Jiayi Yuan

(540) 255-9188 yuancarrieyjy@gmail.com yuanjiayiy.github.io

Education

Master of Science, University of Washington, Seattle, Washington 2022-Present

- Part-time Master in Computational Linguistics
- Courseworks: Deep Learning for NLP, Language Processing Systems and Applications
- GPA 4.0/4.0

Bachelor of Science, Carnegie Mellon University, Pittsburgh, Pennsylvania 2018-2021

- Major in Neuroscience and Computer Science with University Honor
- Courseworks: Intro to Machine Learning for PhDs, Cognitive Robotics, Neural Computations, Computer Systems, Distributed Systems, Advanced Algorithms Analysis and Design

Research Experience

Washington Embodied Intelligence and Robotics Development Lab (WEIRDLab), University of Washington, Seattle, WA, June 2023 - Present

- Co-authored *Scaling Robot-Learning by Crowdsourcing Simulation Environments,* which proposes a pipeline for scaling up data collection and learning generalist policies, by leveraging crowdsource digital twins of real-world scenes using 3D reconstruction techniques and collecting large-scale data in these simulation scenes, rather than in the real-world.
- Engineered an approach of fine-tuning pre-trained foundational models with asynchronous human feedbacks onto the WidowX 250s Robot Arm on short-horizon tasks.

Work Experience

Amazon AGI, Seattle, WA, July 2021 - July 2024

Software Development Engineer II

- Designed, prototyped, and engineered the iDPS CLI, an intuitive command-line interface optimized for initiating and monitoring large-scale dataset processing jobs. This innovation drastically reduced development effort by 80% while enhancing the efficiency of dataset processing.
- Launched and maintained iDPS 2.0, a cutting-edge platform utilizing containerization technology to streamline data processing jobs. This platform significantly accelerated development cycles by 75%, specifically tailored for advanced Natural Language Processing (NLP) models for Alexa.
- Pioneered the development of annotation inconsistency checker tools for Alexa NLU utterance data. Employed advanced SentenceBERT semantic encoding techniques to detect and rectify inconsistencies in labels within training and test datasets. This innovation contributed to improved data quality, vital for training robust NLU models.
- Received distinguished Org-Level Peer Recognition Awards in both 2022 and 2023, acknowledging exceptional contributions to cutting-edge software engineering initiatives.

Technical Skills

Software Engineering: Python, Java, C/C++, TypeScript, MATLAB, Unix, SQL, R, HTML/CSS/Javascript, AWS, GCP, Docker, Git, Conda.

Data Science and Machine Learning: PyTorch, Google JAX, Apache Spark, OpenCV, Tensorflow, Scikit-learn, NumPy, Matplotlib, Pandas.

Robotics and Reinforcement Learning: ROS, Gym, MoveIt, Gazebo, Isaac-sim