

Jingfei Cai-Pincus, Ph.D.

SKILLS

Data Science: Python (including libraries such as NumPy, pandas, SciPy, scikit-learn, statsmodels, matplotlib, seaborn, requests, nltk), Hadoop, Spark, R, SQL, SAS

Development: Git, Linux/Unix command line, shell scripting

Cloud Computing: AWS EC2

EXPERIENCE

GEICO | Chevy Chase, MD | January 2018 – present

Data Scientist | March 2019 – present

- Serve as technical lead in a customer retention project, having completed the first iteration of productionalizing machine learning models on the cloud. This includes creating analytical datasets from raw transactional data, data QC, designing, implementing, automating, and documenting data process workflows, and communicating with stakeholders/internal customers.
- Mentor junior team member and conduct code review.

Product Modeling Analyst | January 2018 – March 2019

Contributed to two rated location fraud prevention initiatives and performed other ad-hoc analyses as needed:

- Performed analyses and provided recommendations that led to the expansion of a rated location fraud prevention initiative to 26 additional states; streamlined and automated a major part of the recurring batch process.
- Evaluated a new vendor product for identity and rated location fraud prevention at point of sale.
- Performed ad-hoc analyses related to customer information, such as moving violations and credit.

General Assembly Data Science Immersive | Washington, DC | *Student* | June 2017 – September 2017

Participated in a 12-week, full-time, project-based training program that focuses on statistics, data analysis and visualization, and machine learning. Projects include:

- **Capstone: Topic Evolution in Life Sciences Research** (extracting data from XML files; Natural Language Processing; dynamic topic modeling).
- **Liquor Sales + Linear Regression** (linear models to predict sales; cross-validation to evaluate model performance; regularization tactics to compare model performance).
- **Web Scraping & Classification** (web scraping; classification models including random forest, logistic regression, k-nearest neighbors, and support vector machine).
- **Predicting Terrorism/Bayesian Inference** (Markov Chain Monte Carlo; filling in missing data).

The Journal of Visualized Experiments | Cambridge, MA | *Scientific Video Editor* | June 2013 – April 2017

- Edited manuscripts and accompanying videos (over 1000) from various scientific disciplines (*e.g.* medicine, biology, neuroscience, immunology, behavior, bioengineering, engineering, chemistry, and environment).
- Addressed comments from authors and enhanced presentation of research methods and results.
- Performed quality control on manuscripts/videos as a final step before publication.

Boston College, Department of Chemistry | Chestnut Hill, MA | *Research Assistant* | July 2006 – April 2013

- Designed and implemented biochemical/biophysical research studies with focus on protein-membrane interactions.
- Analyzed data collected from experiments and summarized research progress in forms of slideshow/poster presentations, written reports, abstracts to scientific conferences and manuscripts to journals.

EDUCATION & TRAINING

Boston College, Graduate School of Arts & Sciences | Chestnut Hill, MA

Ph.D. in Chemistry | May 2013

Peking University, College of Chemistry and Molecular Engineering | Beijing, China

Bachelor of Science in Chemistry | July 2005