
Adam Villarreal

San Diego, CA • +1 559-305-3431 • adamvillarreal100@gmail.com • linkedin.com/in/adam-v99

Machine Learning Engineer

Machine Learning Engineer with 2 years of experience accelerating AI development. Built software tool with 99.94% accuracy, driving two published research initiatives. Guided future internal research strategy with model development. Deployed and maintained web app hosting 16 models. Optimized ML-assisted label generation and data processing, achieving 98.1% accuracy in data QA in 2 months.

WORK EXPERIENCE

Genentech

06/2022 - 06/2024

Pharmaceutical and Biotech Company utilizing AI

Machine Learning Engineer • Full-time

- Integrated 20 public datasets from different data sources to enhance data availability for ML algorithms, enabling multiple research experiments and accelerating model development.
- Engineered software tool in 3 months utilizing AWS Textract to digitize PDF Visual Field Reports with 99.94% accuracy. This data wrangling tool drove two research initiatives from the exported CSV data.
- Led internal research by prototyping models predicting Glaucoma retinal damage, guiding future research by communicating the results of the experiments.
- Built and maintained web app hosting 16 segmentation models for disease detection and Data Visualization, empowering the Ophthalmological community to explore research developments. Utilized Streamlit, Docker, AWS S3, and FastAPI.
- Optimized ML-assisted label generation and Data Pipeline workflow in 2 months, achieving 98.1% testing accuracy through feature engineering, cross-validation, and hyperparameter tuning with Tensorflow and W&B, resulting in higher quality data and more accurate labels.

SKILLS

Languages: Python (Proficient), R, SQL

AI/ML: Machine Learning Algorithms, Deep Learning, Supervised Learning, Unsupervised Learning, Regression, Classification, Clustering, Computer Vision, NLP, Large Language Models (LLMs)

Tools/Technologies: PyTorch, Tensorflow, Keras, Scikit-Learn, HuggingFace, Docker, Streamlit, FastAPI, Weights and Biases, AWS (S3|EC2|Sagemaker|Textract), Git, OpenAI, LangChain, ChromaDB

Data Skills: Data Preprocessing, Data Analysis, Data Visualization, Feature Selection

Soft Skills: Problem Solving, Communication, Collaboration, Ownership

EDUCATION

Bachelor of Science in Data Science

University Of San Francisco

08/2017 - 05/2021

- Data Institute Award for Excellence in Data Science - Math and Statistics Department

PROJECTS

RAG-Powered Personal Assistant LLM Agent

- Built RAG Conversational AI Chatbot that answers personal and professional questions about myself using LangChain
- Used Semantic Search and the Chroma DB Vector Database to store documents
- Used Prompt-Engineering and Query Rephrasing with Chat History to control LLM output
- Used the latest OpenAI's GPT-4o-mini and text-embedding-3-small models
- Built a web-app using Streamlit that provides a UI, and deployed it using an AWS EC2 instance
- Link to app: <https://adamchatbot.hopto.org/>

PUBLICATIONS

Digitalization of Humphrey Field Analyzer Reports into CSV file

AGS Society

Repeatability of a Virtual Reality Headset Perimeter in Glaucoma and Ocular Hypertensive Patients

TVST ARVO