

RAEHASH SHAH

PHONE
+1 (201) 449 4918

EMAIL
raehashshah@gmail.com

LOCATION
Tenafly, NJ

PORTFOLIO
<http://shahraehash.github.io>

EDUCATION

CARNEGIE MELLON UNIVERSITY Master of Science in Computational Biology

Pittsburgh, PA
2025

- Teaching Assistant for Machine Learning in Health Care
- Taken Courses in Machine Learning with Large Datasets; Generative AI; Statistical Genomics; Probabilistic Graphical Models; Image-Based Computational Modeling and Analysis; Deep Learning

Bachelor of Science in Computational Biology University Honors - GPA: 3.7

2024

- Teaching Assistant for Computational Medicine & Great Ideas in Computational Biology
 - President of Computational Biology Undergraduate Society
 - Co-Organizer of ISCB (International Society for Computational Biology) GLBIO Conference Workshop
-

EXPERIENCE

COMPUTATIONAL BIOLOGIST INTERN

01/2024 – 08/2024

Remote // GENIE Lifesciences

- Implemented a disease classification algorithm in Python and a PubMed search engine leveraging NLP (Natural Language Processing) and KG (Knowledge Graph) techniques to streamline drug discovery research by at least 25%

BIOINFORMATICS INTERN

05/2023 – 07/2023

Rensselaer, NY // Regeneron

- Constructed a pipeline consisting of a stable diffusion generative model in a GMP (Good Manufacturing Processes) environment to pseudo-create ONT (Oxford Nanopore Sequencing Technology) signal data from viral and CHO host DNA sequencing data for QC Virology testing

SOFTWARE SYSTEMS ENGINEER INTERN

05/2022 – 08/2022

Bedford, MA // Werfen

- Designed and developed a TDM (Time-division multiplexing) Protocol Sniffer for a RS485 Bus using a Docker environment on a throughput hemostasis testing instrument firmware
- Assisted in leading daily standups as a Scrum master in an Agile work environment (Agile Atlassian Certified)

DATA SCIENCE INTERN

07/2019 – 09/2021

New York, NY // Englander Institute for Precision Medicine

- Analyzed tumor heterogeneity across primary samples and organoids at different points of cell culture through Hierarchical Clustering, PCA and Differential Gene Expression Analysis
-

PUBLICATIONS

“SCN9A: Proposal of Voltage-Gated Ion Channels as a Novel Diagnostic Marker for Alzheimer’s Disease” (2023) Shah D., Shah R., Waldron A., Leonardi D., <https://doi.org/10.1101/2023.05.18.23289925>

- Performed GSEA (Gene Set Enrichment Analysis) and Logistic Regression Analysis on RNA-seq data in R to identify factors associated with Alzheimer’s Disease

“Virome Sniff” (2017) Vargas-Asencio J., Nunez R., Musunuri R., Hao Y., McCormack M., Hagerty P., Shah R., Bradic M., Busby B., <https://osf.io/e9mnf/>

- Generated distribution graphs of the abundances of viruses present in the samples analyzed
-

SKILLS

Computational Tools: BLAST, Bioconductor, Clustal, Pytorch, Anaconda, Pyspark, AWS, Git
Extracurriculars: Scuba Diving, Soccer (Co-Captain), Swim Instructor, DDN Dance Team
Languages: English, Gujarati, Spanish, Hindi