RAEHASH SHAH

PHONE
+1 (201) 449 4918

EMAIL
raehashshah@gmail.com

PORTFOLIO
http://shahraehash.github.io

EXPERIENCE

REGENERON – IOPS 05/2023 – 07/2023
Rensellear, NY // *Bioinformatics Intern*

- Developed a novel machine learning stable diffusion generative model pipeline to create ONT (Oxford Nanopore Sequencing Technology) signal data from viral and host DNA sequencing data.
- Analyzed signal data using Dynamic Time Warping, and a self-developed binary classifier and assessed basecalled signal data by performing sequencing alignment to a reference genome.

WERFEN – INSTRUMENTATION LABORATORY 05/2022 – 08/2022
Bedford, MA // *Software Systems Engineering Intern*

- Designed and developed a TDM (Time-division multiplexing) Protocol Sniffer for a RS485 Bus on a throughput hemostasis instrument firmware.
- Created an automated Peta Linux setup for building firmware for computer architecture of an instrument.

ENGLANDER INSTITUTE FOR PRECISION MEDICINE 07/2019 – 09/2021
New York, NY // *Data Science Intern*

- Analyzed tumor heterogeneity across primary tumor samples and organoids at different points of tumor culture through Hierarchical Clustering, PCA, Differential Gene Expression Analysis, Pathway Analysis
- Compared the use of reference matrices in deconvolution of tumor microenvironments through comparing algorithms of published deconvolution tools and reference matrices.

EDUCATION

CARNEGIE MELLON UNIVERSITY 2020 - 2024
BS Computational Biology - QPA: 3.68
Minor: Machine Learning, Biomedical Engineering
Teaching Assistant for Computational Medicine and Great Ideas in Computational Biology

PUBLICATIONS

"*SCN9A*: Proposal of Voltage-Gated Ion Channels as a Novel Diagnostic Marker for Alzheimer's Disease" Shah D., Shah R., Waldron A., Leonardi D., https://doi.org/10.1101/2023.05.18.23289925

"**The Role of Glycoproteins and HLA Markers in the Fibrosis of Rheumatic Valvular Disease"** Shah R., Leonardi D.,

"Virome Sniff" Vargas-Asencio J., Nunez R., Musunuri R., Hao Y., McCormack M., Hagerty P., Shah R., Bradic M., Busby B., https://osf.io/e9mnf/

SKILLS

Genomics: (sc)RNA-seq analysis, GSEA, Pathway Enrichment Analysis, DESeq2, STAR, Hi-C
ML: Feature Selection, Distributed ML, Hyperparameter Tuning, Regression and Classification
Programming Languages: C/C++, Java, Python, R, Shell/Bash Scripting
Languages: English, Gujarati, Spanish, Hindi