

MICHELLE GILL, PH.D.

Senior AI Scientist, Life Sciences

✉ michelle@michellelynnngill.com

📍 New York, NY

📄 [curriculum vitae](#) and [publications](#)

🌐 [mlgill](#)

🌐 [themodernscientist](#)

in [michellelynnngill](#)

🐦 [modernscientist](#)

EDUCATION

Ph.D. Molecular Biophysics & Biochemistry

2006, Yale University, *New Haven, CT*

B.S. Biochemistry, *Summa Cum Laude*

2001, University of Kansas, *Lawrence, KS*

SKILLS

Programming Libraries: Python, PyTorch, Keras, Scikit-Learn, Pandas, NumPy, SciPy, Shell

Scientific Techniques & Tools: proteomics (Pyteomics, MSFragger, Trans Proteomic Pipeline), cheminformatics (RDKit), molecular docking and modeling (Smina, Glide, Maestro, PyMOL, Chimera, Open Babel)

Databases: SQL, MongoDB, GraphQL

Other: Unix, Git, Docker, Kubernetes, AWS

AWARDS

- [Ruth L. Kirschstein National Research Service Postdoctoral Fellowship](#)
- [NSF Graduate Research Fellowship](#)
- [Barry M. Goldwater Scholar](#)
- Kansas Board of Regents full-tuition merit scholarship

EXPERIENCE

Life Sciences Senior AI Scientist 2019 – Present
NVIDIA

- Tech lead for proteomics and cheminformatics HPC / deep learning projects and collaborations
- Developed deep learning model to predict peptide spectral matches (PSMs) with >95% F1; *manuscript in preparation*
- Used GCNN deep learning model to predict molecular properties; team finished 33rd in [Kaggle competition](#)

Senior Machine Learning Engineer 2019

Senior Data Scientist 2018 – 2019

BenevolentAI

- Built proteo-cheminformatics and deep learning (3D CNNs) platform to prioritize targets based on structure based drug design methods for target identification triage
- Used matrix factorization to predict required drug mechanism of action for disease-target predictions

Senior Deep Learning Consultant 2017 – 2018

NVIDIA

- Designed and implemented proof-of-concept experiments and DL pipeline for clients in pharmaceutical, materials science, and consumer products industries
- Automated setup and maintenance of AWS infrastructure for client work

Senior Data Scientist 2016 – 2017

Metis

- Created 12-week machine learning course for F100 company and curricula on Spark machine learning and NLP
- Instructor for corporate training and quarterly bootcamps

Research Scientist 2014 – 2016

National Cancer Institute

- Developed [NESTA-NMR](#) in C, which uses compressed sensing to enable ~10X faster acquisition of large data sets
- Created website and [documentation](#) for NESTA-NMR

Postdoctoral Research Fellow 2008 – 2014

Columbia University Medical Center

- Elucidated multi-substrate kinetic and thermodynamic pathways for AlkB, a DNA repair enzyme
- Used Monte Carlo simulations to model effect of enzyme dynamics on chemical reaction rates

SELECTED PRESENTATIONS

- "Real Time, GPU-Accelerated Analysis & Visualization in Life Sciences", Invited Keynote, Ken Kennedy Institute Data Science Conference, October 2020 (*forthcoming*), Virtual
- "Artificial Intelligence Driven Drug Discovery", NYC R Conference, Invited Keynote, 2019, New York, NY
- "[Accelerating the journey from data to medicine](#)", *NeurIPS* Expo, 2018, Montreal, Canada
- "[Artificial intelligence as a catalyst for scientific discovery](#)", *JupyterCon*, Invited Keynote, 2018, New York, NY