

KHUSH SHAH

Carnegie Mellon University | B.S. in Computer Science & Mathematics | Expected May 2029
(650) 404-7748 | khushs@cmu.edu | khush.site | [linkedin.com/in/khush--shah](https://www.linkedin.com/in/khush--shah) | github.com/khushshah1
Fully work authorized | U.S. Citizen

EDUCATION

Carnegie Mellon University | B.S. in Computer Science & Mathematics | Expected May 2029

Relevant Classes: Machine Learning, Data Structures & Algorithms, Competition Programming, Computer Systems, Linear Algebra, Multivariable Calculus, Differential Equations

Focus Areas: Machine Learning, Artificial Intelligence, Backend Systems, Data Structures & Algorithms

Los Altos High School

TECHNICAL SKILLS

Languages: Python, Java, C++, C, C#, JavaScript, TypeScript, Rust

Tools & Frameworks: PyTorch, NumPy, Pandas, React, Node.js, Flask, Express, Postgres, ORM, Unity, Git/GitHub

ML: Implemented transformers from scratch; completed ARENA (technical AI Safety curriculum covering RL, evals, interpretability, and alignment); working knowledge of LLM internals, attention, RLHF, and alignment techniques

PROJECTS & RESEARCH

Carnegie AI Safety Initiative (CASI) (2025 – Present)

- Participated in a technical AI safety research reading group, studying fundamentals of ML papers and alignment techniques (transformers in depth, mechanistic interpretability, RLHF, etc.).
- Co-leading a 3-person research project on AI safety alignment robustness, producing empirical research.
- *Geometric Framework for Predicting Fragility of LLM Safety Training Methods* Research Paper:
 - Use LAT-extracted RepE vectors and Hessian eigenvector analysis to show safety concepts occupy lower-rank, higher-curvature subspaces than capability directions
 - Show benign fine-tuning preferentially aligns with fragile safety directions, with representational rank and layer depth as independent fragility predictors
 - Build predictive framework ranking DPO vs PPO-RLHF alignment robustness by geometric properties, validated on HarmBench across LLaMA, Mistral, and Qwen

ScottyLabs Tech Team – CMU Courses Backend (2025 – Present)

- Built automated Rust system to query CMU APIs and recursively discover related course resources.
- Wrote Python ETL to parse/normalize data and maintain Postgres schema + records.
- Exposed processed datasets to the frontend through structured JSON and CSV endpoints.
- Rebuilt the web app in TypeScript + React using TanStack Router; platform used by 100s of CMU students.

Stock Price Forecasting Model (Team Lead) – UC Davis COSMOS (July 2024 - August 2024)

- Implemented ml models from scratch in Python using only NumPy, deriving forward/backprop manually from calculus and linear algebra primitives
- Led 5-person team supervised-fine-tuning sentiment-analysis models for stock-price forecasting; outperformed baselines and presented results at COSMOS Showcase

LEADERSHIP EXPERIENCE

Los Altos Hacks: Organizer & Sponsorship Lead (2021 – 2025)

- Organized the world's largest high-school hackathon. Raised \$15K+ in sponsorships.
- Gave speeches to 1000+ students and mentored many new members

Computer Engineers of the Next Generation (CENG): Senior VP (2022 – 2025)

- Managed 100+ volunteers teaching 24+ coding classes (Python, Java, JS) to underserved schools. Earned *Presidential Volunteer Service Award* for 200+ service hours in a single year.