TIM WANG

MACHINE LEARNING ENGINEER

925-200-1910
timothy.yk.wang@gmail.com
linkedin.com/in/tim-yk-wang
github.com/tyw006
www.tyw006.github.io

SUMMARY

Experienced machine learning engineer with a strong focus on developing and deploying AI applications, particularly using Large Language Models (LLMs). Proficient in Python, LangChain, and vector databases like GCP Vector Search or Qdrant. Demonstrated success in leading AI projects from conception to implementation, optimizing prompts, and leveraging advanced techniques such as Retrieval Augmented Generation (RAG). Committed to continuous learning and innovation, with a proven ability to collaborate effectively in high-performance teams. Passionate about using generative AI to solve complex problems and drive efficiency in real-world applications.

SKILLS

- **Programming Languages:** Python, SQL, Git, Bash
- **OSS/CSS Large Language Models:** RAG, Evaluations, Fine-tuning, Deployment
- **Cloud Services:** AWS, GCP
- **Full-Stack Development:** FastAPI, Streamlit/Gradio
- **MLOps:** Airflow, Github Actions
- **Containerization:** Docker

EXPERIENCE

Machine Learning Engineer
OCT 2023 - PRESENT

Beam Data(Contract)

- Led the full-stack development of a RAG AI assistant for a Fortune 500 client, utilizing Milvus vector database, Langchain, Langchain Expression Language, GPT-4, and Streamlit.
- Successfully delivered a proof-of-concept on a tight deadline that met performance benchmarks, earning executive board approval for further development.

Data Scientist
MAY 2020 - PRESENT

Applied Materials

- Pioneered prompt engineering using open-source LLMs for classification tasks, enabling more detailed analysis of maintenance comments..
- Developed a web application incorporating computer vision algorithms, achieving over 90% accuracy in detecting product defects from images and reduced manual effort by nearly 100%.
- Created a widely adopted web application that automated routine data analysis for a service offering that allowed it to scale by more than 3x.

Equipment Engineer
MAY 2016 - MAY 2020

Samsung Austin Semiconductor

- Developed a novel magnet compensation equation that extended raw material life, resulting in record company output and annual cost savings of $100K.
- Led the end-to-end design and release of a novel part that reduced wafer sliding in the chamber by 33%, significantly decreasing product defects.

PROJECTS

Financial AIdvisor

- Developed an Agentic RAG-based AI assistant web application that intelligently routes finance queries to a Qdrant vector database and other queries to a web search tool. Utilized Langchain, Langserve, Langgraph, Command R+, and Streamlit for building the application. Achieved seamless and scalable deployment by containerizing with Docker and implementing a serverless architecture on Google Cloud Platform.

Olyver AI

- Developed an Olympic Lift Analyzer utilizing a custom-trained Detectron2 computer vision model to detect and estimate barbell speed across different lift phases. Deployed the model using Amazon SageMaker, integrated with API Gateway, Amazon Lambda, Amazon ECR, and Amazon S3 for efficient and scalable performance, and monitored the system with Arize for continuous optimization.

EDUCATION

Machine Learning Bootcamp Certificate
APR 2023 - OCT 2023

WeCloudData

Master of Science, Chemical Engineering
SEP 2014 - DEC 2015

University of California, San Diego

Bachelor of Science, NanoEngineering
SEP 2010 - JUNE 2014

University of California, San Diego