Anikait Singh

✉ anikait@stanford.edu | 🌐 asap7772.github.io | ⓞ asap7772 | 🐦 @Anikait_Singh_ | 🎓 Scholar

Research Interests

My research focuses on advancing decision-making methodologies, particularly in reinforcement learning, and scaling these approaches to broader applications. I am especially interested in multi-step reasoning in complex domains such as code and mathematics, as well as designing alignment algorithms and personalization techniques. My ultimate goal is to develop foundational models for decision-making that leverage diverse, large-scale datasets to achieve robust generalization, efficient rapid learning, and adaptability to individual needs.

Education

Stanford University, Palo Alto, CA   Sep 2023 – Present
Ph.D. in Computer Science   Advisors: Prof. Chelsea Finn and Aviral Kumar

University of California, Berkeley, Berkeley, CA   Aug 2019 – May 2023
B.A. in Computer Science (Highest Distinction)   GPA: 3.986; Technical GPA: 4.0
Selected coursework: Deep Learning, Reinforcement Learning, Convex Optimization, NLP, and Info Theory.

Experience

Microsoft Research, New York, NY   Jun 2025 – Sep 2025
Research Scientist Intern.

Toyota Research Institute (TRI), Los Altos, CA   Jun 2024 – Sep 2024
Research Scientist Intern.

Google DeepMind Robotics, Mountain View, CA   Apr 2023 – Feb 2024
Student Researcher.

X, the moonshot factory, Mountain View, CA   Dec 2022 – Apr 2023
AI Resident.

Selected Publications & Preprints

(* equal contribution; selection)

◇ **Learning to Discover Abstractions for LLM Reasoning**. Y. Qu*, **A. Singh***, Y. Lee*, A. Setlur, R. Salakhutdinov, C. Finn, A. Kumar. Spotlight at ES-FoMo Workshop, ICML 2025.

◇ **Few-Shot Preference Optimization of Synthetic Preferences Elicits Effective Personalization to Real Users**. **A. Singh***, S. Hsu*, K. Hsu, E. Mitchell, T. Hashimoto, A. Sharma, C. Finn. ACL 2025.

◇ **Cognitive Behaviors that Enable Self-Improving Reasoners**. K. Gandhi, A. Chakravarthy, **A. Singh**, N. Lile, N. D. Goodman. CoLM 2025 (Oral).

◇ **Improving Test-Time Search in LLMs with Backtracking Against In-Context Value Verifiers**. **A. Singh**, K. Arora, S. Keh, T. Hashimoto, C. Finn, A. Kumar. ICLR 2025 Workshop (Best Paper Runner-Up).

◇ **Big-Math: A Large-Scale, High-Quality Math Dataset for RL in LLMs**. A. Albalak, D. Phung, N. Lile, R. Rafailov, K. Gandhi, L. Castricato, **A. Singh**, *et al.* CoLM 2025.

◇ **Personalized Preference Fine-tuning of Diffusion Models**. M. Dang*, **A. Singh***, L. Zhou, S. Ermon, J. Song. CVPR 2025.

◇ **Test-Time Alignment via Hypothesis Reweighting**. Y. Lee, J. Williams, H. Marklund, A. Sharma, E. Mitchell, **A. Singh**, C. Finn. Preprint, 2024.

◇ **Towards System 2 Reasoning in LLMs: Learning How to Think With Meta Chain-of-Thought**. V. Xiang, C. Snell, K. Gandhi, A. Albalak, **A. Singh**, C. Blagden, D. Phung, R. Rafailov, N. Lile, D. Mahan, L. Castricato, J.-P. Franken, N. Haber, C. Finn. Position Paper, 2024.

◇ **Adaptive Inference-Time Compute: Large Language Models Can Predict If They Can Do Better, Even Mid-Generation**. R. Manvi, **A. Singh**, S. Ermon. Preprint, 2024.

◇ **Preference Fine-Tuning of LLMs Should Leverage Suboptimal, On-Policy Data**. **A. Singh***, F. Tajwar*, A. Sharma, R. Rafailov, J. Schneider, T. Xie, S. Ermon, C. Finn, A. Kumar. ICML 2024.

◇ **D5RL: Diverse Datasets for Data-Driven Deep Reinforcement Learning**. R. Rafailov*, K. B. Hatch*, **A. Singh**, A. Kumar, L. Smith, I. Kostrikov, P. Estruch, P. J. Ball, J. Wu, S. Levine, C. Finn. RLC 2024.

◇ **Robotic Offline RL from Internet Videos via Value-Function Pre-Training**. C. Bhateja*, D. Guo*, D. Ghosh*, **A. Singh**, M. Tomar, Q. Vuong, Y. Chebotar, S. Levine, A. Kumar. ICRA 2024.

◇ **Open X-Embodiment: Robotic Learning Datasets and RT-X Models**. Open X-Embodiment Collaboration. CoRL 2024.

◇ **Offline RL With Realistic Datasets: Heteroskedasticity and Support Constraints**. **A. Singh***, A. Kumar*, Q. Vuong, Y. Chebotar, S. Levine. NeurIPS 2023.

◇ **Cal-QL: Calibrated Offline RL Pre-Training for Efficient Online Fine-Tuning**. M. Nakamoto*, Y. Zhai*, **A. Singh**, M. S. Mark, Y. Ma, C. Finn, A. Kumar, S. Levine. NeurIPS 2023.

◇ **Pre-Training for Robots: Offline RL Enables Learning from a Handful of Trials**. A. Kumar*, **A. Singh***, F. Ebert*, M. Nakamoto, Y. Yang, C. Finn, S. Levine. RSS 2023.

◇ **RT-2: Vision-Language-Action Models Transfer Web Knowledge to Robotic Control**. Google DeepMind Robotics. ICRA 2023.

◇ **When Should We Prefer Offline Reinforcement Learning Over Behavioral Cloning?**. A. Kumar, J. Hong, **A. Singh**, S. Levine. ICLR 2022.

◇ **A Workflow for Offline Model-Free Robotic Reinforcement Learning**. A. Kumar*, **A. Singh***, S. Tian, C. Finn, S. Levine. CoRL 2021 (Oral).

◇ **A Mobile Application for Keyword Search in Real-World Scenes**. S. Pundlik, **A. Singh**, G. Baghel, V. Baliutaviciute, G. Luo. *IEEE J. Transl. Eng. Health Med.*, 2019.

Teaching & Mentorship

Graduate Student Instructor, CS224r (Spring 2025).
Program Coordinator & Mentor, Deep Learning Portal (2024).
Undergraduate Student Instructor, CS285 (Fall 2021, Fall 2022)
Undergraduate Student Instructor, CS188 (Spring 2022).

Honors

NSF Graduate Research Fellowship (2023–2028).
CRA Outstanding Undergraduate Researcher Award Finalist (2022).
Dean's List, UC Berkeley (2019–2023).

Invited Talks and Service

Invited Talks at BeNeRL, Salesforce Future AI Forum, and ServiceNow
Reviewer (w/honors) for ICLR, NeurIPS, ICML, CoRL, ICRA RAL, AAAI, and CoLM.