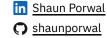
# Shaun Porwal

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



# **Open-Source Projects**

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

August 2021 - Present

(GitHub Link, PyPI Link) over 29k downloads, evaluates binary and survival models; built and maintain instructional website + technical forum, answering advanced user questions (links:  $\underline{1}$ ,  $\underline{2}$ ,  $\underline{3}$ )

`no-more-sql` - https://github.com/juntotechnologies/no-more-sql - Temporarily Privated

March 2024 - Present

Developed a fine-tuned Llama + RAG pipeline on an Nvidia GPU server using Streamlit, FAISS, and Llama 3.1:70B to provide secure text-to-SQL conversion for my organization, servicing 50+ Data Scientists

#### **Experience**

#### Memorial Sloan Kettering Cancer Center

March 2021 to Present

Machine Learning 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 radiology tool with fine-tuned Llama 3.1:70B, trained SAM image segmentation models, & dashboard
- Led statistical analyses for a landmark study on post-chemotherapy RPLND policies in testicular cancer, applying statistical methods (GAM, logistic regression, Kaplan-Meier, CoxPH, DCA) to inform surgeon decisions

## Sema4

February 2020 to February 2021

Bioinformatics Intern

Stamford, CT

- Engineered automated pipelines with Bash, Python, WDL, and AWS to detect structural variants in 100+ 100GB+ BAM files using GitHub tools (Delly, SvABA) and Illumina DRAGEN
- Identified repeat variants in 100+ samples, contrasting insulin vs non-insulin pancreatic neuroendocrine tumors (PNETs) for diabetes therapies and visualized results with Circos, stacked bar, heat map, and scatter plots

### **Education**

Icahn School of Medicine at Mount Sinai

March 2021

**Rutgers University - New Brunswick** 

May 2019

Master's in Biomedical Data Science (MS)

Bachelor's in Biomedical Engineering, Chinese (BE)

## Languages

**English** 

Mandarin Chinese

Hindi

Japanese

Spanish

### **Publications**