Daniel DeFoe
Data Scientist

(650)-773-1428 | danield2255@gmail.com | www.linkedin.com/in/daniel-defoe/ | www.github.com/danield2255

About Me:

I am data scientist and data science educator with a love for problem solving and continuous learning. I am always hoping to learn relevant technologies, as well as apply tried-and-true tools with best principles to develop strong software that does its job well. I am looking for opportunities where there are interesting problems to take on and where people are looking to use data and information technologies for benevolent ends.

SKILLS:

- Python (10 years)
- ML model development (9 years)
- LLM and Gen AI development (1 year)
- Computer Vision (4 years)
- SQL (9 years)
- Statistical Testing (9 years)
- Keras / Tensorflow (5 years)
- Pytorch (2 years)
- AWS (5 years)
- R (6 years)
- Natural Language Processing (5 years)
- Java (2 years)

WORK EXPERIENCE:

Data Science Expert Tutor - TripleTen *Jan 2021 - Present*
- Provided individualized and group tutoring and academic support to over 1000 students enrolled in an online Python-based data science bootcamp
- Facilitated over 500 group workshops covering core data science concepts
- Developed instructional scripts for video-based learning materials
- Conducted over 200 mock interview sessions supporting entry level data scientists enter the workforce

Mid-Level Data Scientist - Whitespace *Oct 2022 - Oct 2025*
- Developed machine learning pipelines for systems within large AWS projects
- Used ensemble boosting models to reduce error of a government model by more than 50%
- Built an entity resolution system that was able to consolidate over 2 million accounts with duplicates down to 75k accounts across government database systems
- Worked on FastAPI systems to conduct operations on large databases in Databricks
- Interviewed and contributed to hiring other data scientists and developers
- Prototyped LLM systems that discussed business documents with customers using RAG

Backend Developer - Computer Vision Engineer - KOS Technologies *Sept 2021 - Oct 2022*
- Developing computer vision models for image classification
- Developed photogrammetry pipeline for mobile application

Computer Vision - Advanced Technology Electronics Engineer - SNSR Fusion Technologies *Feb 2021 - Sept 2021*
- Used Yolov4 models with Tensorflow to make image classification and instance segmentation pipeline on edge devices for automotive safety technology
- Created a system to overlay arrows on pedestrians and cars tracking the direction of their movement in live video

Research Assistant -Cal Poly Corporation *Feb 2020 - Mar 2022*
- Led object detection and instance segmentation research in Python
- Augmented image data and maintained relative annotation placement
- Trained models with Faster R-CNN and Mask R-CNN frameworks on distributed GPU cloud systems
- On a different project, trained clustering models on Uber rides to study rider behavior

EDUCATION & CERTIFICATES:

B.S in Statistics, minor in Data Science *September 2016 - Jun 2020*
California Polytechnic State University San Luis Obispo | San Luis Obispo, CA

Nebius Generative AI Bootcamp Certificate *September 2024 - May 2025*

Tensorflow Certified Developer *December 2020 (Expired December 2023)*

OTHER INTERESTS:

Playing and making music | Audio Engineering | Hobbyist woodworking | Reading