

EDUCATION

Carnegie Mellon University

B.S. in Computer Science; GPA: 3.97

Pittsburgh, PA

Aug 2019 - May 2023

- **Coursework:** ML Systems, Intro to Deep Learning, Compilers, Machine Learning with Large Datasets.
- **Activities:** TA for Parallel Computer Architecture and Programming.

EXPERIENCE

IBM Research Intern

May 2021 - Aug 2021 (Remote)

- Designed a novel algorithm using Ray RPC framework to efficiently train ML pipelines on cloud clusters. Resulted in 3.5x speedup in execution time. Outperforms memory usage of naive implementation by 50%.
- Working on publishing a paper detailing the caching mechanism for common pipeline prefixes.

Bear Robotics Intern

Jul 2019 - Aug 2019 (Redwood City, CA)

- Optimized 3D vision pipeline using CUDA for uniform voxel downsampling.
- Improved global path planning for autonomous robot navigation in C++ and ROS.

PROJECTS

Pystreaming: A Lightweight Python Package for Audio and Video Streaming

Dec 2019 - Present

- Developing Python package for low-latency real-time distributed inference across computer cluster. Leverages ZMQ messaging libraries, utilizes multi-core processing, and implements safe failure patterns.
- Extended project into the summer of 2020 while funded under an REU grant.
- Available for download through PyPI: <https://pypi.org/project/pystreaming>

Warehouse Robots: GPU-Accelerated Training for Reinforcement Learning

Mar 2021 - May 2021

- Wrote custom CUDA kernels to simulate a multi-agent grid-world environment. Built environment that conforms to the API of the OpenAI Gym.
- 17x speedup in overall training runtime, 71x speedup in observation rendering, 103x speedup in environment stepping/reward generation.
- <https://github.com/joseph-x-li/warehouse-robots>

Dynamic Clustering Optimization For Multi-Center Federated Learning

Apr 2021 - May 2021

- Implemented a user clustering based federated learning algorithm that enables varying number of clusters and does not completely rely on participation of all clients.
- Implemented and simulated effect of asymmetric client availability; implemented and simulated effect of weight sharing among clusters.

ScottyLabs

Sep 2019 - Present

- Project lead for printer map and status website that displays the status of all CMU printers.
- Created new API that parses and serves CMU campus printer error codes, signals, and tray statuses.
- Helped build new mentor queue system for TartanHacks hackathon to enable greater mentor participation.

PUBLICATIONS

Detecting an Upward TGF from its Reverse Positron Beam

Jun 2018 - Aug 2018 (UC Santa Cruz)

- Determined the causes of gamma-ray flashes during thunderstorms through Monte Carlo simulation methods.
- Co-authored paper accepted to AGU (American Geophysical Union) Conference, 2018. Published to Journal of Geophysical Research, 2020.

AWARDS

USA Computing Olympiad Platinum Division

AIME 4x Qualifier

USA Physics Olympiad Silver Medal

Physics Bowl Top 10

SKILLS

Languages: Python, C++, C, Rust **Python Frameworks:** PyTorch, TensorFlow, PySpark, Ray, Sklearn, ZMQ