Sai Paresh Karyekar

★Atlanta, GA 🜙 +1 (404) 200-3907 🗷 saikaryekar@gmail.com 🗘 github.com/saikaryekar

in linkedin.com/in/sai-karyekar

EDUCATION

Georgia Institute of Technology, Atlanta, GA

May 2025 GPA: 3.9/4

Master of Science in Electrical and Computer Engineering (ML)

Veermata Jijabai Technological Institute (VJTI), India

June 2023 CGPA: 9.62/10

Bachelor of Technology in Electronics & Telecommunication Engineering

Relevant Coursework

Generative and Geometric Deep Learning, Data and Visual Analytics, Machine Learning, Technology Entrepreneurship (Teaching Assistant)

EXPERIENCE

NVIDIA, Santa Clara | GenAl Intern

June 2024 - Aug 2024

- Developed & fine-tuned an LLM (Llama3) to power a Generative AI-based Retrieval-Augmented Generation (RAG) tool, improving blog generation efficiency.
- Deployed the application via Gradio on an AWS EC2 instance, creating an interactive frontend for AI-generated content enhancing adoption by 90%.
- Integrated LangChain & LlamaIndex to enhance document retrieval, chunking and indexing.

Georgia Institute of Technology, Atlanta | Student Researcher 🔗

Aug 2023 - Dec 2023

- Built predictive ML models using scikit-learn to analyze large-scale education datasets, improving academic performance prediction by 5%.
- Applied feature engineering & data analysis (ANOVA, OLS regression) to derive insights on student learning behaviors.

PROJECTS

Optimizing Mathematical Reasoning in SLM§

Sept 2024 - Dec 2024

- Fine-tuned a T5-small transformer model (60M parameters) with LoRA & Chain-of-Thought (CoT) prompting to optimize math reasoning.
- Benchmarked quantized T5-base models, reducing memory consumption by 5% and improving accuracy by 12%

CityScout: Statistics driven Safe City Recommendation System &

Sept 2024 - Dec 2024

- Developed a crime risk prediction model using Random Forest & TF-IDF, for region-specific safety analysis.
- Processed & analyzed **2GB+** of crime data using PySpark & Pandas on **Google Cloud.**
- Built an interactive **Tableau dashboard**, to visualize city-level crime risk scores.

ApplyCation: Automated Job Application Platform §

Oct 2024

- Built an automated job application system using Selenium & Claude API, reducing application time by 50%.
- Designed a Streamlit-based UI for user profile management, integrating Flask backend & GCP.

SKILLS

Machine Learning & AI

LLM Fine-Tuning, RAG, LangGraph, OpenAI API, Anthropic API

Backend & Cloud Deep Learning

Python, C++, Flask, AWS, GCP, Docker, PostgreSQL, Linux, React

PyTorch, TensorFlow, Hugging Face, OpenCV, Git

Extra-Curriculars

- Contributor, Workout of the Day Prediction (GT Data Science Club) ML-based fitness analysis project.
- Social Chair, Grad Society of Women Engineers (Grad SWE) Organized networking events to foster in-person connections.