

# MICHELLE GILL, PH.D.

Data scientist, biophysicist

📍 New York, NY

✉ michelle@michellelynnngill.com

📄 curriculum vitae and publications

🔗 [mlgill](#)

🌐 [themodernscientist.com](#)

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:** Python, Matlab, C, R, Shell

**Big data:** PySpark

**Machine learning:** Supervised algorithms, Clustering, Natural language processing, Principal component analysis, Signal processing, Computer vision, Neural networks, Deep learning, Compressed sensing

**Statistics:** Regression (linear, logistic, non-linear), Monte Carlo simulations, Bayesian methods

**Databases:** PostgreSQL, MongoDB, GraphQL, Neo4J

**Other:** Unix, Git, AWS, Docker, Kubernetes, Flask, LaTeX

## 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

**Senior Data Scientist**

2018 – Present

*BenevolentAI*

- Used matrix factorization and graph convolutional neural networks (GCNNs) to understand drug mechanisms
- Utilized deep learning (3D CNNs) and chemoinformatics to predict ligand pose and affinity for a given target
- Toolkit: TensorFlow, PyTorch, RDKit, Docker, Kubernetes

**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*

- Conducted corporate trainings and co-instructed 12-week bootcamps focused on Python and Spark
- Developed 12-week machine learning course for F100 company and Spark machine learning and NLP curricula

**Research Scientist**

2014 – 2016

*National Cancer Institute*

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

**Postdoctoral Research Fellow**

2008 – 2014

*Columbia University Medical Center*

- Used Monte Carlo simulations to model effect of physical changes on enzyme activity
- Studied anti-oncogenic associated mechanisms of substrate binding to AlkB, a DNA repair enzyme

**Consultant**

2006 – 2007

*The Boston Consulting Group*

- Developed Excel-based statistical tools and Access database for organizational streamlining of pharmaceutical client
- Part of team that developed municipal bond investment strategy for financial services client
- Member of team awarded 2007 Global Strategy Olympics Prize for pharmaceutical client work

## PRESENTATIONS

- "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
- "From structural biology to AI: a holistic approach to studying molecular machines", *Brookhaven National Laboratory*, Invited Presentation, 2018, Upton, NY