Dhruv Tarsadiya

dhruvtarsadiya15@gmail.com | 213-812-4945 | dhruvtr.github.io | github.com/dhruvtr | Los Angeles, CA

Education

University of Southern California, Masters in Computer Science — Aug 2024 – May 2026 — GPA: 4.0/4.0
- **Selected Coursework:** Optimization for Machine Learning, Applied Natural Language Processing, Privacy Preserving Machine Learning, Computer Graphics

Indian Institute of Technology Guwahati, Bachelors in Computer Science — Nov 2020 – May 2024 — GPA: 9/10
- **Selected Coursework:** Databases, Compilers, Networks, Operating Systems

Publications

f-INE: A Hypothesis Testing Framework for Estimating Influence under Training Randomness | *Submitted to **ICLR** 2026* — **[paper]**

Sample, Align, Synthesize: Graph-Based Response Synthesis with ConGrs | *Accepted at **NuerIPS** ER 2025 and SCALR @ **COLM** 2025* — **[paper]**

Experience

USC NLP Group | Research Intern — May 2025 – Aug 2025
Advised by: **Sai Praneeth Karimireddy**
- Devised data influence estimation method based on differential privacy and robust to model training randomness
- Scaled method to fine-tune 8 billion parameter LLMs and outperformed existing SOTA methods in detecting data poisoning and attributing model behavior by up to 20% and improving reliability of scores by up to 60%

DILL Lab, USC | Research Assistant — Jan 2025 – present
Advised by: **Swabha Swayamdipta**
- Designed a graph-based data structure to encode semantic variations in LLM responses, thereby proposing a novel inference time scaling method, improving factuality by 31% and mathematical reasoning capability by up-to 6% over baselines for frontier 70 billion parameter Llama and Qwen models

Hanyang University, South Korea | Software Intern — May 2023 – July 2023
- Implemented a new lower bound interval type-2 (IT2) fuzzy clustering algorithm in the Fuzzy MATLAB Toolkit

Projects

Diverse Code Generation with Large Language Models [report][code] - Utilized Quality Diversity and Reinforcement Learning based optimization techniques for more diverse semantic and syntactic code sampling using LLMs, leading to 10% improvement in generated code accuracy over baselines

Ray Tracer [code] - Wrote a custom multi-threaded ray-tracer from scratch in C++, additionally handling recursive and spectral reflection and soft shadows. Enhanced anti-aliasing through gaussian super-sampling

Majuli Island Virtual Reality Tour [code] - Led, coordinated and managed a team to develop an Interactive VR Tour of Majuli Island in Unity, leveraging over 500 360-degree images and 20 360-degree videos and deployed on Meta Quest 2

xv6: Unix-based Operating System [code] - Implemented a Shortest-job-first and Round-robin based hybrid CPU scheduling algorithm, paging, lazy memory allocation and dynamic page swapping using C language

Skills

Programming Languages and Software Tools : Python, C, C++, C#, Javascript, MATLAB, Git, Unity, Verilog, Docker
Libraries and Packages : PyTorch, HuggingFace, Transformers, OpenGL, OpenMP, vLLM, Ollama, Scikit-Learn, Django

Awards

- Secured national rank 370 (top 0.1 percentile) in Joint Entrance Examination, 1 million+ candidates (2020)
- Achieved national rank 97 in KVPY-SA and received national research fellowship scholarship (2019)