

RIGVED SHIRVALKAR

+1 (718) 668-5932 | rigvedrs@gmail.com | New York, NY, USA | linkedin.com/in/rigvedrs

EDUCATION

New York University

September 2024 - May 2026

Master's, Computer Engineering

GPA: 4

- Computer Systems Architecture, Real Time Embedded Systems, Machine Learning, MLOps, Big Data, Deep Learning

Amrita School of Engineering

October 2020 - May 2024

Bachelor's, Electronics and Computer Engineering

GPA: 8.66

- Data Structures and Algorithms, Deep Learning, Machine Learning (Mathematics and Statistics), Operating Systems, DBMS (SQL), Data Science

SKILLS

Skills: Pytorch, LangChain, llama index, streamlit, Tensorflow, FastAPI, Natural Language Processing (NLP), Computer Vision, NumPy, Pandas, Jupyter, git, React.js, Django, AI, C/C++, Python, CUDA, Linux/Unix, Docker, Kubernetes, Cloud Computing

PROFESSIONAL EXPERIENCE

NYU IT

AI Developer

December 2024 - Present

- Working on GenAI solutions with the Research and Technology team at NYU IT, building custom RAG pipelines, testing and integrating with various language models for enhanced QA systems.
- Integrated token tracking, cost tracking, and other metrics to monitor the performance and cost of the models used.

Comsense Technologies

AI Intern

February 2024 - June 2024

- Built a custom Retrieval-Augmented Generation (RAG) pipeline using large language models (LLMs) and embedding model fine-tuning, achieving improved high hit rates by 22% and reliable answers compared to generic solution
- Integrated and tested with various language models for enhanced QA systems, and designed and implemented the product's backend using FastAPI and integrating LlamaIndex for better document processing.
- Collaborated with cross-functional teams to implement genetic algorithms for travel path optimization for a client.
- Built multiple proof of concepts using Generative AI and deep learning solutions for the clients
- Skills used: Python, NLP, PyTorch, FastAPI, LlamaIndex, Deep Learning, Gen AI

Red Hen Labs

Open Source Contributor

June 2023 - November 2023

- Developed a multi-model deep learning pipeline to generate captions for speech gestures in Christian Art images, utilizing vision-language models and YOLO V8 pose detection.
- Applied style transfer and fine-tuned the YOLO V8 model to adapt to art datasets, improving pose detection.
- Skills used: Vision-Language Pre-training, Style Transfer, YOLO V8, Data creation, Python, Deep Learning, GPU Clusters

Traboda

Full Stack Web Developer

March 2021 - July 2021

- Enhanced and streamlined user interface (UI) and user experience (UX) by adding new features and resolving issues on both backend and frontend.
- Utilized CSS, ReactJS, and Django to develop and maintain web applications.
- Implemented and integrated new backend features to support additional functionalities and improve performance.

PROJECTS

Mitigating Bias in Image Classification

- Pretrained models on diverse datasets for improved generalization. Showed method to mitigate bias using data manipulation.
- Paper accepted at the 7th International Conference on Electronics, Materials Engineering, and Nanotechnology.

Underwater Image Enhancement Transformer using Generative Adversarial Network

- Combined GANs and transformers using encoder-decoder structure to enhance underwater images. Research under Prof. Sumi Suresh, CSE Dept, Amrita University.
- Combined methods from multiple research papers to achieve near state of art results of PSNR: 25.87, SSIM:0.91, UIQM:3.12

EigenCAM for YOLO V11

- Developed a package for applying **EigenCAM** on YOLO V11 that got starred by over 100 people on GitHub
- Visualized model attention regions in classification, segmentation, and object detection tasks.