

Samar Josyula

📍 Washington, D.C. 📞 571.528.1718 📩 samarjosyula@outlook.com

Profiles

👤 [SamarJ03](#)

linkedin [samar-josyula](#)

Education

University of Pittsburgh

May 2027

Computational Biology

Projects

LLM4DD

Anti-PGCC Drug Discovery Using High-Throughput Screening and LLM-Augmented In Silico Modeling

Designed and implemented a hybrid pipeline integrating empirical single-cell imaging and machine learning to identify novel inhibitors of therapy-resistant polyploid giant cancer cells (PGCCs). Leveraged the LLM4SD framework to synthesize phenotypic, transcriptomic, and cheminformatic data for enhanced feature generation, compound prioritization, and mechanistic insight. Validated novel PGCC inhibitors and contributed to the development of a scalable, interpretable AI-driven drug discovery approach.

QSAR4PD

PPAR-γ QSAR Modeling for Anti-Parkinson's Neuroprotective Drug Discovery

Designed and implemented a QSAR pipeline to predict pEC50 bioactivity of PPAR-γ agonists for neuroprotection in Parkinson's disease, using RDKit, Mordred, and xTB for descriptor generation. Performed ChEMBL-based data curation (EC50 filtering, SMILES cleaning, Lipinski's Rule), feature selection (low-variance filtering, RFE, correlation pruning), model training (scikit-learn, XGBoost), and virtual screening of drug-like compounds.

Experience

UPMC Montefiore

May 2024 – June 2025

Pittsburgh, PA

Patient Care Technician

- Provided routine patient care and supported the departmental services in accordance with established policies and procedures.
- Worked in UPMC Montefiore unit 10EW. A 30 bed acuity adaptable med-surg care unit. A diverse patient population with ranging cases from hepatology, urology, endocrinology, pulmonary infectious diseases, and more.
- Applied a diverse skillset including phlebotomy, EKG administration, and other clinical procedures while adapting to the demands of a high-acuity environment.

MedStar Health

May 2023 – August 2023

Washington, D.C.

Data Analyst Intern

- Worked under Dr. Nawar Shara, predominantly on cardio-metabolic risk prediction in maternal health.
- Various risk factors, including race/ethnicity, previous disease history and test results of varied bodily measures were databased in order to better provide medical care to those at risk for maternal death and infections.
- Focused on using multiple data-dictionaries together and writing software in python and R to better portray these statistics in a larger, more broad database.

Skills

Programming Languages

Python, Java, R, SQL

Libraries and Frameworks

Pandas, Polars, MongoDB, NumPy, seaborn, matplotlib, Requests

Bioinformatics and Cheminformatics

RDKit, Mordred, PaDEL, xTB, PubChemPy, OpenBabel, ChEMBL, Uniprot, MACCSkeys, ECFP4

DevOps and Tools

Git, GitHub, Conda, Docker, Jupyter, RStudio

Machine Learning and LLM

openai SDK, Transformers, Pipeline, LangChain, Ollama, scikit-learn, XGBoost, LightGBM

Modeling and Simulation

AutoDock, Vina, PyMOL, SwissADME, admetSAR