Sahil Arora
linkedin.com/in/sahilarora3
arolihas.github.io

Georgia Institute of Technology Aug 2015 – May 2021
MS Machine Learning GPA 3.88/4.00
BS Computer Science (Intelligent Modeling and Simulation) Highest Honors
CPE Lyon – Certificate in Molecular Biochemistry May 2017 – Jun 2017

EXPERIENCE

Ordaos Bio *AI Scientist II* Oct 2022 – Current
- Incoming role

Roivant Sciences *Technical Rotation Analyst* Jul 2021 – Oct 2022

Sumitovant Biopharma *Machine Learning Engineer* Aug 2022 – Sep 2022
- Implemented on-the-fly near duplicate detection on batch data with sBERT embeddings for competitive intelligence web app

VantAI *Machine Learning Engineer* Aug 2021 – Aug 2022
- Developed geometric deep learning models that characterize the molecular surface of proteins and identify likely non-partner specific binding sites by training on structural features with curated datasets
- Used contrastive learning on dMaSIF embeddings of point cloud representations to achieve 86% accuracy on observed PPI
- Created command line utility and Dockerized endpoint for efficient prediction of potential binding sites on human proteome
- Trained pairwise alignment-free compatibility model that predicted interaction likelihood of interfaces between surface patches on the interface, achieving 99% sensitivity and 65% specificity on key internal targets against E3s
- Validated compatibility model on STRINGdb and BioGRID protein interaction databases, achieving 0.95 and 0.15 enrichment factor respectively (fraction of E3:target interactions with predicted probability > 0.65 that are present in the database)
- Built dashboard for visualizing surface patches, learned embeddings, and chemical features in predicted protein-protein interactions of key targets to accelerate prioritization screening with external clients in drug discovery process

Georgia Institute of Technology *Head Teaching Assistant* Aug 2020 – May 2021
- Managed team of 15 TAs and created assignments and quizzes for 500+ person online masters Deep Learning course

IBM – CIO *Software Engineering Internship* Jun 2020 – May 2021
- Wrote and optimized Selenium tests in Java for a Cucumber test automation framework in a Salesforce platform

Merck – Manufacturing Division Jun 2019 – May 2019
- Developed CNN for classification of defective vaccine vials to capture potentially up to $0.5M per year in lost revenue
- Created Flask web application with SQLite3 backend to upgrade ticket management system for contracting service

Virginia Systems and Technology *Software Engineering Internship* Jan 2019 – May 2019
- Implemented features for an Angular/Typescript/Java web app simulating flight path triangulation of targets

Novo Nordisk – CMR Operations *Data Analytics Internship* May 2018 – Aug 2018
- Built modal interface for modifying budget line items with Excel VBA to optimize Business Management team processes

RESEARCH

Biological Systems and Engineering Laboratory Aug 2019 – Aug 2020
- Applied feature selection with scikit-learn to identify minimal metabolite concentration sets and classify cellular phenotypes

BioMEMS and Biomechanics Laboratory Jan 2018 – Aug 2019
- Implemented fast online peak detection algorithm for AFM probe-surface contact with force measurement data
- Attached gold nanorods to silica particles via thiols that move under directed laser light to study photothermal propulsion

PROJECTS

Motif Embedded Disease Pathway Prediction in Protein-Protein Interaction Networks (Tensorflow+Keras, stellargraph)
- Embedded motif frequency distribution in GCN to improve node classification task from 8% to 56% F1-micro score

Knee MRI Diagnosis Modeling (PyTorch, fastai)
- Applied multiheaded attention and semi-supervised learning to improve AlexNet classification of tears to 95% AUC

SafeSpace Mobile App (Kotlin, Google Cloud, CouchDB, wireshark)
- Built app that reports how busy a place is with WiFi tracking and crowdsourced locations for COVID-19 hackathon

LEADERSHIP

Data Science at GT *External Affairs Member* Aug 2019 – Dec 2020
- Created marketing plan and contacted sponsors as part of a team for Hacklytics data science hackathon

International Society of Pharmaceutical Engineering *Cofounder and Vice President* Aug 2017 – Aug 2019
- Led organization of over fifty members and held networking events to connect students to faculty and industry

SKILLS

Machine Learning: PyTorch, PyTorch Geometric, TensorFlow, scikit-learn, NumPy, SciPy, pandas, matplotlib, plotly
Web: JavaScript (Angular, Node), TypeScript, Flask, d3
Software Engineering: Java, R, SQL, Google Cloud, Argo, Docker, jupyter, conda, git, vim, tmux