JENNIFER STISO, PHD

data scientist ⋄ jeni.stiso@gmail.com ⋄ jenniferstiso.com ⋄ github.com/jastiso

SKILLS

Programming: Python (sklearn, pandas), R (tidyverse, shiny), SQL, Javascript
Fields: machine learning, statistics, graph theory, clinical bioinformatics, social science

EXPERIENCE

Data Scientist, Myriad Genetics   *Feb 2023 - present*

- Independently managed 5 stakeholder relationships, identified 4 new project opportunities, defined 4 project scopes, executed 4 data-driven solutions, and consistently implemented agile methodologies (Jira for project management) contributing to strategic business objectives.
- Designed a comprehensive project plan for a generative AI use case, presented by the CTO at an industry conference, showcasing strategic vision and innovation in AI technology.
- Utilized SQL (Snowflake), Python, MLFlow, and Dataiku to create reusable features from intricate healthcare and insurance data, contributing to enhanced data accessibility and model performance.
- Developed a random forest model predicting payment amounts with less than 5% monthly error, experiment tracking, feature engineering, and model explanations to increase the efficiency of the accounting team.
- Used causal modeling to evaluate the payor team's model of contract prioritization and propose improvements, ultimately increasing the value of the contracts won by the team.
- Leveraged both internal and external data sources to identify the affinity of different demographic groups to different marketing channels, allowing for more effective ad targeting.
- Led the completion of the data science team's first project, presenting results to 7 varied business teams (including product, revenue, technology, and marketing), which increased awareness and understanding of data science among business leaders.
- Conducted comprehensive code reviews, ensuring collaboration, code quality, and adherence to best practices within the team.
- Mentored an intern in designing and managing two consecutive projects, one of which resulted in the segmentation of IT security roles, saving the IT team more than 100 hours of work.

Data Scientist, Children's Hospital of Philadelphia   *March 2022 - Jan 2023*

- Designed and implemented two new features on a dashboard summarizing the environmental impact of anesthetics for clinical stakeholders, resulting in a notable 28% reduction in greenhouse gas emissions post-deployment.
- Developed and assessed a flexible pipeline for causal inference (double machine learning) of medicine-based interventions in cardiac events advancing the recommendation capabilities of an existing decision support tool.
- Conducted a comprehensive review of 25 explainable AI papers, actively contributing to a structured review authored by a team of four members.
- Established, documented, and managed a Postgres database housing electronic health record (EHR) data for over 58,000 patients, supporting and streamlining projects led by two different colleagues.
- Tested and documented user demos for two decision support products, both featured at the American Medical Informatics Association's AI showcase.

Data Scientist, Univ. of Pennsylvania Bioengineering   *Jan 2021 - March 2022*

- Led a development team to design and launch web-based behavioral experiments, engaging thousands of participants through Amazon Mechanical Turk, utilizing JavaScript, SQL, and Heroku, enabling the collection of larger datasets.
- Collaborated with the Perelman School of Medicine to quantify gender and racial bias in medical student evaluations, contributing to the assessment of equitable trainee feedback.

- Led the development of a Python software package with unit tests for network control theoretic analysis, enhancing the accessibility and usability of the tool.
- Formulated an agent-based network model to analyze citation behaviors within the scientific community, providing insights into drivers of inequality in science, creating a simulated environment to test the impact of potential interventions.
- Directed development and deployment of suite of tools (Google Chrome Application and Gender Diversity Statement and Code Notebook) to reduce gender bias in scientific citations. These tools have garnered over 500 users, resulting in an average reduction in user citation bias of 20%.
- Led a team of 5-10 contributors in the creation of tools for citation transparency during the Organization for Human Brain Mapping Hackathon, emphasizing the importance of openness and clarity in scholarly references.

PhD Candidate, Univ. of Pennsylvania Neuroscience *Aug 2017 - Jan 2021*

- Collaborated with four labs across three universities to pioneer a novel multi-modal, dynamic, graph model for identifying links between gene expression and brain activity, allowing for novel methods to identify gene targets.
- Led a team of 2 technicians to design and implement experiment to probe human relational learning and fit an information-theoretic model of human behavior, testing the validity of fundamental assumptions in theories of learning.
- Developed a processing pipeline with custom quality checks for large publicly available dataset of neural EEG timeseries in over 300 individuals (.5Tb) with epilepsy, allowing for broader use of the dataset.
- Conducted hypothesis tests, akin to A/B tests, to identify models that best fit neural data, enhancing our understanding of the principles guiding brain activity.
- In collaboration with leading mathematician, successfully developed tools from graph theory, network control theory, and non-negative matrix factorization to model the spread of electrical stimulation and BCI control in the brain, contributing to a data-driven method for developing therapies.
- Authored > 20 research articles (5 first author), with > 500 citations.

Intern, Johns Hopkins Applied Physics Lab - Intelligent Systems Group *July 2020 - Oct 2020*

- Independently led research project modeling activity spread in biological neural networks, expanding existing knowledge of static activity.
- Contributed open-source Python code implementing common graph models in larger connectomics datasets (10^6 connections) facillitating faster and more robust studies.
- Advised interns on writing, data visualization, research program investigating the impact of biological neural connection motifs on weight-agnostic artificial neural networks.

Consultant for Medical Device Startup, Univ. of Pennsylvania *Aug 2019 - Jan 2020*

- Quantified early adoption market and execution strategy for small health-tech startup in the Philadelphia area that specializes in neurofeedback devices.

EDUCATION

University of Pennsylvania - PhD in Computational Neuroscience *Jan 2021*
University of California at Berkeley - BA in Molecular Biology & Cognitive Science *Aug 2016*

SELECT INVITED TALKS AND PODCASTS (5/10)

Obtaining Unbiased Datasets in Healthcare *Virtual. 2023*
Data Unchained Podcast.
Building Ethical AI Models in Healthcare. *Salt Lake City, UT. 2023*
BioBytes Talk Series.
Effects of Interictal Discharges on Functional Connectivity. *Philadelphia, PA. 2020*
Women in Data Science Conference.
Network Models of Brain Structure, Function, and Control. *Rome, Italy. 2019*

Organization for Human Brain Mapping: Data Science in Neuroscience Symposium.
Using Control Theory to Model Direct Electrical Brain Stimulation. *Paris, France. 2018*
Network Science: Networks in Big Data and Personalized Medicine Satellite.

LEADERSHIP AND TEACHING

Instructor for Advanced Methods in Graph Analysis, JHU Applied Physics Lab *2023*
Lectured and led code tutorial on network control theory
Guest Lecture Introduction to Data Science, UC Berkeley *2023*
Lectured and led discussion on Best Practices for Building Ethical Models
Co-Director of Graduate Led Initiatives and Activities, Univ. of Pennsylvania *2020*
Negotiated an increase in funds totaling 41% of initial budget ($8,000) from three separate funding sources within the University of Pennsylvania
Python Data Science Bootcamp *2019*
Designed and delivered one 3 hour lecture on the Pandas package, and one 1 hour lecture on machine learning with the SciKit Learn package for graduate students attending the Python Data Science Bootcamp
Guest Lecture and Teaching Assistant, University of Pennsylvania *2018-2021*
led discussion section for 25 students in introductory neuroscience course; gave 3 graduate-level guest lectures on mathematical models of brain activity

RECENT AWARDS

Teammate Appreciation Award, Myriad Genetics *2023*
NGG APICAL Honorable Mention, University of Pennsylvania *2021*
program level award for leadership in outreach programs
Ruth L. Kirschstein National Research Service Award, University of Pennsylvania *2020*
national level PhD funding totaling $46,000