Vansh Kapoor

+1 412-606-5875 • vanshk@cs.cmu.edu • vansh28kapoor.github.io
vansh28kapoor

Education

Carnegie Mellon University (CMU) Pittsburgh, PA
Master of Science in Machine Learning | *GPA:* ***4.18/4*** Dec'25

- Key Courses: Deep RL (A+), Generative AI (A+), Statistics (A+), ML Optimization (A+), Adv ML (A+)
- Teaching Assistant for *Deep Reinforcement Learning*; delivered recitations, and mentored 8+ team projects

Indian Institute of Technology Bombay (IITB) Mumbai, India
Bachelor of Technology in Electrical Engineering with Honors May'24
GPA: ***9.76/10*** *(Department Rank* $\mathbf{6^{th}}$ *amongst 200+ students)* *Honors GPA:* ***10/10***

- Received *Undergraduate Research Award* for outstanding research conducted in Partially Observable MDPs
- Awarded AP Grade (*Course Topper* among 200+) in *Reinforcement Learning*, Adv Deep Learning, Adv ML

Publications

- V. Kapoor,..., A. Kumar. "*i*WAM: Action Prediction without Future Generation" *ICLR'27 Submission*
- V. Kapoor,..., A. Kumar."TRIM: Hybrid Inference in Reasoning Tasks." *ICLR'26* (Website)
- V. Kapoor, J. Nair. "MDPs with State Sensing Costs." *AISTATS'26* (Paper)

Industry & Research Experience

CMU CS Research Assistant | *Guides: Prof. Aviral Kumar & Prof. Max Simchowitz* Jan'26 - Dec'26

- Developed a **flow-map-based World Action Model (WAM)** for single-step video and action generation, achieving **over 6× faster inference** with no loss in performance compared to IDM-style WAMs on real-world robotic tasks.

Amazon | *Applied Science Intern* (ICLR 2026) | San Francisco, CA May'25 - Aug'25

- Developed **step-level LLM routing framework** for reasoning tasks (math, code generation), selectively routing critical steps to stronger models and delegating routine continuations to smaller models for budget-aware inference.
- Demonstrated **5x higher cost efficiency** than SOTA routers on AIME, with **uncertainty-aware RL and POMDP** methods matching the performance of a large expert model while using up to **80% fewer expensive tokens**.

Google DeepMind | *Research Collaborator* | Bangalore, India Jul'23 - Aug'24

- Designed look-ahead algorithms for **early rumor detection** on social graphs and adapted them for large-scale networks to nullify **bot attacks** and **adversarial misinformation propagation** on content platforms like Youtube.

IIT Bombay | *Research Assistant* (AISTATS 2026) | Mumbai, India Jul'23 - Jan'25

- Designed the novel **Selective Policy Improvement (SPI)** algorithm to compute optimal policies for **MDPs with state sensing costs**, outperforming state-of-the-art Partially Observable MDP solvers across all sensing costs.

Google | *Silicon Engineering Intern* | Bangalore, India May'23 - Jul'23

Selected ML Projects

Tetris RL Agent | ***#1 on CMU RL Course Leaderboard*** Feb'25 - Apr'25

- Benchmarked model-free and model-based RL algorithms spanning DQN, **Soft Actor-Critic (SAC)**, Twin Delayed DDPG (TD3), Evolution Strategies (ES), and **Model Predictive Control (MPC)** with beam-search for Tetris.
- Achieved a leaderboard-best **5.3M average reward** (over **2,000× the DQN baseline**) using beam-search MPC, value function trained via ES, and action-space engineering for efficient long-horizon decison-making.

Imitation Learning from Denoising Diffusion Models Sep'24 - Nov'24

- Developed **transformer-based diffusion policies** for imitation learning in BipedalWalker OpenAI Gym environment, achieving **98% of expert performance** and benchmarking against algorithms like DAgger and Behavior Cloning

Text-to-Image Diffusion Models with Enhanced Semantic Understanding Jan'24 - May'24

- Devised SUR (Semantic Understanding & Reasoning) architecture utilizing **LLAMA-based prompt enrichment**
- Boosted **multi-modal** visual question answering accuracy (counting/color/action) by **20%** over baseline CLIP

Skills

Programming Languages: Python (Proficient), C++, Julia, Java, MATLAB, HTML
Libraries: PyTorch, JAX, TensorFlow, Gymnasium, Hugging Face, PyTorch-Geometric, SciPy, OpenCV