

Wan-Chun(Elena) Liao

elenaliao1002@gmail.com | San Francisco, CA | www.linkedin.com/in/elenaliao1002/ | github.com/elenaliao1002 | (530)-7809327

EDUCATION

M.S. in Data Science, University of San Francisco (USF), USA 2022 - Jun. 2023

Relevant Coursework: Machine Learning, Deep Learning, Linear Regression, Time Series, A/B Testing, EDA, MLOps, Distributed Computing (Apache Spark), Relational Databases (SQL), NoSQL(MongoDB)

Award: Schmidt Family Foundation Fellowship

B.S. in Agricultural Economics, National Taiwan University (NTU), Taiwan 2015 - Jan. 2019

Award: Top-class Scholarship(1%)

PROFESSIONAL EXPERIENCE

The Nature Conservancy

California, USA

Machine Learning Engineer Intern

Oct 2022 - Present

- Implemented the EfficientNet-B3 classifier to retrained the multi-class image classification models resulting in 91% average precision. Planning to deploy the model within the existing AWS infrastructure. [\[GitHub\]](#)
- Deployed an end-to-end data pipeline using Github Actions, including data acquisition, preprocessing satellite data, and modeling to estimate duration of farmlands flooding to protect migrating birds. [\[GitHub\]](#)
- Developed a PySparkML K-means model for predicting flooded area coverage. Automated the data pipeline using Airflow, Google Earth Engine API, MongoDB, and Databricks.[\[GitHub\]](#)

Bank SinoPac

Taipei, Taiwan

Data Analyst(Risk Analyst)

Apr. 2020-Aug.2021

- Evaluated clients' probability of default by analyzing customer profiles, loan data, collateral asset to helps loan officers to making sound risk decision commensurate with business units risk appetite for grant loans.
- Provided critical information to the senior managers to oversee the quality of the bank's assets by adopting RAROC methodology, a risk-based profitability framework.

Data Warehouse Engineer

Mar. 2019-Apr. 2020

- Developed an ETL-based transaction analysis system and dashboard to visualize customer segmentation, provide actionable insights for product managers resulting in a 30% increase in operational efficiency.
- Conducted EDA, data cleaning, feature engineering, and data visualization with R, SQL, and PowerBI to apprise bank performance for Senior Manager and Planning Dept.
- Contributed to data warehousing architecture by gathering requirements, defining user acceptance testing, documenting changes, and monitoring results for improved system performance.

FruitPay

Taipei, Taiwan

Purchasing Analyst Intern

Apr. 2018-July 2018

- Devised agricultural procurement strategies to lower storage cost to nearly zero and decrease spoilage rate by 30% through applied ARIMA & Holt-Winter time series models with R.
- Analyzed RFM, customer fruit preference, delivery method, and retention to customize subscriber's "Unique Fruit Box" and reduce cancellation subscription rate by 20%.

SKILLS

Programming Languages: Python(Pandas, Numpy, Scikit-learn, Plotly), SQL(Postgres, MS SQL), R

Skills: AWS (S3, EC2, Lambda), PyTorch, git, NoSQL (MongoDB), Docker, Apache Spark, Kafka, Google Cloud Platform, Databricks, Power BI, Tableau, Flask

PROJECTS

Netflix Homepage Browsing Time Optimization (Hypothesis Testing, A/B Testing)

- Designed a series of multivariate experiments to investigate the main & interaction effects of 4 design factors and determine optimal configuration and influential design factors. [\[GitHub\]](#)

BBC News Search Engine

- Cleaned 7k+ news files, implemented search engines in Python with object-oriented hash tables and word2vec from scratch, building Flask-based website (jinja) for representing searched keywords and relevant sentences.

Other Featured Projects

- Machine Learning Algorithms (Decision Tree, Random Forest, Gradient Boosting, etc) from scratch in Python. [\[GitHub\]](#)
- Feature Importance Selection Methods (Column-Drop, Impurity-Drop, Permutation Importance, etc.).