AWS S3, Google Analytics/Firebase, Agile

Packages: jupyter, pandas, sklearn, numpy, scipy, dagster, networkx, matplotlib, seaborn, genism, nltk, rpy2, beautiful soup, flask