Yuliang (Leo) Chen

La Jolla, CA 92093 | 805-637-4650 | yuc204@ucsd.edu

EDUCATION

University of California, San Diego — **San Diego, CA**
Candidate for Master of Science in Data Science — Sep 2023 – Jun 2025 (Expected)
- Current GPA: 3.96/4.0

University of California, Santa Barbara — **Santa Barbara, CA**
Bachelor of Science in Mathematics and Statistics & Data Science — Sep 2018 - Jun 2022
- Achieved a GPA of 3.43/4.0

Relevant Coursework:
Computer Science, Distributed System, Advanced Machine Learning, Interpretable ML&AI, Computer Vision, Probability Theory, Statistics Inference, Linear Algebra, Mathematical Analysis, Stochastic Process Analysis, Time Series Analysis, Data Visualization, Database Management, Big Data, Data Processing, Optimization

PROFESSIONAL EXPERIENCE

MOSAIC Lab — **San Diego, CA**
Research Assistant — Mar 2024 – Current
- Leveraging large language models (LLMs) for multimodal and multitask analysis on multi-variate time-series data to develop time series foundation models for various downstream tasks.
- Contribute to the development of tiny foundation models tailored for deployment on edge devices like Jetson Nano, emphasizing efficient resource utilization and low computational footprint.

University of California, San Diego — **San Diego, CA**
Research Assistant — Oct 2023 – Jan 2024
- Independently collecting and preprocessing medical image datasets comprising more than 500,000 publicly available images.
- Employing self-supervised learning, like Masked Autoencoder, to train foundational medical image models on Kubernetes clusters.
- Finetuning LVMs on low-resource datasets, aiming high accuracy in downstream tasks such as segmentation and detection for disease classification.

Micro Ingredients — **Montclair, CA**
Supply Chain Analyst — Oct 2022 – Sep 2023
- Spearheaded a three-member team in the implementation of a Random Forest Model, achieving a notable 83% accuracy in demand forecasting and subsequently realizing a 27% enhancement in product availability.
- Collaborated with the software engineering team to established database system and employed web scraping tools to collect and store data effectively for subsequent data analysis.
- Proposed and collaboratively implemented a Python-based Streamlit dashboard to visualize data, thereby effectively transforming raw data into actionable insights.

PROJECT EXPERIENCE

Image-to-Image Search Using CLIP — Nov 2023 – Dec 2023
- Independently utilized PyTorch to develop an image-to-image search engine, incorporating advanced models such as CLIP and ResNet for feature extraction.
- Demonstrated superior performance of CLIP in handling semantic information through comparative analysis.

Predictive Models for Presidential Election — Jan 2021 – Mar 2021
- Utilized the R programming language to perform extensive Exploratory Data Analysis (EDA) and Principal Component Analysis (PCA) on election voter demographic data, extracting valuable insights into voter behaviors.
- Developed and fine-tuned QDA, logistic regression, and random forests for county-level prediction models, resulting in a 92% accuracy rate with the random forests model.
- Systematically assessed candidate model sets using diverse statistical metrics to achieve peak performance while maintaining computational efficiency.

SKILLS

Computer: Python (PyTorch,Pandas,Ray,TensorFlow), R, MySQL, C++, SAS, Microsoft Office Suite