Benjamin E. Jordan

bej9@cornell.edu | 607-339-1740
LinkedIn | Website | GitHub

Skills

Programming: Python, PyTorch, NumPy, C++, CUDA, Scikit-learn, Docker, Java, C#, YAML, Angular
Other Skills: Git, Linux, Unit Testing, Agile Development, LaTeX

Experience

Machine Learning Engineer @ Northrop Grumman AIR Lab **[Jan 2024 - Present]**

- Applied machine learning research and engineering role
- Worked on adapting large language models for use as accurate, domain specific assistants
- Experimented with various knowledge injection techniques on models up to 70B parameters in size
- Relevant technologies used include Python, PyTorch, FSDP, Docker, Git, and Linux

Machine Learning Engineer Intern @ KLA **[May 2023 - Aug 2023]**

- Independently researched and implemented semiconductor defect detection algorithms
- Finetuned and quantized vision transformers for efficient and robust classification with limited data
- Applied bayesian optimization to tune hyperparameters
- Presented project during poster board session and was invited to give a second technical talk

Software Engineer Intern @ Carestream **[May 2022 - Aug 2022]**

- Developed and maintained C# backend functionality in Carestream's ImageView x-ray software
- Gained professional experience with unit testing, agile, git, large codebases, and OOP design

Research Software Developer @ RIT **[Aug 2022 - Jan 2023]**

- Re-hired part-time by faculty to develop software for spatial audio research during final semester
- Independently created a program with 3D graphics for data collection on audio interpretation
- Technologies used include Three.js, Angular, and Typescript

Projects

Generative Text-To-Audio Synth [Personal]

- Adapted Meta's AudioCraft open-source codebase to create a generative text-to-audio synthesizer
- Trained 3.3B parameter autoregressive language model using 3 nvidia A6000s
- Prepared 160GB dataset of audio samples with multi-label and natural language text descriptions
- Analyzed MusicGen & Encodec research to guide hyperparameter, training, and data choices
- Implemented UI with Angular, and used runpod serverless compute for model serving/inference
- Working on implementing infrastructure for collecting human preference data
- Plans to implement efficient, genre specific synthesis with low-rank adapters (LoRA)

Distributed Systems Labs & Framework [Coursework]

- Created a Google spanner-esque distributed key-value store in Java
- Implemented Paxos for replica group consensus, 2PC to achieve atomic commit for distributed transactions, and dynamic load balancing of shards to handle server reconfiguration

GPU Programming Deep Dive [Personal]

- Personal endeavor to solidify knowledge of gpu programming and hw/sw optimization
- Read the book "Programming Massively Parallel Processors: A Hands On Approach, 4th Edition"
- Implemented all CUDA kernels from the book, along with optimizations like tiling

Education

Cornell University [Graduated Dec 2023]
M.Eng. in Computer Science - Concentration in Machine Learning - [3.76 GPA]

Rochester Institute of Technology [Graduated Dec 2022]
B.S. in Computer Science - Presidential Merit Scholarship - [3.94 GPA during second half of degree]

Relevant Coursework

ML: Large Scale Machine Learning, Machine Learning Hardware and Systems, Reinforcement Learning, Mathematical Foundations of Machine Learning, Computer Vision, Natural Language Processing
Sys: Distributed Systems, Parallel Computing, Computer Architecture, Operating Systems, Networks
Math: Matrix Computations, Graph Theory

Activities

RIT Varsity Track and Field (15-20 hrs / week commitment) **[March 2019 - Dec 2022]**