James Ngai

Personal Website | LinkedIn | GitHub
Location: Seattle, WA
Email: jamesdngai@gmail.com | Mobile: 425-429-8347

Education

Carnegie Mellon University
Bachelor of Science, Major in Computer Science, Minor in Mathematics — *Aug 2022 – May 2026*
GPA: 3.91/4.0

Work Experience

Machine Learning Researcher — Jan 2023 – Present
Carnegie Mellon ZUZ Lab — *Pittsburgh, PA*

- Research under Professor Seth Goldstein in developing pricing algorithms for **multi currency crypto transactions**.
- Develop **Agent-Based Model** with **Reinforcement Learning** and **Game Theory** to simulate **real-world** trading.
- Reduced average **convergence time** by **43%** and increased **model reliability 26%** with **Policy Gradient Methods**.
- Implemented **Linux** script on **Google Cloud** for **automated hyperparameter testing**, improving speed by **500%**.
- Named **National Science Foundation SURF Fellow**, awarded **$4500 grant** for **Machine Learning research**.

Teaching Assistant — Feb 2022 – Sept 2022
Expii — *Remote*

- Taught advanced **combinatorics** and **probability** courses for **100+** students in **several different classes**.
- Successfully trained a team of **15+ TAs**, resulting in consistently **outstanding student reviews**.
- Utilized **Data Analysis** skills in **R** to analyze student test results, **significantly enhancing feedback** for each student.

Software Engineering Intern — Jun 2018 – Jun 2022
Allstate Hale Insurance — *Seattle, WA*

- Led development of mailer ad software for **70,000** prospective clients **saving $10000 annually**.
- Corrected **15%** of all **ad data** by **cross-validating** SalesGenie housing data with King County **REST APIs**.
- Improved **Ad Efficacy by 27%** with **XML Requests** and runtime by **75%** with **multithreading** in Python.

Projects

eBay 2023 Machine Learning Competition (Top 10 Result) *PyTorch, NLP, AWS, Google Cloud* Source Code Confidential

- Spearheaded **data preprocessing** for the **eBay Machine Learning Challenge**, leveraging **Pandas** and **Hugging Face** to create training/evaluating/testing pipeline in Python.
- Optimized **NLP** model training scripts using **LoRA**, **Quantization**, and **Deepspeed** to facilitate **efficient Multi-GPU training**, resulting in a **55% reduction** in memory consumption while preserving competitive performance levels.
- Conducted hyperparameter tuning and **K-Cross Fold Validation** on **11 billion parameter models** in distributed training setups on **AWS EC2** and **Google Cloud**.

11-785 PhD Deep Learning Capstone Project *PyTorch, Reinforcement Learning, Computer Vision* Source Code

- Created **Reinforcement Learning** Agent with **Computer Vision** to solve Jigsaw reassembly challenges.
- Developed and implemented **PyTorch Dataclasses** and Datasets, streamlining code through concise structure.
- Improved model generalizability by integrating **AlphaZero** using **OpenAI Gym** framework, surpassing established benchmarks in **both performance and adaptability** by **10%**.

Technical Skills

Languages : Python, R, SQL, C, C++, Java, JavaScript, HTML, CSS, Git, Julia, Linux, LaTeX

Technologies : Pandas, Sklearn, GCP, AWS, TensorFlow, PyTorch, CUDA, Kubernetes, Spark

Relevant Coursework

11-785 PHD Intro Deep Learning , **15-213** Intro Computer Systems, **21-241** Linear Algebra, **21-266** Vector Calculus