

EMPLOYMENT

Data Science Associate **Pew Research Center** **2017 - Present**

Data Labs

- Led a team of 4 in a research project from the initial idea generation to the design of back-end infrastructure and data analysis. Trained and deployed a neural network model to more than 100K images using Transfer Learning. Python/R/Django/AWS.
- Conceptualized and developed a geodatabase that allows researchers to study how attitudes of survey respondents overlay with geographic components.
- Authored and coauthored blog posts that aim to introduce data science/programming to social scientists <https://medium.com/pew-research-center-decoded>
- Collaborated with subject teams to develop joint research projects.
- Organized and presented in Brownbag to introduce data science to Center-wide audience to complement their survey work.

Fellow **Data Science for Social Good** **2016**

- Built a machine learning prediction model to predict students at risk of interacting with the juvenile criminal justice, and identify key risk factors.
- Contributed end-to-end data pipeline that include data cleaning, ETL process, model training, deployment and inference. Model performs 87% better than the baseline in identifying students at risk. Python/SQL/Command Line Tools

Teaching Assistant **UC San Diego** **2012 - 2017**

- Served as a teaching assistant for 11 undergraduate courses, primarily on introductory and intermediate microeconomics sequences

Analyst **BlackRock** **2008-2010**

Solutions Center

- Supported external users of BlackRock's risk management software enterprise system. Specific tasks included writing SQL queries to identify trades that were booked, provided software training to external users.
- Worked with Relationship Management group to support sales and implementation for clients

Engineering Intern **Intel** **2006**

Manufacturing Engineering Group

- Developed metrics to measure equipment and factory capacity and output, identified run rate improvement projects, established accident prevention measures and planned training programs for personnel

RESEARCH

Non-Peer Reviewed

- **Gender and Jobs in Online Image Search** Pew Research Center (with Stefan Wojcik, Brian Broderick and Adam Hughes) Covered by: MarketWatch, The Washington Post, Mother Jones

Academic

- **Advertisers Capture: Evidence from Hong Kong** (Submitted) Covered By: Marginal Revolution
- Initiate and finish project from beginning to end Code in R to construct a measure of political polarization on the newspaper articles
- Code in python to scrape news articles, clean and extract useful information, run multinomial logistic regressions to estimate and calculate the financial loss that a pro-Democracy newspaper suffered
- **Measuring Subjectivity in History Textbooks** (Submitted) Covered By: BBC

- Write Python scripts to clean the text, count adjectives, calculate the ratio of positive to negative words and employ word embedding (word2vec)

Conferences

- **2019:** AAPOR, ICA
- **2018:** APSA, APSA Computational Social Science Pre-Conference
- **2016:** Tobin Project

OTHER PROJECTS AND EXPERIENCES

Quantitative Researcher **UC San Diego** **Dec 2016-Apr 2017**

- Built a machine learning prediction model to identify factors leading to academic success using data from local high schools

Statistical Consultant **Wisers** **Jan 2017-Feb 2017**

- Guided the construction of a social media sentiment index on political candidates in Hong Kong

Instructor and Organizer **UC San Diego** **Nov 2016**

- Taught two machine learning workshops to economics PhD students that covered basic concepts and different classification algorithm, and application using sklearn

EDUCATION

- UC San Diego: PhD in Economics, May 2017
- University of Tokyo: Masters in Public Policy, Apr 2012
- UC Berkeley: BEng in Industrial Engineering, Dec 2007

LANGUAGES AND SKILLS

- Programming: Python, R, SQL, Django, AWS
- Data Science: Machine Learning, Deep Learning, Machine Vision, Natural Language Processing
- Language English, Chinese, Japanese.