

Pranav Varshney

US Citizen | pvarshh.me | pvarsh@umich.edu | linkedin.com/in/pvarshh | github.com/pvarshh

EDUCATION

University of Michigan

B.S. Computer Science

Dec 2026

Ann Arbor, MI

EXPERIENCE

Uber

Software Engineering Intern (Winter) | Rideshare Optimizations

Jan 2026 - Apr 2026

San Francisco, CA

Scale AI

Generative AI Intern | Eval Pipelines

July 2025 - Oct 2025

Remote

- Developed a retrieve and rerank RAG pipeline to automate LLM inputs for data annotation via candidate filtering
- Built a high-throughput benchmark system to score 8K+ model responses daily, reducing 60% annotation errors
- Created an LLM-as-a-judge system alongside adjudication measures to improve data labeling correctness to 96%

Amazon Web Services

Software Engineering Intern | Sagemaker

May 2025 - July 2025

Bellevue, WA

- Created a Go-based CLI for mapping 1M+ SageMaker endpoints to adjacent AWS Services for troubleshooting
- Developed a real-time system health monitoring framework, yielding an on-call response time reduction of 70%
- Built containerization tooling for audit systems through AWS copilot, enabling support over 10,000 queries/wk

University of Michigan

Instructional Assistant | Python Programming

Aug 2024 - Dec 2024

Pontiac, MI

United Wholesale Mortgage

Software Engineering Intern

May 2024 - Aug 2024

Pontiac, MI

- Normalized databases using 3NF, reducing duplicates by 45% to accelerate GCP migration timeline by 3 weeks
- Saved \$20K by implementing a random forest model to classify loan outcomes, accepting or rejecting the proposal

Ratna Global Technologies

Software Engineering Intern

May 2023 - Aug 2023

Newark, CA

- Developed a RAG based customer support system using company data to resolve 80% inquiries with 90% accuracy
- Retained \$1.2K/month in API request costs by implementing a redis cache in react, reducing redundant API calls

RESEARCH

Precision Imaging × AI Laboratory | [Prof. Joyce Wang](#)

Oct 2025 - Present

- Engineered models for automated segmentation of cardiac chambers and quantification of ejection fraction
- Building a standardize CVD diagnosis platform, targeting interoperability with portable ultrasound devices

Self Operating Networks Laboratory | [Prof. Victor Liu](#)

Jan 2025 - May 2025

- Designed AST-aware transformer using constrained attention mechanisms, achieving 92% code generation accuracy
- Designed contrastive learning pipeline for automated dataset expansion, increasing training samples by 250%

PROJECTS

Cvrve

- Sold at 7 figures; helped develop the infrastructure observability behind [job board](#) (prev. cvrve) used by 100k+

Sharded Key-Value Store | Go, Paxos, Concurrency

- Built a fault-tolerant, sharded key-value store using Paxos for consensus and coordination between replica groups
- Implemented dynamic reconfiguration to safely transfer shards between replica groups during changes in topology

SKILLS

Languages: Python, C++, Go, Java, C#, Typescript, Javascript, SQL

Technologies: Docker, Flask, Next, React, Node, Express, Tailwind CSS, Pandas, NumPy, Plotly, Sci-kit Learn

Concepts: Full-Stack Development, Mobile Development, Cloud Computing, Parallel Programming, Machine Learning, NLP, Computer Vision, Data Analytics, Big Data, Embedded Systems, Quantum Computing, Computer Security, Linux