

Caleb E. Strait

Data Scientist

(413) 884 - 4767
caleb.strait@gmail.com
Ohio, USA

calebstrait.com
github.com/calebstrait
linkedin.com/in/calebstrait

Skills

Communication

Data Visualization

Statistical Modeling

Machine Learning: K-Means, PCA, SVM, Linear & Logistic Regression, Artificial Neural Network, Decision Trees, Naive Bayes

Matlab: Psychtoolbox

Python: Flask, Pandas, Numpy, Jupyter Notebook, SciKit-Learn, Matplotlib

Web: HTML, CSS, JS, API

OS: Mac, Windows, Linux

SQL: SQLAlchemy, PostgreSQL

Java, C++/C#, SPSS

Education

University of Rochester

Ph.D., Brain & Cognitive Sciences
May 2016

M.A., Brain & Cognitive Sciences
January 2014

Oberlin College

B.A., Psychology

Concentration, Cognitive Science
Minor, Computer Science
May 2011

Experience

Data Science Fellow **Insight Data Science** 2017

- Built "Nextgame," a web application that gives increasingly personalized video game recommendations by prompting user feedback for each of a series of recommended games. Wrote a custom recommendation algorithm using collaborative filtering and reinforcement learning, collected data with web scraping and the igdb.com API, cleaned data with regular expressions, and stored data in a PostgreSQL database hosted with AWS. Built a front end in Flask that queries this database for each user using SQLAlchemy. (**Python**)
 - Hosted at nextgame.site
 - Repo at goo.gl/2dGQ3C
- Analytically validated the app's performance after saving usage data from 42 users. Used a binary cumulative density function test on logistic regression coefficients to show that there were significantly more users than we would expect by chance for whom recommendations were increasingly positively received given more feedback. (**Python**)

Senior Scientist **University of Rochester** 2011-2016

- Designed and programmed 30+ custom experiments to characterize neural signal processing during human decision-making. (**Matlab, C++**)
- Led a team of three laboratory technicians and four research assistants in recording neurophysiological data during the 4 experiments that produced the cleanest behavioral data.
- Estimated mid-decision neuronal spike frequencies from extracellular voltage measurement time series data using PCA with K-Means. (**Matlab, OmniPlex**)
- Wrote custom sliding-window logistic regression models of neuronal spike frequencies in reward-specialized brain regions as functions of decision parameters and choice behavior data. Demonstrated that neurons encoding an option's value tended to do so relative to its best alternative, suggesting that these regions encode decision-relevant information in a way that computationally subtracts an option's pros from its cons. (**Matlab**)
- Created a Signal Processing Toolkit for use with standard-format decision-making experiment datasets with neural data. Fits a set of sliding-window logistic regressions and outputs a series of figures characterizing how the recorded neurons tended to encode options' values relative to each other. (**Matlab**)
 - Repo at goo.gl/1TndRV
- Published four 1st-author peer-reviewed articles in some of the highest impact-factor journals in psychology and neuroscience.
 - Hosted at calebstrait.com
- Wrote and successfully defended a doctoral thesis around these works, "Neural Mechanisms of Reward-Based Choice," through a series of public departmental talks.
 - Hosted at goo.gl/VT4KNJ
- Organized, chaired, and spoke at a symposium, "The Neural Basis of Economic Choice," at neuroscience's biggest conference, the Society for Neuroscience's Annual Meeting.