Rohan Siva

669-294-7436 | rohansiva@utexas.edu | linkedin.com/in/rohan-siva | | rohan-siva.github.io | Google Scholar

Education

University of Texas at Austin Austin, TX
B.S Electrical & Computer Engineering Honors, DS minor (GPA: 3.97) Expected Graduation: May 2027
Relevant Coursework: Data Structures, Data Science Principles & Lab, Probability & Statistics, Matrices, Intro to Computing, Discrete Math, Embedded Systems, Linear Systems, Linear Algebra, Diff Eq, OOP
Extracurriculars: Association for Computing Machinery (ACM) - Internal Officer, UT Convergent - AI Dev

Experience

Cisco - *Machine Learning Engineering Intern* May 2025 – Nov 2025
- Developed kRAIG, an AI agent converting natural language into executable Kubeflow Pipelines for ETL workflows
- Deployed a RAG-based pipeline generator using **Elasticsearch** and safety guardrails for database operations
- Integrated **MCP** Server w/ **PostgreSQL & AWS** tooling, automating direct **ETL** into enterprise data storages
- Implemented **CI/CD** pipeline for **Docker** container deployment & automated unit tests w/ 90%+ code coverage

Center for Autonomy @ UT Austin - *Undergraduate Researcher* Aug 2024 – Present
- Refined autonomous driving pipeline in **CARLA** sim using **conformal prediction** for uncertainty scoring
- Engineered **LLaVA** finetuning pipeline w/ **PyTorch**, gaining **5%** higher performance & **40%** lower variance
- Designed intervention framework via formal methods and active sensing, w/ **95% traffic safety compliance**

PKU Lab @ Peking University - *Undergraduate Researcher* May 2025 – Present
- Conducting embodied AI research under Prof. Hao Tang, optimizing reasoning for vision-language models (**VLM**)
- Created novel parallelizable diffusion-based **PyTorch** architecture for VLM reasoning, **cutting latency by 34%**

Canyon Technologies LLC - *Software Engineering Backend Intern* May 2024 – Oct 2024
- Built a cloud-based nzESL display management dashboard w/ **SpringBoot**, **Docker** & **Postman** for 1000+ users
- Integrated display tags and BLOZI base stations to enable real-time updates, QR-codes, and label management
- Designed and deployed a secure customer mailing & password system with **Java** and **React.js** using **OWASP**

Keitt Lab @ UT Austin - *Undergraduate Researcher* Nov 2023 – July 2024
- Built low-power sensor mesh using **XBee** radios & time-synced **Raspberry Pis** for remote wildlife monitoring
- Developed **PyTorch** acoustic detection system to identify bird species & location from chirp data w/ live analysis
- Implemented **distributed computing** architecture for remote data collection, **reducing latency by 46**%

Selected Publications

Bhatt, N. P., Yang, Y., **Siva, R.**, Milan, D., Topcu, U., & Wang, Z. (2024). **Know Where You're Uncertain When Planning with Multimodal Foundation Models: A Formal Framework**.Accepted @ **MLSys 2025**. [**arXiv**] [**Code**]

Projects

MoodScribe – Social Sidekick | *GitHub*
- Supervised fine-tuned Llama 3.1 model for sentiment analysis with a synthetically generated dataset using LoRA adapters, optimizing training with Unsloth for memory reduction and FlashAttention for a 44% speedup.

Medicina.ai – Health Risk Predictor | *GitHub*
- Developed a full-stack health-risk prediction app with Python, HTML/CSS/JS & Flask, using Random Forest & Logistic Regression to estimate diabetes & heart disease risk from user input, including preprocessing & inference.

Blackjack-RL – Deep RL Blackjack Agent | *GitHub*
- Implemented a DQN-based RL agent in PyTorch using RLCard with experience replay, ϵ-greedy exploration, target networks, and reward shaping to learn optimal strategies in Blackjack environments.

Technical Skills

Languages: Python, Java, C, C++, JavaScript, SQL, TypeScript, Swift, HTML, CSS, R, Assembly
Technologies: SpringBoot, Flask, FastAPI, Node.js, PostgreSQL, Docker, Linux, Kubernetes, React.js, CI/CD, AWS
AI/ML: PyTorch, TensorFlow, JAX, OpenCV, Carla, Elasticsearch, LangChain, Spark, Transformers, Unsloth, MCP
Concepts: Full-Stack, Databases, Deep Learning, NLP, ML, CV, Autonomy, Perception/Planning, SLAM, Sys Design