

Rohan Ashish Potdar

rohanpotdar138@gmail.com [GitHub](#) [Linkedin](#) [Website](#)

EDUCATION

Purdue University, West Lafayette, IN

August 2020 – December 2023

BS in Computer Engineering; Minor in Mathematics; GPA: 3.81

WORK EXPERIENCE

Anyscale (Software Engineering Intern):

January 2023 – July 2023

- Finetuned [GPT-J-6B](#) using RLHF on distributed clusters using Ray Train
- Built RayGPT, a tool to write optimized Ray code using retrieval-augmented generation
- Scaled batch inference for OpenAI's [Whisper](#) speech-to-text transcription model using Ray Data
- Integrated Whisper with retrieval pipeline to augment internal automated customer insight platform

Anyscale (Software Engineering Intern):

May 2022 – August 2022

- Worked on [RLlib](#), a library for large-scale distributed reinforcement learning, built on top of [Ray](#)
- Implemented Doubly Robust Off-Policy Evaluation in RLlib to evaluate policies on offline data
- Designed [offline RL API](#) for industry customers with recommender systems workloads
- Prototyped end-to-end RL on a GPU using JAX and RLlib, with 10x speedup over CPU-based environments

Teaching Assistant:

- ECE 57000 - Artificial Intelligence
- ECE 36800 - Data Structures

August 2023 – December 2023

August 2021 – December 2021

RESEARCH

Neural-MMO @ Purdue:

August 2021 – December 2021

- Led a student research group focused on multi-agent reinforcement learning for the [Neural-MMO Challenge](#)
- Wrote grant for Google Cloud research credits under Dr. [Mahsa Ghasemi](#) and Dr. [Shreyas Sundaram](#)
- Trained transformer architectures for reinforcement learning using [Proximal Policy Optimization](#) in RLlib

Purdue SWARMS Group:

January 2021 – August 2021

- Simulated drone swarms and tested control algorithms [using Microsoft AirSim](#) under Dr. [Shreyas Sundaram](#)
- Researched multi-agent reinforcement learning through Summer Undergraduate Research Fellowship (SURF)
- Extended Counterfactual Multi-Agent Policy Gradient algorithm to off-policy RL and continuous actions

PROJECTS

Machine Learning:

- Reproducing optimized LLM inference on GPUs using [Staged Speculative Decoding](#)
- Reimplemented Robotic Transformer 1 (RT-1) paper in [PyTorch](#)

Service:

- Contributed to open-source libraries under the Farama Foundation such as [OpenAI Gym](#) and [PettingZoo](#)
- Selected papers and organized presentations on LLMs and RL at the [Purdue ML Reading Group](#)

SKILLS

- Languages: Python, C, C++, Java
- Libraries: RLlib, Ray, DeepSpeed, PyTorch, TensorFlow, Wandb

RELEVANT COURSEWORK

CURRENT: * GRADUATE: +

Reinforcement Learning Theory⁺(A), Artificial Intelligence⁺(A), Robotics⁺(A),
Computational Complexity(A), Large Language Models⁺, Algorithms*