

Harish Vijayasarangan

harishvijayasarangan@gmail.com | [portfolio](#)
[linkedin](#) | [github](#)

Skills

Languages: Python, SQL, JavaScript.

Technologies & Tools: AWS, EC2, PyTorch, TensorFlow, Fine-tuning, MCP, LangChain, AI agentic frameworks, Microsoft AutoGen, CrewAI, LLM, Git, GitHub, Docker, Image Segmentation, Jenkins, DevOps CI/CD pipeline, RAG, Vector databases, Kubernetes, REST APIs, WebSockets, Nginx, Postman, Anaconda, Flask, FastAPI.

Work Experience

Sri Sairam Techno Incubator Foundation, India

Jan 2025 - Present

Research And Development Intern

- Trained **DenseNet121**(CNN-based architecture) to classify retinal images to find the severity of diabetic retinopathy (Gradual vision loss caused by **diabetes**) .
- Collaborated with **Ramachandra Medical Hospital**, acquiring high-quality labeled retina images.
- Worked hand-in-hand with ophthalmologists to better understand clinical features like microaneurysms, exudates, and hemorrhages, ensuring medically accurate model predictions.
- Handled class imbalance using focal loss and data augmentation.
- Integrated Attention Module(CBAM) to Focuses on spatial and channel-wise importance.
- Converted the trained PyTorch model to ONNX for production.
- Deployed on **AWS** EC2 instances and DigitalOcean Droplets using FastAPI and ONNX Runtime.

Securden, India

Jun 2024 - Jul 2024

SDE Intern

- Automated credential management via Jenkins CI/CD and API, streamlining creation, update, and deletion processes.
- Built custom Jenkins Pipelines for job automation and user management; improved DevOps efficiency.

Project Work

RAG : Developed a **RAG** PDF chatbot using Hugging Face embeddings with Chroma vector DB , Achieved over 90% retrieval relevance and sub-1s response time optimizing chunking, embedding, and real-time vector similarity search.

RetinAI : AI application to detect Diabetic retinopathy(Gradual vision loss caused by diabetes), Architected a decoupled system where the ML model is hosted independently on a FastAPI server, exposing REST APIs for inference.

DSA-Finetune : Fine-tuned(LoRA) Llama 3.1 8b with an alpaca dataset with 20k rows to solve DSA problems with Python. My Model has gotten 200+ downloads in Hugging Face.

DST : A website with a custom enterprise chatbot built using a vision model (Llama 3.2) with custom knowledge, built using Flask and JavaScript to provide an interactive user experience dealing with EEG scent interference .

Education

Sri Sairam Institute of Technology

B.E. in Computer Science and Engineering

Global Certification

Red Hat Certified System Administrator (RHCSA)

Expires Mar 2027