

Education

- **University of Michigan**, Ann Arbor Aug. 2025 – May. 2027
Master of Science in Computer Science and Engineering
- **SRM Institute of Science and Technology**, Chennai, Tamil Nadu Aug. 2021 – Aug. 2025
Bachelor of Technology in Computer Science and Engineering
CGPA: 9.47/10, First Class with Distinction

Publications

- *Pagoda: Roofline Characterization of Energy and Time for DNN Inference and Training on Edge Accelerators*: 3 authors, **Gupta** et al., Yogesh Simmhan. Under review at IEEE TPDS. ([link to preprint](#))
- *Fulcrum: Power and Latency-aware Optimization of Concurrent DNN Training and Inference on Edge*: 2 authors, **Gupta** et al., Yogesh Simmhan. Under review at ACM TECS. ([link to preprint](#))
- *VARs: Vision-based Assessment of Risk in Security Systems*: **Gupta** et al. Accepted at ICIoT 2025.
- *ViDAS - Vision-based Danger Assessment and Scoring*: **Gupta** et al. ICVGIP 2024 ([link to paper](#))
- *ECHO - Environmental Sound Classification with Hierarchical Ontology-Guided Semi-Supervised Learning*: **Gupta** et al. IEEE CONECCT 2024 ([link to paper](#))
- *ISAApp - Image Based Attendance Application*: 4 authors, **Gupta** et al. International Conference on AAIMB, 2023 ([link to paper](#)).
- *Managing Congregations of People by Predicting Likelihood of a Person Being Infected by a Contagious Disease like the COVID Virus*: **Gupta** et al. IEEE CCEM, 2020 - **Best Short Research Paper Award** ([link to paper](#)).

Work Experience

1. **Research Intern, Stanford University**, PI - [Dr. Teddy Akiki](#) Jun. 2024 – Jun. 2025
 - Developed autoregressive and diffusion models (**Chronos**, **TimeLLM**) for time-series forecasting to uncover latent “language-like” patterns in fMRI brain activity.
 - Designed and trained a **VQ-VAE + GPT2/LLaMA2** pipeline to discretize continuous fMRI signals into latent tokens and predict future sequences.
2. **AI Intern, Kiwi** Apr. 2025 – May. 2025
 - Constructed a **LangGraph-driven multi-agent framework** handling **100+** customer service **interactions per day** across varied query types like Transactions, Rewards, and Card applications.
 - Framework included **7+ LLM agents** with **15+ tools** and Freshdesk and Mixpanel integrations delivering accurate results in **< 1 min**.
3. **Research Intern, Indian Institute of Science**, PI - [Dr. Yogesh Simmhan](#) Aug. 2024 – Feb. 2025
 - Proposed and improved baseline neural networks through a **penalizable MAPE loss function** in optimizing concurrent training and inference on edge accelerators from MAPE scores of 10% to 2%.
 - Built a **novel Active Learning technique** that samples the best power modes on Jetson devices for DNN workloads with losses close to optimal (0%).
4. **Research Intern, Samsung R&D**, PI - [Mr. Shwetank Choudhary](#) Sept. 2023 – Jul. 2024
 - Devised a **novel** two-stage pipeline that trains a CNN with semi-supervised learning(ResNet50, EfficientNet) on the label ontology of environmental sound classes as a **pretext** task.
 - Produced an accuracy **increase** of 1-8% compared to baseline models, achieving 97.5% **state-of-the-art** accuracy in the ESC-10 benchmark dataset.
5. **AI Intern, Upthrust** Jun. 2023 – Aug. 2023
 - Integrated **PageSpeed API** and **Screaming Frog API** to produce and fix website issues with **Langchain** and **LLMs**.
 - Linked the **Meta API** with **GPT** to create a chatbot providing **recommendations to optimize ads**.
6. **Data Science Intern, AarogyaAI** Jun. 2020 – Sept. 2021
 - Deployed a dashboard to display datasets and model results stored in AWS S3 bucket, showcasing what drugs are resistant to the given variant of Tuberculosis bacteria.

Positions of Responsibility

- **Odyssey Lab** Jan. 2024 – Jan. 2025
 - Started a lab to help students across the best universities in India to work on research.
 - Built a network of mentors from IISc, IIIT-H, Amazon, IBM, and Samsung and created 3 teams to work on Multimodal and LLMs reasoning projects.
 - Published a paper to ICVGIP 2024, India’s top Computer vision Conference.