Vidit Naik

3429 Canyon Crest Dr, Riverside, CA | (951) 425-7229 | viditnaik@gmail.com | linkedin.com/in/viditnaik

Professional Summary

Master's in Computer Science with hands-on experience in full-stack development and cloud infrastructure. Skilled in building real-time APIs, scalable systems, and CI/CD pipelines using Java, JavaScript, and AWS. Strong team collaborator with a focus on performance and agile delivery.

Education

University of California, Riverside March 2025

Master of Science in Computer Science (GPA: 3.63)

- **Relevant Coursework:** Operating Systems, Agile Development, SDLC, Data Structures and Algorithms, Object Oriented Programming, Machine Learning, Artificial Intelligence, Networking, Design Patterns, Distributed Systems

Vellore Institute of Technology, Chennai August 2023

Bachelor's in Computer Science and Engineering

Skills

Languages Java, Python, JavaScript, C++, SQL

Database MySQL, MongoDB, NoSQL

Frameworks and Libraries React.js, React Native, Hadoop, Node.js, Django, Express, RESTful APIs, Flask, SpringBoot

Testing PyTest, JUnit, Selenium

AI and ML Large Language Models (Qwen, BERT, Llama3), RAG, PyTorch, Fine-Tuning (LoRA), Ollama

Cloud and DevOps AWS (S3, Glue, Lambda, EC2), Docker, Kubernetes, CI/CD Pipelines, Jenkins, Git/GitHub, Jira

Tools and Methodologies Microservices, Tableau, Distributed Systems, RESTful API, Agile/Scrum, Jira, OAuth 2.0, Linux, Unix Shell Scripts

Experience

Center for Robotics and Intelligent Systems October 2024 - March 2025

Student Researcher *Riverside, CA*

- Developed an AI-powered system that integrates drones with Large Language Models (LLMs) for user-driven control through natural language.
- Utilizing Retrieval-Augmented Generation (RAG) to provide the LLM with contextual knowledge from drone manuals, increasing command accuracy by 68%.
- Collaborating in a Scrum-based Agile development process to design features with a core commitment to quality and scalability.

Kent Cam May 2022 - June 2022

Machine Learning Engineer *Noida, India*

- Utilized data analytics techniques to assess and optimize the integration of new features in company cameras, significantly improving core functionalities and enhancing overall product performance.
- Automated cross-region data pipelines, reducing data-validation errors by 35% and ensuring 99.9% data integrity across multiple data streams.
- Conducted thorough product testing, identifying and resolving critical software issues to ensure seamless and efficient performance.

StuDetails July 2020 - February 2021

Software Engineer *Noida, India*

- Built a scalable web application using React and Flask to manage scheduling and inspections for enterprise clients, supporting 5,000+ monthly users across 50+ distributed locations.
- Led cross-functional collaboration to optimize AWS S3 data workflows using AWS Glue, improving data accuracy by 25% and enabling real-time integration with client-facing systems.
- Implemented automated testing and CI/CD pipelines, cutting deployment failures by 30% for apps processing 1TB+ data monthly.

Projects

InsureSearch: RAG based AI Chatbot

- Deployed RAG and fine-tuned LLaMA3 to boost response accuracy, integrating seamlessly with Together AI's API.
- Reduced token usage by 65% with top-k BERT search, improving response efficiency and enhancing personalization.
- Enhanced LLaMA-3 8B with 4-bit quantization and LoRA adapters, fine-tuning on insurance policy data to boost accuracy for user prompts. Handled the source control of the project through Git.

CitySafe: Chicago Crime Insights Dashboard

- Built an analytics dashboard processing 1.5M+ records using SQL, Spark, and PostGIS with ETL pipelines for automated data extraction and crime trend insights.
- Designed geospatial visualizations with Python libraries and React.js, enabling self-service monitoring of crime hotspots.
- Reduced query response time by 25% by optimizing data models, indexing, and building a Python API for efficient data delivery.

CreateFlow - CalHacks 11.0

- Designed a multi-agent system with LangGraph linking AI agents for content, scheduling, & analytics, cutting creation time by 50%
- Enhanced LLaMA-3 8B with 4-bit quantization and LoRA adapters, fine-tuning on LinkedIn data to boost engagement predictions