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**
Master of Science in Electrical and Computer Engineering (ML)   *GPA: 3.9/4*

Veermata Jijabai Technological Institute (VJTI), India   **June 2023**
Bachelor of Technology in Electronics & Telecommunication Engineering   *CGPA: 9.62/10*

RELEVANT COURSEWORK

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

EXPERIENCE

NVIDIA | *Developer Marketing Intern (Product Marketing)*   June 2024 - Aug 2024

- Engineered a Generative AI-powered RAG tool for technical blog generation using Llama3 model on NVIDIA NIMs, accelerating content creation and demonstrating expertise in deep learning and AI/ML techniques.
- Optimized outputs through advanced prompt engineering, producing detailed technical blog outlines, comprehensive content, and effective social media posts while ensuring 90% adoption within the team.
- Curated developer resources for the contest with LlamaIndex, thereby improving support for developers.

Georgia Institute of Technology, Atlanta | *Student Researcher* 🔗   Aug 2023 - Dec 2023

- Applied statistical techniques (ANOVA, OLS regression) and machine learning models to analyze a large-scale dataset of student performance in a microelectronics course.
- Developed predictive models using scikit-learn and performed feature importance analysis, resulting in a 5% improvement in course success rates.

PROJECTS

Enhancing Mathematical Reasoning in Small Language Models (SLMs) 🔗   Sept 2024 - Dec 2024

- Optimized the T5-small transformer model (60M parameters) for math reasoning on the GSM8K dataset using techniques such as Low-Rank Adaptation (LoRA), Chain-of-Thought (CoT) prompting, and full fine-tuning.
- Compared CoT-based methods to a quantized T5-base model, demonstrating 5% lower memory usage while maintaining comparable performance.

ApplyCation: Automated Job Application Platform 🔗   Oct 2024

- Developed an automated job application system leveraging Selenium and dynamic script generation via Claude API, streamlining job applications on various portals and reducing the time per application by 50%.
- Designed a responsive Streamlit-based frontend for managing user settings, resume uploads, and job tracking, improving user experience and accessibility.

SKILLS

Languages	Python, C++, SQL, R, MATLAB
Tools	Git, Linux, GCP, AWS, Docker
Frameworks	Tensorflow, PyTorch, Scikit-learn, OpenCV, LangChain, LlamaIndex, Selenium, Streamlit

EXTRA-CURRICULARS

- As a member of the Data Science club at GT, contributed to the Workout of the Day prediction project by estimating workout durations and comparing performance with CrossFit enthusiasts.
- Served as the Social Chair of the Grad Society of Women Engineers (Grad SWE).