Animesh Srivastava

animeshs2808@gmail.com | +1(323)-620-7865 | GitHub | LinkedIn | Portfolio | Los Angeles, California

Education

University of Southern California | Los Angeles, USA   **August 2023 – May 2025**
Master of Science in Computer Science (GPA: 3.6/4)

MS Ramaiah University of Applied Sciences | Bangalore, India   **August 2017 – June 2021**
Bachelor of Technology in Computer Science and Engineering

Work Experience

Software Engineer, Easley Dunn Productions | Los Angeles, USA   **June 2025 – Present**
- Built VR surgical training prototype in Unity(C#) with JSON-based session replay & diegetic UI, improving user interaction by 20%.

Full Stack Software Engineer(Founding Engineer), Redvest | Los Angeles, USA   **May 2024 – October 2024**
- Architected multithreaded **Python** ingestion pipelines for real-time financial analytics, reducing end-to-end data latency by 40%.
- Designed and deployed containerized backend services (**Node.js/Python**) on **AWS EKS**, implementing scalable KPI monitoring systems that eliminated 60% of manual reconciliation work.
- Led core feature development across backend and API layers, improving efficiency and boosting user engagement by 20% and presented product at TechCrunch Disrupt 2024.

Software Engineer, American Express(via TCS) | Bangalore, India   **January 2021 – July 2023**
Worked on two projects, **IDL (Indian Data Localization)** and **Travel and Lifestyle Services**
- Engineered distributed **ETL** pipelines using **Python, PySpark, PL/SQL, and Hive** to process multi-source financial datasets with 100% daily reconciliation accuracy across **OFSAA, Oracle, and MongoDB**.
- Automated **RBI-compliant** data localization validation by migrating 30+ regulatory rules, reducing compliance risk by 50%.
- Designed and optimized **PL/SQL** procedures and data quality validation layers across production tiers, improving consistency and reducing recurring ingestion failures by **20%**.
- Built internal **REST APIs** and **Java**-based automation tools supporting 50K+ daily transactions, eliminating **20+ manual** engineering hours per week, and also **eliminating manual report comparisons**.

Projects

Animatrix AI Agent - Product Review Aggregator
- Built a **distributed AI product** intelligence system that aggregates multi-source product reviews and pricing data, performs **LLM**-based sentiment and feature extraction, and generates structured product insights in near real time.
- Designed a **Retrieval-Augmented Generation (RAG)** pipeline using **Qdrant** for vector search over embedded review corpora, enabling context-aware summarization and feature comparison grounded in retrieved evidence.
- Implemented **LLM-driven AI agent** task orchestration with **Celery** and **Redis** to coordinate multi-step workflows (scraping, embedding generation, retrieval, LLM reasoning, and price normalization), ensuring scalable and fault-tolerant AI processing.

US Stock Market Portfolio
- Designed and implemented **RESTful** backend services in **Node.js** and **Express** for real-time stock data ingestion, portfolio tracking, and historical price analytics, optimizing MongoDB queries to reduce API response time by 20%.
- Built a **React**-based web application and a companion **Android** app that consumed a shared backend API layer, designing the backend to support multi-client access with consistent logic and scalable deployment on **AWS (EKS, EC2, S3, CloudFront)**.

Multi-LoRA(Low Rank Adaption) Integration Pipeline
- Developed a **multi-adapter LoRA** blending system for diffusion models to preserve character identity across AI-generated images, reducing feature mixing by 58%.
- Implemented region-based attention control and adapter blending on the **FLUX diffusion model** and validated improvements using **CLIP alignment** and **ArcFace** identity metrics across large-scale GPU experiments.

Research Papers

- **Trip Planner and Recommender using Flutter and Tensorflow** - "IEEE, 2022"
- **Conceptual Model Simulation to Detect and Report City Traffic Violations Using Distributed Intelligent Agents** - "IEEE, 2021"

Skills

Languages: Python, Java, C, C++, SQL, JavaScript, TypeScript, React.js, Angular, C#, Dart
Core: Distributed Systems, Multithreading, Memory Management, REST APIs, GraphQL
Backend: Node.js, Flask, Django, Express.js, Event-Driven Architecture, WebSockets
AI Systems: Generative AI(LLMs), RAG, LangChain, LangGraph, Vector Databases (Qdrant), Agentic AI, Diffusion Models
Cloud & DevOps: Google Cloud(GCP), AWS(EKS, EC2, S3, Lambda), Microsoft Azure, Docker, Terraform, CI/CD(Github Actions)
Data Engineering: PySpark, Hive, PostgreSQL, MongoDB, Redis, PL/SQL, MySQL