

AJAY KALLEPALLI

✉ ajay.r.kallepalli@gmail.com  [linkedin.com/in/ajaykallepalli/](https://www.linkedin.com/in/ajaykallepalli/) ☎ (424) 354-7363

EXPERIENCE

Software Engineer, GenAI

San Francisco, CA

ArangoDB

Oct 2024 - Present

- Developed **GraphRAG** microservices for Intel's open-source enterprise AI platform using FastAPI, Docker, and LangChain, enabling vector-based information retrieval, knowledge graph generation with LLMs, and chat history management.
- Fine-tuned an LLM-powered Copilot using Pytorch, achieving a **50% performance improvement** over the base model in producing Arango Query Language, enabling ArangoDB customers to query and generate insights from their graph database
- Generated synthetic training datasets in Rust for model fine-tuning, improving text feature extraction and knowledge graph generation, ultimately achieving a **30% improvement** in accuracy compared to the baseline model

Machine Learning Engineer

San Diego, CA

Qualcomm

Dec 2021 - Jun 2024

- Architected end-to-end computer vision data pipelines leveraging Python, AWS (S3, Lambda), and Bash, orchestrating secure and scalable handling of **petabyte-scale** multimodal datasets for facial modeling and synthetic avatar generation
- Achieved a **16x speedup** in CV data processing workflows by parallelizing preprocessing tasks (Tif Conversion, Color Correction, Normal Map Generation, etc.) using multiprocessing and distributed computing strategies
- Designed and implemented a custom image processing pipeline for facial avatar reconstruction using RealityCapture, OpenCV, NumPy, and RawTherapee across an 87-camera photogrammetry rig, eliminating **\$35,000** in third-party consulting costs
- Trained and led a remote annotation team, developing modular video tutorials and automating Quality Control protocols using Python scripts, ultimately resulting in **4000 hours saved** for engineering staff

Data Scientist, Consultant

Los Angeles, CA

UCLA Extension

Jan 2021 - Jun 2021

- Automated the migration of **30+** disparate excel sheets into a scalable PostgreSQL database using python, enabling efficient data analysis
- Leveraged R and Python to analyze marketing channel effectiveness through sentiment analysis and regression models, identifying key engagement factors that **reduced** missed classes **by 33%**

EDUCATION

University of San Francisco (USF)

San Francisco, CA

M.S. in Data Science, Data Engineering Concentration

Expected Jun 2025

University of California, Los Angeles (UCLA)

Los Angeles, CA

B.S Statistics; B.A. in Economics

September 2017 - June 2021

PROJECTS

Scalable Restaurant Data Pipeline & Analytics Dashboard

Tech Stack: MongoDB, Airflow, Google Cloud Storage, BigQuery, PySpark, Google Looker Studio

- Built a scalable Airflow pipeline to ingest Michelin ratings, and Google reviews from the Google Restaurants API into Google Cloud and MongoDB Atlas
- Conducted an analysis of Michelin-starred restaurants, integrating Google ratings, Sentiment Analysis, Trends, and Michelin data into interactive dashboards with Looker Studio

LangChain Hackathon, Customer Support Agent (2nd Place)

Tech Stack: LangChain, LangGraph, Streamlit, Pydantic, Qdrant, Docker

- Developed an AI-driven customer support agent utilizing retrieval-augmented generation (RAG) to automate issue resolution, achieving an 80% reduction in manual ticket handling
- Engineered an agentic workflow using LangGraph, chain-of-thought reasoning, and prompt engineering

SKILLS

Languages: Python, R, SQL, JavaScript, PowerShell, Bash, Matlab, TypeScript

Frameworks & Tools: PyTorch/TensorFlow, Git, Docker, Kubernetes, AWS, Apache (Hadoop, Spark, Kafka), Selenium

Libraries: NumPy, Pandas, scikit-learn, Hugging Face Transformers, LangChain, LlamaIndex, MLFlow, OpenCV