Sai Paresh Karyekar

Atlanta, GA (404) 200-3907 saikaryekar@gmail.com saikaryekar sai-karyekar

Education

Georgia Institute of Technology, Atlanta, GA May 2025
MS in Electrical and Computer Engineering, (Machine Learning) *GPA: 3.85/4*

University of Mumbai, India June 2023
B.Tech in Electronics & Telecommunication Engineering *CGPA: 9.62/10*

Relevant Coursework

Generative and Geometric Deep Learning, Data and Visual Analytics, Regression Analysis, Machine Learning, Digital Image Processing, Tech Entrepreneurship **(Teaching Assistant)**

Experience

NVIDIA | *Developer Marketing Intern* June 2024 - Aug 2024

- Developed a Generative AI powered internal tool for end-to-end technical blog generation using the Llama3-70b-instruct model hosted on NVIDIA NIM API Catalog.
- This tool employs Retrieval-Augmented Generation (RAG) with the FastEmbed Embedding model, using nvidia/rerank-qa-mistral-4b for reranking and FAISS for local vector embedding storage.
- Implemented advanced prompt engineering techniques to generate comprehensive blog outlines, detailed content, and corresponding social media posts from provided documents and URLs.

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

- Performed exploratory data analysis using statistical methods such as **ANOVA, OLS regression** to predict final grades of students in a course.
- Identified key factors influencing student performance through **feature importance analysis**, allowing for targeted interventions and improving overall course success rate by 10%.

Projects

Enhancing Math Reasoning in SLMs via Fine-Tuning and Compression Sept 2024 - Present

- Fine-tuning T5-small on math-specific reasoning datasets (GSM8K, MATH) using LoRA and prompt-based methods, aimed at improving reasoning performance while maintaining computational efficiency.
- Implementing quantization-aware training and knowledge distillation to create student models, optimizing for reduced inference time and memory usage without sacrificing performance in reasoning-intensive tasks

ML based Portfolio Management with Macro-Financial Indicators Jan 2024 - May 2024

- Developed an ML-based portfolio management system integrating macro-financial indicators to optimize portfolio allocation. Utilized various ML models, with DBSCAN achieving a high Silhouette score of 0.737
- Implemented K-means, ARIMA, and LSTM models to predict stock prices, resulting in improved predictions and a higher Sharpe ratio.

Deep Learning-Based Forest Fire Classification with VGG19 Jan 2023 - May 2023

- Developed a CNN based stubble fire detection system, using transfer learning on a modified VGG-19 architecture, yielding a **95.2%** accuracy in binary classification on the DeepFire dataset.

Skills

- **Programming Languages**: Python, C++, SQL, R, MATLAB, JavaScript
- **Tools**: MySQL, Tableau, PowerBI, Git, Linux, OpenRefine, D3.js, OpenGL, Azure, AWS, Docker
- **Frameworks**: Tensorflow, PyTorch, Scikit-learn, OpenCV, LangChain, LlamaIndex PySpark, Scala, CUDA

Achievements and Extra-Curriculars

- As the Social Chair of the Grad Society of Women Engineers (Grad SWE), organized events, managed RSVPs and finances, and planned activities including volunteering, arts and crafts, movie nights, and off-campus outings.
- As a member the Data Science club at GT, contributed to the Workout of the Day prediction project by estimating workout durations and comparing performance with fellow CrossFit enthusiasts at local affiliates.