Jessy Song

Email: jessys@stanford.edu | Phone: +1(415) 688-0886 | LinkedIn: @jessy-song-1102 | GitHub: @Jessyjias

Summary: ML/AI Engineer and Data Scientist with experience in developing deep learning models (CNN, GNN), specializing in biomedical data analysis, data app development, big data, cheminformatics, and computer vision. Visit my GitHub for recent side projects and past research and LinkedIn for complete work experience.

Work Experience

AI/ML Engineer, Frontier Medicines 06/2022 - 02/2024 | South San Francisco, CA

Machine Learning Model Development

- Developed, trained and fine-tuned a new synthesis complexity scoring model with graph neural network to provide chemical compound rankings on Frontier's AI Drug Design platform, used by all company R&D pipelines.
- Performed rigorous data pre-processing from retrosynthesis software, feature engineering for analyzing biases towards molecular substructures to evaluate model robustness and generalizability.
- Utilized SageMaker, MLFlow and additional AWS suite to conduct extensive training experiments, optimizing performance with feature augmentation using small molecule LLMs such as MegaMolBART.

Data Science and Cheminformatics

- Led development of a compound-centric similarity data dashboard for fragment-based drug discovery, integrating graph database (Neo4j), RDKit, and Plotly Dash for intuitive chemical data visualization and structural analysis.
- Conducted comprehensive analysis of large chemical libraries with distributed parallel computing tools (Spark/ Dask) to manipulate molecular structures, improving internal library diversity and relevance to drug pipelines.
- Spearheaded large-scale chemical similarity search with vector database (Pinecone) to drive efficient compound exploration and improvements of generative models for small molecule design.
- Investigated novel approaches in generative AI for small molecule design and protein structural prediction tools such as AlphaFold, contributing to company's capabilities in target discovery and compound generation.

Medical Device Algorithm R&D Intern, Intellijoint Surgical Inc 05/2021 – 08/2021 | Kitchener, Canada

- Prototyped the robot integration to Intellijoint's surgical navigation system and evaluated potential of robotics-assisted total knee replacement surgery.

Software Engineer Intern, PointClickCare - API Architecture 05/2019 – 08/2019 | Waterloo, Canada

- Implemented RESTful APIs using Spring framework in Java to be used by over 100 vendors and product management teams to track client application and activations.

Research Experience

Computational Neuroscience Research Assistant, BRAIN Lab 01/2023 – 04/2023 | University of Waterloo

- Research project on MouseNet (a biologically-constrained convolutional network modelling mouse visual cortex) trained with Unity-generated videos modelling mouse vision. | Supervisor: Dr. Bryan Tripp

Deep Learning Research Assistant, UW VIP LAB 01/2022 – 12/2022 | National Research Council Canada

- Developed an explainable, deep convolutional neural network with prototypical learning for few-shot classification to detect COVID-19 patients with ultrasound images.
- Manuscript: COVID-Net USPro | Abstract at Med-NeurIPS 2022 | Supervisors: Dr. Ashkan Ebadi and Dr. Alexander Wong

Undergrad Researcher, Computational Biology and Biophysics Lab 01/2022 – 06/2022| Queen's University

- Analyzed mutation patterns for DNA Polymerase Epsilon in endometrial cancer patient genome based on mutational signature exposures using clustering/classification algorithms. | Supervisor: Dr. Anna Panchenko

Education

M.S, Biomedical Data Science, Stanford University 09/2023 – today | Stanford, CA
Part-time via Honors Cooperative Program

B.A.Sc, Honours Biomedical Engineering, University of Waterloo 09/2018 – 05/2023 | Waterloo, Canada
With Distinction; Neural Engineering Specialization, Computing Option

Skills

Python, Pytorch, R, AWS, Spark, Dask, SageMaker, MLFlow, Google Cloud Platform (GCP), Graph DB, SQL, Vector DB, Plotly Dash, Streamlit, C++, MATLAB, React, Git