

# VIRAJ SANAP

1-919-345-5267 | [vasanap@ncsu.edu](mailto:vasanap@ncsu.edu) | [Linkedin](#) | [GitHub](#) | [Portfolio](#) | Raleigh, NC (Open to relocate)

## EDUCATION

### North Carolina State University, Raleigh, NC

Master of Science in Electrical and Computer Engineering

May 2025

GPA 3.93/4.0

**Relevant Coursework:** Advanced Deep Learning, Automated Learning and Data Analysis, Automated Software Engineering, Cloud Computing Technology, Computer Vision, Digital Imaging Systems, Neural Networks, Object Oriented & Design Development

### Savitribai Phule Pune University, Pune, India

Bachelor of Engineering Electronics and Telecommunication

May 2020

CGPA 7.93/10

## WORK EXPERIENCE

### NCSU - College of Agriculture and Life Sciences

Aug 2024-May 2025

Research Intern

Raleigh, NC

- Engineered an interactive dashboard using Python, Streamlit, and Folium that enhanced yield predictions and optimized planting dates by 15% in soybean production
- The dashboard is currently helping over 100 farmers across North Carolina make data-driven decisions, enhancing productivity and managing resources more efficiently.

### NCSU - Active Robotic Sensing Lab

Apr 2024-Dec 2024

Graduate Student Researcher

Raleigh, NC

- Architected Generative AI models in PyTorch for anomaly detection in Blueberry crops, elevating accuracy by 20%
- Leveraged High-Performance Computing (HPC) to accelerate model training by 30%. Utilized Git for version control, collaborating with a team of 5 researchers on codebase management, ensuring smooth integration and deployment of research components.

### Infosys Limited

Nov 2020-May 2023

Senior Systems Engineer

Pune, India

- Optimized backend systems using Java, Spring Boot, and SQL, reducing server response time by 30%
- Designed and implemented microservices and REST APIs, boosting performance by 23% through multithreading
- Orchestrated microservices deployment on AWS, integrating Kafka for real-time data streaming and establishing CI/CD pipelines
- Developed dynamic front-end applications with Angular and Node.js, enhancing UI readability by 25%.
- Spearheaded Agile development initiatives, accelerating bi-weekly service deployments.

## TECHNICAL SKILLS

**Languages :** Python, C++, R, Ruby, Java, Typescript, MATLAB

**Frameworks :** TensorFlow, PyTorch, SpringBoot, Angular, ReactJS, Rails, NodeJS

**Data Tools:** Git, Tableau, PowerBI, Docker, Kubernetes, Apache Kafka, Spark, JIRA

**Cloud and DevOps:** AWS, Microsoft Azure, Gitlab, CI/CD Pipelines, Github Actions

**Database and OS:** MySQL, MongoDB, PostgreSQL, Linux, Windows

**Certifications:** Microsoft Azure Cloud Fundamentals AZ-900, AWS Cloud Practitioner CLF-C02

## PROJECTS

### MedQUAD - Medical Question Answer System

- Fine-tuned LLaMA and TinyLLaMA models using the QLoRA method, leveraging the MedQuad dataset for training and evaluation.
- Implemented performance evaluation metrics (EM, BLEU) to assess and optimize model accuracy for medical question answering, achieving a 10% improvement over baseline models.

### Financial Fraud Detection Using RAG and LLM

- Built a RAG-based fraud detection system using MongoDB Atlas Vector Search and Google Gemini 1.5 Flash LLM for scam identification. Engineered a pipeline to transform user input into vector embeddings and retrieve top similar fraudulent records.
- Developed a React.js and FastAPI-based platform with user reporting to enhance fraud detection accuracy.

### Formula1: Data Pipeline and Analysis using Azure Databricks

- Implemented end-to-end data pipeline using Azure Databricks, Azure Data Factory, Azure Data Lake Gen2, and Spark.
- Utilized PySpark and SparkSQL for ETL tasks on a single-node cluster within Databricks environment on Formula1 datasets. Designed comprehensive PowerBI Dashboard to generate actionable insights from Formula1 datasets

## ACHIEVEMENTS

### NC Plant Science Initiative Hackathon

Raleigh, NC

First place

Oct 2024

- Amplified synthetic image data pipeline performance by integrating AWS SageMaker and S3 for dataset generation.
- Implemented advanced data augmentation techniques and leveraged YOLOv8, boosting accuracy in real-world plant classification.