
Adam Villarreal

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

Machine Learning Engineer

Past Data 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 scalable 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 • Remote

06/2022 - 06/2024

Data 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, driving 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 in 2 months, achieving 98.1% accuracy through 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

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

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

Soft Skills: Problem Solving, Communication, Collaboration, Ownership

EDUCATION

Bachelors Of Data Science

University Of San Francisco • GPA: 3.86

08/2017 - 05/2021

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

PROJECTS

SpongeBob Script Generator

- Fine-tuned a GPT-2 LLM model with 393 SpongeBob episode scripts using HuggingFace
- Built with an intention to teach teens NLP basics like text-vectorization and transformer models
- Implemented a web app using Streamlit to generate a whole page of a generated script. This app included the feature of being able to choose a character to start the script
- [Link to App here](#) and [Code here](#)

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