Abbie M. Popa, Ph.D.

abbiepopa@gmail.com
401-440-5228
linkedin.com/in/abbiepopa
github.com/abbiepopa

Experience

Merck Remote
Associate Director of Data Science September 2021 - Present

- Led build by a team of business analysts and software engineers of a python-based web-app that replaced a cumbersome excel model used for market budget optimization of Merck's suite of oncology products, speeding time to value by 50%.
- Led design and implementation of a robust parallel processing engine for use by different web apps within the company, in one case decreasing runtime from over 6 hours to under 15 minutes per run.
- Built an internal interactive web-app that generates range forecasts based on Monte Carlo simulations for Merck's HPV and RSV product-lines. Users generated over 50 simulations and visited the app over 3,000 times during one cycle of long-range operations planning.

App Annie San Francisco, CA
Staff Data Scientist April 2021 – August 2021
Senior Data Scientist August 2019 – April 2021

- Developed machine learning models and led small team of data scientists and data engineers to enhance App Annie's core b2b product.
- Increased the accuracy of our core model in predicting extremely well-performing app launches, resulting in a significant decrease in customer support tickets (Patent US-20220311834-A1).
- Modeled user acquisition by apps across four paid and organic channels (blog post).
- Modeled likely causes of anomalous events in the Insights (formerly Data Stories) product (blog post).

Cogitativo (healthcare consulting start-up) Berkeley, CA
Data Science Intern April 2019 – June 2019

- Identified non-recommended procedures from millions of records on AWS Redshift.
- Built a Tableau dashboard to visualize anomalous behavior of home healthcare providers.

The Data Institute at the University of San Francisco San Francisco, CA
Data Science Post Doctoral Fellow August 2018 – July 2019

- Described nodes in brain networks of patients with schizophrenia and healthy controls using self-supervised machine learning (feature embeddings) resulting in a Network Science publication (video).
- Consulted on student intern projects with private sector companies on topics including computer vision, natural language processing, and smart-bidding.
- Managed projects resulting in two published IEEE COMPSAC proceedings using PySpark and SparkML to make predictions on EEG data (pdf 1, pdf 2).

University of California at Davis Davis, CA
Ph.D. Researcher September 2012 – June 2018

- Used mixed-effects linear modeling, k-means clustering, and ICA to interpret participants' behavior and brain activity resulting in 4 publications and 18 conference presentations (google scholar).
- Trained and mentored 6 junior research assistants and 7 volunteer interns.
- Member of the UC Davis Data Science Initiative.
- Founding member of the Davis Incubator Group, a student-run community for Ph.D. students interested in transitioning to data science careers.

Technical Skills

Python | R | git | PostgreSQL | PySpark | scikit-learn | pandas | NumPy | Agile Development | Stakeholder Communication
Avid proponent of open source, contributed to sklearn and pandas, focusing on updating existing code to the latest gold standards in python. Co-author on the paper behind the opensource package multi-node2vec.

Education

Ph.D., University of California at Davis, Neuroscience
Honors Sc.B., Brown University, Cognitive Neuroscience