Ayush Jain

Research Scientist
Applied Reinforcement Learning, Meta

https://ayushj240.github.io
ayushj240@meta.com

Education

University of Southern California Dec 2024
PhD in Computer Science (Advisors: Joseph J. Lim & Erdem Bıyık)
GPA: 4.0/4.0

University of Southern California May 2024
MS in Computer Science
GPA: 4.0/4.0

Indian Institute of Technology Delhi June 2016
B.Tech in Electrical Engineering
GPA: 8.99/10

Research Interest & Experience: Reinforcement Learning

PhD Thesis: Decision Making in Complex Action Spaces
My goal is to develop *adaptive agents* for both physical and virtual environments that can operate in large, unseen, and complex action spaces. Applications include robotics, organic and ad recommendation systems, and instruction-following agents for environments like Minecraft and Android.

Publications

[ArXiv] X. Li, **A. Jain**, Z. Yang, Y. Korkmaz, E. Bıyık. "When a Robot is More Capable than a Human: Learning from Constrained Demonstrators".

[ArXiv] A. Samanta, A. Magesh, Y. Yu, R. Wu, **A. Jain**, D. Jiang, B. Vidolov, P. Sajda, Y. Efroni, K. Hassani. "Internalizing Self-Consistency in Language Models: Multi-Agent Consensus Alignment".

[**EACL** 2026 Findings] R. Wu, A. Samanta, **A. Jain**, S. Fujimoto, J. Kwon, B. Kretzu, Y. Yu, K. Hassani, B. Vidolov, Y. Efroni. "Imbalanced Gradients in RL Post-Training of Multi-Task LLMs". *19th Conference of the European Chapter of the Association for Computational Linguistics.*

[**NeurIPS** 2025] Y. Korkmaz*, U. Bhuwania*, **A Jain**†, E. Bıyık†. "Actor-Free Continuous Control via Structurally Maximizable Q-Functions". *Neural Information Processing Systems.*

[**RLC** 2025] **A. Jain**, N. Kosaka, X. Li, K. Kim, E. Bıyık, J. Lim. "Mitigating Suboptimality of Deterministic Policy Gradients in Complex Q-functions". *Reinforcement Learning Conference.*
Outstanding Paper Award on Empirical Reinforcement Learning Research

[**ICLR** 2025] G. Zhang*, **A. Jain***, I. Hwang, S. Sun, J. Lim. "QMP: Q-switch Mixture of Policies for Multi-Task Behavior Sharing". *International Conference on Learning Representations.*

[**ICLR** 2022] **A. Jain***, N. Kosaka*, K. Kim, J. Lim. "Know Your Action Set: Learning Action Relations for Reinforcement Learning". *International Conference on Learning Representations.*

[**ICML** 2020] **A. Jain***, A. Szot*, and J. Lim. "Generalization to New Actions in Reinforcement Learning". *International Conference on Machine Learning.*

[COLING 2018 Workshop] **A. Jain**, V. Singh, S. Ranjan, R. Rajkumar, S. Agarwal. "Uniform Information Density Effects on Syntactic Choice in Hindi". *Workshop on Linguistic Complexity and NLP.*

[DGfS 2017] **A. Jain**, V. Singh, S. Agarwal, and R. Rajkumar. "Uniform Information Density models for language production". *39th Annual Conference of the German Linguistic Society, DGfS 2017.*

Industry Experience

Meta, Sunnyvale, USA — Feb 2025 – Present
Research Scientist in Applied Reinforcement Learning Team, Monetization Org
→ Reinforcement learning for optimizing ad bidding and auction as a contextual bandit.
→ Reinforcement learning to optimize the recommendation of ads to users on Meta Platforms.

Meta Reality Labs Research, Redmond, USA — May 2024 – Dec 2024
Research Intern with Nitin Kamra
→ Reinforcement learning based virtual agents for instruction-following in **Android** devices.

Microsoft Research, Montreal, Canada — May 2023 – Jan 2024
Research Intern with Eric Yuan and Marc-Alexandre Côté
→ Combine discrete prompt optimization in large language models (LLM) with gradient-optimization in neural networks for natural-language reasoning tasks and instruction-following **Minecraft** agents.

Naver AI Research, Seoul, South Korea (remote) — May 2022 - Aug 2022
Research Intern with Kyung-Min Kim (Naver CLOVA), Joseph J Lim — June 2021 – Dec 2021
→ Reinforcement Learning in **recommender systems** for large and varying action spaces, like streaming news recommendations, using graph attention networks to address varying listwise slate-actions.

Samsung Research, Seoul, South Korea — Sep 2016 – June 2018
Engineer at Data Analytics Lab with James Geraci, Yunsu Lee
→ Market share prediction and data unification for various Samsung appliances using time series models.
→ Failure prediction with time-series anomaly detection and information extraction from text modeling.

Samsung Electronics, Suwon, South Korea — May 2015 – July 2015
Software Engineer Intern with Sungmok Seo
→ Design of smart thermostat algorithm that learns and adapts to user schedules.

Patent

J. Lee, M. Kim, **A. Jain**, T. Hwang, J. Kim, H. Cho. "Method and Apparatus for Managing Operation Data of Appliance for Failure Prediction". U.S. Patent No. 11,182,235. 11/23/2021.

Teaching & Reviewer Services

Teaching Assistant, University of Southern California (USC)
Course: Deep Learning and its Applications, CSCI-566

Joseph J. Lim: Fall 2019, Sprint 2019, Fall 2020

Jesse Thomason: Spring 2023

Yue Zhao: Spring 2024

Yan Liu: Fall 2024

Reviewer Services:

International Conference on Learning Representations (ICLR): 2023, 2024, 2025, 2026

Conference on Neural Information Processing Systems (NeurIPS): 2023, 2024, 2025

International Conference on Machine Learning (ICML): 2025

Conference on Robot Learning (CoRL): 2021, 2022, 2023, 2024

Reinforcement Learning Conference (RLC): 2025

Association for the Advancement of Artificial Intelligence (AAAI): 2026

Achievements

Outstanding Paper Award on Empirical Reinforcement Learning Research at Reinforcement Learning Conference (2025)

Director's Award for ranking in the top 7% at IIT Delhi (2013-14).

All India Rank of 198 in IITJEE (2012).

All India Rank of 91 in AIEEE (2012).

Central Board of Secondary Education Merit Certificate for securing top 0.1% (2012).

KVPY Fellowship Award and National Talent Search Examination (NTSE) Scholarship Award.

All India Rank of 3 in the National Science Talent Search Examination (NSTSE) (2011).

Top 1% nationally in Physics, Chemistry, and Astronomy (2011).

Top 30 in the Indian National Astronomy Olympiad (2012).