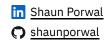
# Shaun Porwal

🖲 New York, NY 🌭 (732) 318-7592 🚇 shaun.porwal@gmail.com 🔗 https://shaunporwal.com 🏳 U.S. Citizen





May 2019

#### Education

Icahn School of Medicine at Mount Sinai

Master's in Biomedical Data Science (MS)

March 2021 **Rutgers University - New Brunswick** 

Bachelor's in Biomedical Engineering, Chinese (BE)

## **Projects**

`dcurves` Python Library - https://decisioncurveanalysis.org

- Developed Python package (GitHub Link, PyPI Link) implementing Decision Curve Analysis to evaluate binary and survival models; over 25k downloads
- Built and maintain decisioncurveanalysis.org, an instructional website with a technical forum for DCA users, answering advanced user questions (links:  $\underline{1}$ ,  $\underline{2}$ ,  $\underline{3}$ )

# **Experience**

### Memorial Sloan Kettering Cancer Center

March 2021 to Present

Data Analyst/Engineer

New York, NY

- Engineered data pipelines with SQL-extracted institutional data and wrote statistical code for Amplio, allowing sarcoma, melanoma, whipple, gastrectomy, and liver surgeons to see patient outcomes
- Developed AI-driven radiology pipeline, fine-tuning Llama3.1 and image segmentation models, with results visualized via an interactive RShiny dashboard
- Led statistical analyses for a landmark study on post-chemotherapy RPLND policies in testicular cancer, applying cuttingedge methods (GAM, logistic regression, Kaplan-Meier, CoxPH, DCA) to produce actionable insights (coming soon)
- Conducted a hospital-wide patient comorbidity analysis using survey data, validating with Posit's 'pointblank,' and demonstrated the survey's effectiveness in improving billing accuracy for patient visits

#### Sema4

February 2020 to February 2021

Bioinformatics Intern

Stamford, CT

- Automated large mutation detection for whole genome sequencing (WGS) data with Bash, DRAGEN, and AWS S3/EC2
- Designed a pipeline to clean and compare results from mutation detection algorithms (e.g., DRAGEN) to explore the origin different pancreatic tumors
- Visualized genomic data using R & Python to present findings to my research group and capstone defense committee

# Languages

Mandarin Chinese Hindi Spanish Japanese 

#### **Publications**