Jake Sauter

78 Old Sylvan Lake Rd (845)-891-9910
Hopewell Jct, NY, 12533 jns4001@med.cornell.edu

EDUCATION

Weill Cornell Graduate School of Medical Sciences, New York, NY February, 2022
M.S., *Computational Biology*
Molecular Biology, Analysis of Next Generation Sequencing Data, Applied Machine Learning, Quantitative Genomics and Genetics, Data Structures and Algorithms for Computational Biology

State University of New York at Oswego, Oswego, NY December, 2019
B.S., *Applied Mathematics, Computer Science*
Statistical Genetics Capstone, Natural Language Processing, Systems Programming, Topics in Artificial Intelligence, Non-parametric Statistics

SKILLS

Data Analysis: NGS Data QC and Analysis, Data Visualization, R Tidyverse
Machine Learning: Deep Learning with Keras and Pytorch, Scikit Learn, Tidymodels
Code and Collaboration: Linux + Git, R + RStudio, Python + VScode, R Shiny Apps

RELEVANT EXPERIENCE

Northeast Information Discovery, Canastota, NY March 2019 - August 2020
Data Scientist

- Designed and validated production ready deep-learning models for applications to radio communications
- Developed robust and feature-rich in-house **R packages** for rapid model development, data augmentation and model performance analysis
- Tested and debugged R and **Python packages** through **Git workflows**, facilitating efficient collaboration and quick turn-around times

Center for the Neural Basis of Cognition, Pittsburgh, PA June 2018 - August 2018
Research Intern

- Documented research progress thoroughly, developing resources for successful project handoff and continuation within the lab
- Collaborated with lab members through thoughtful research-focused discussions during weekly lab meetings, paper presentations, and project updates
- Presented a scientific poster entailing study findings to the CMU research community during an open-door research poster session

Depaul Medical Informatics REU, Chicago, IL May 2017 - August 2017
Research Intern

- Analyzed effects of multi-rating ground truth aggregation techniques on model performance of diagnostic prediction interpretations trained on leading Lung Nodule database
- Showcased researched outcomes in a poster presentation for the scientific community as primary author at SPIE Houston 2018 conference