Alan Wang

858-539-5564 | Authorized to work in U.S. with CPT/OPT | zew@seas.upenn.edu | zew013.github.io

Education

University of Pennsylvania	*M.S. in **Data Science***	2024 - 2026
University of California San Diego	*B.S. in **Data Science**, B.S in **Applied Math & Economics***	2020 - 2024

***Relevant Coursework*:** Algorithms, Statistics, Regressions, Hypothesis Testing, Data Mining, Forecasting, Data Visualization, Machine/Deep Learning, Computer Vision, Database, Big Data Analytics, Optimization, Time Series, Large Language Model.
***Programming language*:** Python (NumPy, Pandas, Matplotlib, Plotly, Scikit-Learn, PyTorch, Spark, Dask, etc.), SQL, R, Java
***Others*:** Tableau, Git, AWS, Unix shell scripting, Kubernetes, Microsoft Office Suite

Professional Experience

Data Modeling Intern | TE Connectivity   **May. 2024 – Aug. 2024**

- Composed large datasets by collecting and consolidating data from various sources. Collaborated with multiple teams to access and integrate historical data, which streamlined the data collection process for future projects.
- Launched an Auto-ML pipeline on AWS SageMaker to improve cost estimations. Reduced cost estimation time from hours to 10 minutes, allowing cost analysts to focus on strategic decision-making rather than manual estimates.
- Designed cost models and generated statistical insights using Excel-VBA and retrieving cloud data for real-time updates.

Machine Learning Intern | Grant Street Group   **May. 2023 – Aug. 2023**

- Proposed and developed a machine learning powered monitoring system. Tested models like Random Forest, ARIMA, Prophet, and Temporal Fusion Transformer for self-supervised anomaly detection.
- Manipulated hundreds of millions of data points using SQL and Python, and composed the database program for automated model retraining and update. Led a team of four to implement the new system which improved the F1 score from 0.15 to 0.6 to replace the previous static threshold-based system.
- Proficient at using SQL, Tableau, Python to deliver data visualizations and statistical analysis for daily operations.

Machine Learning Head Teaching Assistant | Halicioglu Data Science Institute, Penn Engineering   **Mar. 2023 – Present**

CSE-PACE Program Designer | UCSD CSE Department | *Webpage*   **May. 2022 – Sep. 2022**

- Implemented the NSF-supported peer-mentor program as part of the CS curriculum that supported over a thousand students.

Research Experience

Data Science Capstone Owner | Prof. Alex Cloninger   **Oct. 2023 – Apr. 2024**

Diffusion Models for Image and Data Generation | *GitHub, Webpage*

- Investigated how scene representations are generated during the diffusion process. Demonstrated that 3D properties are learned early in the denoising stage before human visual recognition. Achieved by inserting probing classifiers into self-attention blocks.

Research Assistant | Rappel Laboratory   **Feb. 2023 – Oct. 2023**

Image Segmentation and Propagation Analysis Program for cAMP Waves in Cell Aggregation Stage | *Slides Demo*

- Developed a two-stage Python program that segments more than 60 GB of images and videos, and applies an unsupervised clustering algorithm for data cleaning and constructs velocity vector fields for scientific analysis.

Research Assistant | Prof. Richard Carson & Prof. Dale Squires   **Dec. 2021 – Dec. 2022**

Data-Driven Analysis of Ethical Preferences in UN Membership Policies & Assumptions in Conditional Logit Model

- Developed an ETL data mining pipeline using Python and AWS to create a large dataset from 70 years of documents, improved efficiency and accuracy, especially for handwritten records. Performed statistical analysis to generate insights.

Leadership & Projects

Data Analyst, Tech VP | Lumnus Consulting (Student Enterprise) | *Webpage*   **Nov. 2021 – Feb. 2023**

- Organized alumni talk events, fostering knowledge sharing. Led team projects with data analysis and visualizations, facilitating clear communication of data insights. Launched and maintained the company website using React.js and Heroku.

Deep Learning Projects

YOLO model Implementation for object detection & SOLO model Implementation for instance segmentation.

Language Intention Classification Using BERT & Model Compression for Web Deployment via Heroku, REST api | *Webpage*

Using CNN and LSTM models for Image Captioning on COCO Dataset | *GitHub, Report*

Building Neural Network from Scratch & Building Transformer in PyTorch | *GitHub, Report*