# Fahim Tajwar

Website: https://tajwarfahim.github.io/ Email: ftajwar@cs.cmu.edu

## EDUCATION

**Carnegie Mellon University**                                               **Pittsburgh, PA**
Doctor of Philosophy (PhD), Machine Learning                                      2023 -- Current

**Stanford University**                                                        **Stanford, CA**
Master of Science (MS), Computer Science (AI/ML)                                    2022 -- 2023
Bachelor of Science (BS) with Distinction, Mathematics                              2017 -- 2022

## PUBLICATIONS (* Equal Contribution)

Offline Retraining for Online RL: Decoupled Policy Learning to Mitigate Exploration Bias                    2023
Max Sobol Mark*, Archit Sharma*, Fahim Tajwar, Rafael Rafailov, Sergey Levine, Chelsea Finn
Preprint, 2023

Conservative Prediction via Data-Driven Confidence Minimization                    2023
Caroline Choi*, Fahim Tajwar*, Yoonho Lee*, Huaxiu Yao, Ananya Kumar, Chelsea Finn
ICLR workshops: TrustML, ME-FoMo, 2023

Surgical Fine-Tuning Improves Adaptation to Distribution Shifts                    2023
Yoonho Lee*, Annie S Chen*, Fahim Tajwar, Ananya Kumar, Huaxiu Yao, Percy Liang, Chelsea Finn
International Conference on Learning Representations (ICLR), 2023

When to Ask for Help: Proactive Interventions in Autonomous Reinforcement Learning                    2022
Annie Xie*, Fahim Tajwar*, Archit Sharma*, Chelsea Finn
Neural Information Processing Systems (NeurIPS), 2022

Do Deep Networks Transfer Invariances Across Classes?                    2022
Allan Zhou*, Fahim Tajwar*, Alexander Robey, Tom Knowles, George J. Pappas, Hamed Hassani, Chelsea Finn
International Conference on Learning Representations (ICLR), 2022

No True State-of-the-Art? OOD Detection Methods are Inconsistent across Datasets                    2021
Fahim Tajwar, Ananya Kumar*, Sang Michael Xie*, Percy Liang
ICML Workshop on Uncertainty & Robustness in Deep Learning (UDL), 2021

Scalable deep learning to identify brick kilns and aid regulatory capacity                    2021
Jihyeon Lee*, Nina R. Brooks*, Fahim Tajwar, Marshall Burke, Stefano Ermon, David B. Lobell, Debashish Biswas, Stephen P. Luby
Proceedings of the National Academy of Sciences, Apr 2021, 118 (17)

## RESEARCH EXPERIENCE

**Research Intern, Stanford Artificial Intelligence Laboratory**            March 2020 – June 2023

- (Prof. Chelsea Finn and Percy Liang) Relationship between distribution shifts and the layer of a neural network that should be fine-tuned (surgical fine-tuning) on the unsupervised (test-time) adaptation setting, specifically adapting earlier layers outperform adapting later/all layers for corruption datasets like CIFAR-10-C and ImageNet-C (*Under review in ICLR, 2023*)
- (Prof. Chelsea Finn) Deep reinforcement learning for irreversible environments with applications to episodic, autonomous, and continuous learning setups (*NeurIPS, 2022*)
- (Prof. Chelsea Finn) Generative model-based algorithm that produces performance boost of 1-2% when combined with other state-of-the-art methods on long-tailed versions of datasets like CIFAR, GTSRB, etc. (*ICLR, 2022*)
- (Prof. Percy Liang) Demonstration of out-of-distribution (OOD) detection problem being too broad by showing that many well-known methods don't perform consistently on a comprehensive suite of benchmark datasets (*ICML UDL Workshop, 2021*)

## Computer Vision Research Intern, Stanford University

March 2019 – June 2020

- (Prof. Steve Luby and Stefano Ermon) Built a system using convolutional neural networks (CNNs), that detects environmental regulation violations in the form of brick kilns and produces their co-ordinates from satellite imagery in South Asia (*PNAS, 2021*)
- Detected nearly 10,000 brick kilns in Bangladesh which directly affect the lives of at least 1 million people using the trained CNN in *TensorFlow,* with the possibility of extending the project over to India (with 100,000 kilns) and other developing countries
- Created a classifier to distinguish between environment-friendly (Zigzag) and unfriendly (FCK) type of brick kilns

## TEACHING EXPERIENCE

Teaching Assistant, <u>Math 20 (Calculus)</u>, Stanford University · · · · · · · · · · · · · · · · · · · · · · · · · · Jan 2023 – March 2023
Teaching Assistant, <u>CS 330 (Deep Multi-Task and Meta Learning),</u> Stanford University · · · · · · · · Sept 2022 – Dec 2022
Academic Tutor, Athletic Academic Resource Center (<u>AARC</u>), Stanford University · · · · · · · · · · · Sept 2021 – June 2022
Academic Tutor, Stanford University Mathematical Organization (<u>SUMO</u>) · · · · · · · · · · · · · · · · Sept 2019 – June 2020
Trainer, National Physics Olympiad Team, Bangladesh (<u>BdPhO</u>) · · · · · · · · · · · · · · · · · · · · · · · Feb 2017 – June 2018

## INDUSTRY EXPERIENCE

### Software Engineer Intern, Meta Platforms (formerly Facebook)

June 2022 – September 2022

- As part of the Ads Core ML Eng team, designed and implemented components for state-of-the-art ML recommendation systems in *Python*
- Experimented with various knowledge distillation techniques to improve performance of computationally cheaper ML networks
- Designed new modules with scalability in mind to make sure they work well with extremely large datasets and can also be trained efficiently

### Software Engineering Intern, Cadence Design Systems

June 2020 – September 2020

- Using *C++,* designed efficient data storage systems for graph neural networks (GNN), which reduced memory usage by 86%
- Using *Python* (*NumPy*), designed and implemented efficient data pre-processing modules
- Designed time and memory efficient data-loader classes in *Keras* and *TensorFlow;* researched memory-time tradeoff to find the optimal design

## SKILLS

Programming Languages: Python, C, C++, Java, Matlab
Frameworks: Unix, PyTorch, Caffe2, TensorFlow

## TALKS & PRESENTATION

- Neural Information Processing Systems (NeurIPS) · · · · · · · · · · · · · · · · · · · · · · · · · · · · · November 2022
- International Conference on Learning Representations (ICLR) · · · · · · · · · · · · · · · · · · · · · · · · April 2022
- ICML Workshop on Uncertainty & Robustness in Deep Learning (UDL) · · · · · · · · · · · · · · · · · · July 2021
- Stanford Earth Summer Undergraduate Research (SESUR) · · · · · · · · · · · · · · · · · · · · · · · · August 2019
- Stanford EE Research Experience for Undergraduates (REU) · · · · · · · · · · · · · · · · · · · · · · · August 2018

## AWARDS

University Distinction, top 15% of the graduating class, Stanford University · · · · · · · · · · · · · · · · · June 2022
Tau Beta Pi Engineering Honor Society · · · · · · · · · · · · · · · · · · · · · · · · · · · · · · · · · · · · May 2020
Bronze Medal, 48th International Physics Olympiad, Indonesia · · · · · · · · · · · · · · · · · · · · · · · July 2017
Bronze Medal, 47th International Physics Olympiad, Switzerland Liechtenstein · · · · · · · · · · · · · · July 2016

## SERVICE

- Reviewer, Neural Information Processing Systems (NeurIPS) · · · · · · · · · · · · · · · · · · · · · · · · · · 2023
- Reviewer, NeurIPS Workshop on Distribution Shifts (DistShift) · · · · · · · · · · · · · · · · · · · · · · · · 2023
- Reviewer, International Conference on Learning Representations (ICLR) · · · · · · · · · · · · · · · · · · · · 2024

## SELECTED COURSEWORK

**Mathematics:** Real Analysis, Functional Analysis, Differential Topology, Measure Theory, Probability Theory, Graph Theory, Abstract Algebra, Linear Algebra & Matrix Theory, Statistical Inference, Numerical Computing

**Computer Science:** Machine Learning, Reinforcement Learning, Convex Optimization, Deep Learning for Computer Vision, Artificial Intelligence: Principles and Techniques