Kamalesh Kumar

413-472-6246 kamalvidya.2002@gmail.com kkk2002 kkamaleshkumar kkamaleshkumar.github.io

Education

University of Massachusetts Amherst **Sep. 2024 – May 2026**
Master of Science in Computer Science (GPA 4.00/4.00) *Amherst, MA*

Indian Institute of Technology (IIT) Madras **Jul. 2020 – May 2024**
B.Tech in Civil Engineering, Minor in Artificial Intelligence & Machine Learning *Chennai, India*

Publications

- Breaking Free from Hand-Crafted Rewards: A Genetic Programming Framework for End-Goal-Driven **Reinforcement Learning**
- SketchCleanGAN: A generative network for improving 3D CAD model retrieval systems (*Computer & Graphics'24*) [DOI]
- SketchCADGAN: A generative approach for query sketch completion of 3D CAD models. (*Computer & Graphics'23*) [DOI]

Technical Skills

Languages: Python, Rust, C++, MATLAB, LaTeX, C, SQL
Libraries: PyTorch, TensorFlow, JAX, vLLM, triton, transformers, VeRL, DeepSpeed, TensorRT-LLM, OpenCV
Technologies/Frameworks: LangChain, ROS, Ray, Linux, Git, Spark, Hadoop, Kubernetes, Docker, Azure

Experience

Machine Learning Intern **May 2025 – Aug. 2025**
KLA Corporation *Milpitas, California*

- Mitigated catastrophic forgetting by ∼99%, retaining ≥97% of baseline defect-count performance on previously seen wafer processes during sequential fine-tuning, compared to >250% degradation under vanilla fine-tuning.
- Reduced raw defect count by ∼98% on previously seen wafers (∼85K → ∼1.5K–4K) while preserving performance on new wafers using interference-aware replay and gradient-constrained optimization.
- Cut fine-tuning time by 50% (7 → 3.5 min) by freezing ∼90% of model parameters, identifying variation-sensitive layers via Fisher Information Matrix over 443K parameters, with negligible loss in accuracy across old and new wafers.

Research Intern (RL) **May 2024 – Aug. 2024**
Mitacs Globalink *Antigonish, Canada*

- Developed a genetic programming framework that improved agent fitness scores by up to 218% in high-dimensional MuJoCo environments (e.g., Humanoid, HalfCheetah) compared to standard human-engineered reward functions.
- Optimized PPO's learning efficiency, enabling agents to reach peak performance in ∼200,000 time-steps versus the >800,000 required by the baseline, effectively reducing training time by 80%.
- Validated the statistical significance of the results ($p < 10^{-4}$) across the 11 tested environments, and surpassed all competing baselines in 82% of tasks, and achieved 22× gains in task alignment coefficient (TAC) over the PPO baseline.

Research Intern (RL) **May 2023 – Aug. 2023**
Paris AI Research Institute *Paris, France*

- Developed a unified theoretical framework for adversarial RL, formalizing and proving equivalences between state, action, and policy perturbation models under Wasserstein uncertainty (W_∞).
- Introduced and analyzed non-stationary adversarial MDP formulations, establishing conditions for optimal robust policies via minimax value functions and optimal transport-based arguments.

Research

Query-Conditioned Agentic LLM Workflow Optimization via RL **Dec 2025 – present**
Embodied AGI Lab (Advisor: Prof Chuang Gan) *UMass Amherst*

- Extending compile-time Pareto workflow optimization to a query-adaptive setting, formulating workflow selection (model choice, reasoning budget, structure) as a constrained/cost-aware RL problem over the accuracy–latency trade-off space.
- Developing and evaluating RL/bandit policies for dynamic per-query workflow selection, leveraging compile-time sub-agent cost models for state representation and sample-efficient policy optimization.

Foundations of Continual Reinforcement Learning via Hindsight Rationality **Sep 2025 – present**
Autonomous Learning Laboratory (Advisor: Prof Bruno Castro Da Silva) *UMass Amherst*

- Reframing RL beyond the traditional MDP paradigm by replacing expected-return maximization with deviation regret minimization as the primary evaluation principle for continually learning agents.
- Challenging the training–deployment dichotomy in standard RL by formulating learning as a perpetual adaptation process grounded in hindsight rationality, eliminating reliance on optimal atemporal artifacts.
- Bridging the theoretical groundwork between partially observable history processes (POHP), RL, and game theory, where agents are evaluated by adaptive behavioral rationality instead of stationary value optimality.

Improving Sketch Queries for Robust Retrieval of 3D CAD Models **Aug 2022 – Dec 2023**
Advanced Geometric Computing Lab (Advisor: Prof Ramanthan) *IIT Madras*

- Ideated a conditional Wasserstein GAN with encoder–decoder generators and PatchGAN-style discriminators, trained on paired sketch data to perform stroke denoising and sketch-to-geometry translation under severe input noise and sparsity.
- Optimized training with multi-term objectives (adversarial + pixel-wise reconstruction + structural/perceptual constraints), and validated improvements using geometry reconstruction metrics against GAN and non-GAN baselines.