

Tharun Kumar Tiruppali Kalidoss

+1 (310) 961 6434

tharunk@princeton.edu

Senior at Princeton University

tharunkumar.xyz

Education

Princeton University
B.S.E. Computer Science
Aug 2022 - May 2026

Relevant Courses: Theory of Algorithms w/ Dr. Robert Tarjan, Distributed Systems, Real Analysis, Algebra, Stochastic Systems, Reinforcement Learning, Economics in Computing, Intro ML, Data Structures and Algorithms, Discrete Math, Linear Algebra.

Technical Skills

C++, Python, Typescript, Javascript, Java, SQL, NoSQL, Pytorch, JAX, TensorFlow, Azure, AWS, Mobile Development, Full-stack Development, Swift, Machine Learning, CI/CD, Kubernetes

Relevant Experience

Research Engineer — Google DeepMind, Princeton, NJ

Sep 2025 – Present

- Researched and implemented novel State Space Model architectures for efficient long-context language modeling **under Prof. Elad Hazan** in the Princeton Google DeepMind Lab.
- Trained 500M, 1B, and 7B parameter LLMs from scratch, benchmarking performance against transformer and SSM baselines, reaching SOTA for long-context retrieval, training stability. Planning to publish to ICML.

Software Engineering Intern — Amazon Web Services, Seattle, WA

Jun 2025 – Sep 2025

- Built low-level Python and Go bindings to monitor server hardware health and automate management of a **50,000+ node fleet**, ensuring regional availability by maintaining host redundancy while repairing failed nodes.
- Increased server availability by 15%** and **eliminated 30% of on-call tickets** after shipping to production with CI/CD pipelines
- Received top performance evaluation for contributions to large-scale infrastructure automation.
- Led organization-wide GenAI talks to train teams on leveraging generative AI tools for productivity optimization.

Member of Technical Staff — UniversalAGI, San Francisco, CA

Jan 2025 – Apr 2025

- Built AI solutions for Fortune 500 companies and governments. Flying back/forth between Princeton, SF, and Dubai.
- Developed a **state-of-the-art Text-to-SQL LLM agent** achieving **high accuracy on Spider 1.0 benchmark**; successfully deployed the solution **on-premise for the UAE government** w/ Kubernetes on private employee data.
- Developed an **LLM agent for Apple** that builds no-code applications w/ data integrations to external services (e.g. Slack, Gmail).

Research Engineer — Hooglee, San Francisco, CA

May 2024 – Dec 2024

- Worked with **Sebastian Thrun, and Bichen Wu** to incubate a funded consumer video generation startup.
- Engineered synthetic media generation system utilizing wav2lip fine-tuning for audio-driven animation, FLAME 3D facial model integration, and MediaPipe pose estimation, deployed on iOS with Metal acceleration.
- Developed parameter-efficient LoRA training pipeline on distributed H100 infrastructure for personalized image generation models, creating a cloud-based API pipeline for real-time image synthesis in social media applications

Software Engineering Lead — Nora Music, Contract, Remote

Aug 2023 – Mar 2024

- Developed a state-of-the-art webcrawler with Selenium with GPT data-cleaning pipeline. Fully automated in AWS EC2.
- Invented a proprietary encryption system and audio fingerprinting strategy to prevent song/audio piracy.

Projects

Multi-Modal Comic-to-Animation Engine

Mar 2023

- Built an end-to-end comic-to-anime generation engine using multi-modal agents, Stable Diffusion in-painting, and stable video diffusion with automated lip-syncing and ElevenLabs voice synthesis
- Worked with comic creators with 200k+ readers, landed two animation companies as clients, and got YC interview for this project. (Demo: https://www.youtube.com/watch?v=Hb_hQ3CVrM)

Jul 2024

RL-based Obstacle Avoidance for Drone Swarms

- Formulated obstacle avoidance problem as an MDP, and trained a CNN via PPO to teach drones in a swarm to navigate various environments while maintaining swarm formation, used OpenAI gym environment. Won Project Eagle hackathon at AGI House.

Award-winning GPT-based Arcade Wagering Platform built on Ethereum

Mar 2023

- Collaborated with small engineering team to develop wagering platform for web-based arcade games using React, Flask, Firebase, Solidity, Arbitrum. Smart contracts built on the Arbitrum test-net. (github.com/tharunkumartk/cadia)
- Won a total of \$25,000 in prize money, winning Princeton Pitch, Princeton's DeSo Hackathon, and LionHack.