Alan (Jia Lin) Yuan

alan.yuan.jly@gmail.com | linkedin.com/in/jalnyn | github.com/jalnyn

Work Experience

Amazon Jun 2023 – Aug 2023

Software Developer - Intern

Toronto, Ontario

- \bullet Designed and implemented a precompute layer to increase recommendation speed by 99% from 150ms
- Created automated data analysis tool to ensure predictions are above 75% accuracy

PAIR Lab (Robotics)

Sep 2021 – Present

Researcher

Toronto, Ontario

- Exploring the usage of MPC, MultiModal LLM in long-term planning of low level task through constant re-evaluation of plans, improving on state of the art: arXiv:2401.04157
- Worked on meta-controllers to assist multi-policy agents to plan long horizon tasks using sparse rewards
- Assisted with **Orbit**, a robot learning framework built on top of **NVIDIA**'s **Isaacsim** resulting in a publication
- ullet Designed and built **GPU** parallelized state systems with low overhead allowing a **4x** speedup over the cpu solution

Amazon

May 2022 – Aug 2022

Software Developer - Intern

Vancouver, British Columbia

- Engineered a modular microservice in Java to send notifications to customer of cashback on select products
- Utilize AWS webservices such as Lambda, SQS and SNS to ensure scalability of the notification system

Intel May 2021 – May 2022

 $Software\ Engineer$ - Intern

Toronto, Ontario

- Developed support software to generate 4000+ of completely random test-cases for edge-case testing
- ullet Optimized support tool's Ram templates to reduce false positives and failing cases by around 70%

Centivizer Apr 2020 - Sep 2020

Software Developer - Part-time

Toronto, Ontario

 Designed and wrote backend application using Node.JS and SimplePeer to connect users via video call EDUCATION

University of Toronto

-cGPA

MSc in Applied Computing

Sept 2023 - Dec 2024

University of Toronto

 $3.84~\mathrm{cGPA}$

HBSc Computer Science Specialist, Major in Mathematics

Sep. 2018 - May 2023

Publications (* equal contribution)

M. Skreta*, Z. Zhou*, **J. L. Yuan***, K. Darvish, A. Aspuru-Guzik, A. Garg. RePLan: Robotic Replanning with Perception and Language Models, *Submitted to (ICLR) 2024 [under review]*

M. Mittal, C. Yu, Q. Yu, J. Liu, N. Rudin, D. Hoeller, **J. L. Yuan**, R. Singh, Y. Guo, H. Mazhar, A. U. Mandlekar, B. Babich, G. State, M. Hutter, A. Garg. ORBIT: A Unified Simulation Framework for Interactive Robot Learning Environments, (RA-L) 2023 PROJECTS

MultiModal AI Story teller (Co-Founder) | private repo

Jul 2023 – Present

- Implemented Multimodality using LLMs and Latent Diffusion Models to build a interactive story teller.
- Lead efforts to increase LLM efficiency by utilizing SOTA quantization to decrease VRAM usage by 87.5%.
- Utilized NLP techniques to summarize context to reduce context size, reducing inference time by up to 10%.
- Grew project from 0 to 50 daily active users resulting in 2k in profits.

CaNetDa: Deep Learning for GeoGuesser in Canada | Link: GitHub

Jan 2021 – Apr 2021

• Mined dataset and trained an ensemble of Computer Vision models: ResNet, EfficientNet and Vision Transformer resulting in a 47% improvement over random agent in predicting province of image in Canada

Machine Learning Course Competition | Link: GitHub

Sep 2020 – Dec 2020

• Achieved the 5th highest score in the unsupervised movie recommendation competition based on Netflix data

Tron UDP Multiplayer | Link: GitHub

Sep 2019 – Dec 2019

• Created a four player game for local networks using the UDP network protocol and C++

• Utilize epoll for both client and server to monitor the socket as well as the timer (server) and stdin (client)

BF-Interpreter | Link: GitHub

Mar 2018 - Nov 2018

• Built a BF shell that runs all example BF programs found on Wikipedia in C TECHNICAL SKILLS

Tech: Deep Neural Networks, Large Language Models, Latent Diffusion Models, Python, C++, JavaScript, Java, Rust Tools: Huggingface, Oobabooga, PyTorch, Git, React, Node.js, MongoDB, SQL, Numpy, GraphQL, Robotics,