Stanley Z. Hua

Toronto, Canada
437-986-3444 | stanley.hua@mail.utoronto.ca | linkedin | github | website

Education

University of Toronto — Toronto, CA
H. BSc. in Computer Science (GPA: 3.86/4) — *Sept. 2019 – May 2024 (Expected)*

Technical Skills

Languages: Python, SQL, Shell Script, Git, C/C++, Java, HTML/CSS, R, Assembly
Software Tools: GitHub, Azure ML
Data Skills: Data Cleaning, Deep Learning, Computer Vision, Data Visualization, Clustering, Dimensionality Reduction
Python Libraries: Pandas, NumPy, Scikit-Learn, PyTorch, Tensorflow, Keras, Matplotlib, OpenCV, Dask

Experience

Software Engineer Intern — May 2022 – Present
Intel Corporation — *Toronto, CA*

- Full-stack development to create/maintain internal software tools to assist internal teams in development and performance benchmarking.

Junior Data Scientist — May. 2022 – Present
The Hospital for Sick Children — *Toronto, CA*

- Productionized MLOps code for model deployment, forecasting the number of patients in the ER hourly.
- Improved data validation (quality checks, unit-testing) of time-series data from the ER.
- Implemented saving/loading of a complex nested model to pickle.

AI Research Student — Sept. 2021 – May. 2022
The Hospital for Sick Children (Goldenberg Lab) — *Toronto, CA*

- Adapted deep learning methods from video modeling (Conv.Pooling, CNN-LSTM, TSM) to improve prediction of a kidney disease from medical images (ultrasound) taken over multiple hospital visits using PyTorch.
- Assessed model performance based on key metrics (AUROC, AUPRC) with bootstrapped confidence intervals.

AI Research Student — May 2021 – Sept. 2021
University of Toronto (Moses Lab) — *Toronto, CA*

- Created a large-scale dataset **CytoImageNet** (890K images, 894 classes) from 20 TB of open-source microscopy images (and tabular metadata) using Python (pandas, numpy, opencv).
- Pretrained deep convolutional models (EfficientNetB0) on CytoImageNet using Tensorflow, providing biologists a new means to extract information from microscopy images.
- Paper published and poster presented at NeurIPS 2021 LMRL workshop.
- The CytoImageNet dataset has attracted attention on Kaggle (3716 views, 398 downloads).

ML Research Student — Jul. 2020 – Jul. 2021
University of Toronto (Tyrrell Lab) — *Toronto, CA*

- Investigated the effect of dimensionality (PCA, neural autoencoder) on clustering (K-Means, DBSCAN, Agglomerative) of medical images under small sample sizes using Python (scikit-learn, tensorflow keras).

Conference Presentations

CytoImageNet: A large-scale pretraining dataset for bioimage transfer learning
NeurIPS 2021 (Learning Meaningful Representations of Life Workshop)

Honors & Awards

2021 **University of Toronto CSB Undergraduate Research Award**, $4000

2020-22 **Dean's List Award**