Monisha Gottam

408-390-3506 | monisha.gottam@gmail.com | linkedin.com/in/monisha | github.com/monisha-g7

Summary

AI Full Stack Engineer with 3+ years' experience in LLMs, multimodal learning, and cloud-scale deployment via AWS, GCP, and MLOps. Skilled in building robust AI systems with Docker and Kubernetes.

Education

San Jose State University — San Jose, CA
Masters of Science in Applied Data Science — *Aug. 2023 – May 2025*
- **Relevant Coursework:** Machine Learning, Deep Learning, Artificial Intelligence, Data Mining, Database Systems, Data Visualization, Mathematics for Data Science

Jawaharlal Nehru Technological University — Hyderabad, India
Bachelors of Technology in Electronics Engineering — *Aug. 2014 – May 2018*
- **Relevant Coursework:** Data Structures and Algorithms, Statistics, IoT, Computer Networks, C, Java

Experience

AI Engineer Intern — Jan 2025 – Present
Codersgenie — *San Jose, CA*
- Fine-tuned **LLMs (GPT, LLaMA, Mistral)** using **RAG** and **prompt engineering**, increasing model accuracy by **85%** and decreasing hallucinations by **35%**.
- Built **AI-powered autonomous agents** with **LangChain**, **LlamaIndex**, **FAISS**, and **Pinecone**, automating **50+ workflows** and reducing manual load by **65%**.
- Deployed scalable **AI applications** with **Flask**, **FastAPI**, **React.js**, **PostgreSQL**, **Docker**, and **Kubernetes**, maintaining **99.98% system uptime**.

AI Solution Engineer — Jun. 2022 – Jul. 2023
Certa.ai — *Hyderabad, India*
- Developed frontend solutions using **React.js**, **JavaScript**, and **JSON**, reducing manual effort by **60%** across **5+ enterprise clients**.
- Built and integrated **20+ RESTful APIs**, boosting perfomance and reducing load time by **40%**.
- Designed an internal **AI support bot** for prompt engineering, accelerating workflows by **70%** and delivering demos to **15+ stakeholders**.

System Engineer - Data Science — Jan 2021 – Jun 2022
Tata Consultancy Services — *Hyderabad, India*
- Built ML models in **Python**, **R**, and **Java** for fraud detection and churn prediction, achieving **92% accuracy** and reducing false positives by **35%**.
- Developed scalable **ETL pipelines** using **Python**, **SQL**, and **Apache Airflow**, processing **10TB+ daily**.
- Automated model deployment via **MLOps pipelines**, reducing manual retraining time by **50%**.

Projects

AI-Powered Personal Assistants for Specialized Industries – *Python, LLMs, RAG, LangChain, Streamlit* – GitHub
- Developed domain-specific assistants using LLMs and RAG to streamline workflows in Legal and Healthcare sectors.

Autonomous Driving Systems with LVLM – *Python, PyTorch, LLaVA, CLIP, Hugging Face, OpenCV* – GitHub
- Simulated autonomous driving perception and reasoning using LLaVA and CLIP-based vision-language models.

Publications

- **An Outlook on the Status of Security Performance in the Light of Intelligence** – Google Scholar
- **How Machine Learning Can Be Used to Improve Predictive Analytics** – Google Scholar

Technical Skills

Data Analysis & Visualization Tools: SQL,PostgreSQL, Python PySpark, Power BI, Tableau,MS Excel, AWS, GCP
Technical Expertise: LLM Fine-Tuning, Machine Learning, Deep Learning, Statistical Modeling, MLOps, Feature Engineering, RAG, Prompt Engineering
Functional Expertise: Experimentation,Visualization, Cross-Functional Collaboration, Stakeholder Communication