Arjun Sharma

arjun.sharma@caltech.edu | github.com/ArjunS07 | linkedin.com/in/ArjunS07 | arjuns07.github.io

Education

California Institute of Technology *(GPA: 4.2/4.0)* Pasadena, CA

B.S. in Computer Science; minor in Information & Data Science *Graduation: June 2028*

- Coursework *(* denotes graduate-level, achieved A+/A in all listed classes)*
  - Deep Learning*, Machine Learning & Data Mining*, Computer Graphics*, Computing Systems, 3-D Deep Learning*, Data Structures & Algorithms, Software Design in C
  - Linear Algebra (proof-based), Applied Linear Algebra*, Probability, Differential Equations, Multivariable Calculus
  - *Scheduled before June 2026: GPU Programming*, Large Language & Vision Models*, Computational Cameras**
- Activities: Localization & mapping @ Caltech Lunabotics, Venerable House Webmaster, Board of Control member

Experience

Caltech Yisong Yue Group, Research Assistant Jan. 2026 – Present

- Researching reinforcement learning post-training of large language models for complex codebase-level agentic tasks
- Converting code benchmarks to open-source environments for RL with continuous & verifiable rewards

Caltech CMS Group (CERN collaboration) @ Fermilab, Research Assistant June – Dec. 2025

- Designed novel pipeline with physics-informed flow matching / diffusion model to generate point clouds in 4-dimensional space, guided by global attributes generated by a normalizing flow (state-of-the-art log-likelihood)
- Trained equivariant message-passing graph attention network for flow matching; implementing residual architecture, classifier-free guidance, coordinate transformations, and variable-sized point cloud generation

Skills: Generative modeling, flow & diffusion models, PyTorch, JAX, NumPy, Python, Kubernetes

Cambridge University, ML + Radio Astronomy Student Researcher June 2023 – July 2024

- Identified novel method (positive unlabeled learning) for flagging anomalies in fast radio burst data; improved recall by 4.62% with classifier ensemble; submitted 66 future observation candidates
- Published first-author paper in the Monthly Notices of the Royal Astronomical Society

Skills: Semi-supervised learning, Python, scikit-learn, SciPy, Pandas, NumPy

BAM.money, Full-stack Software Engineering Intern June – Dec. 2022

- Developed 4 new features for a Python webscraper to extract insights from 200k+ news articles/day in MongoDB; deployed to cronjob on AWS ECR; designed back-end and front-end pipeline to display insights to clients

Skills: Node.JS, Python, ReactJS, Keras, scikit-learn, Pandas, Bash, AWS (EC2, ECR, S3), Docker

Teaching

Teaching Assistant, Caltech CS 2 (Data Structures & Algorithms) Jan. 2026 – Present

- Host 6 hours of 1-on-1 interaction for debugging and review every week; perform quality assurance for assignments

Awards

Caltech Physics 11 Awardee: 1 of 6 students selected for class based on solutions to open-ended modeling problems; awarded full funding ($8000 grant) to pursue summer research

Winner, Govt. of India Smart India Hackathon 2022: 1 of 50 winners from 3000+ submissions; 1 of 5 winners in the education category. Built full-stack platform with Django, Flutter, & AWS to enable secure and scalable tutoring

Projects

Complete 3-D graphics rendering pipeline | *C++, OpenGL, GLSL*

- Developed renderer from scratch in pure C++ with perspective projection, Phong shading, attenuation, reflectance, backface culling, depth buffering, line rasterization, arcball rotations. Re-implemented with OpenGL and GLSL
- Wrote a ray tracer using superquadrics, implementing lighting models for shadowing, reflections, and refraction
- Implemented a mesh processor for implicit fairing using the discrete Laplacian and halfedges

Full-stack platform to host puzzle hunts | *Python, Django, TailwindCSS, PostgreSQL* GitHub

- Architected & implemented a platform to host collaborative puzzle-solving events; served 300+ concurrent users for 72+ hours for 3 consecutive annual events; managed Django + Postgres deployments to Google Cloud & AWS

Skills

Languages: Python, C, C++, Java, Golang, MATLAB, JavaScript, TypeScript, Dart, Swift, SQL, HTML/CSS, Bash

Technologies: PyTorch, Tensorflow, JAX, sklearn, OpenGL, Django, Flask, TailwindCSS, CUDA, Flutter, AWS, Docker