Yu Cheng

yu@yu-cheng.co
github.com/fischcheng
yu-cheng-boston
(786) 503-3366

PROFILE

Driven Data Scientist with a strong background in machine learning algorithms, predictive modeling, and data analysis. Proven success in assembling and analyzing data sets from diverse sources, developing data pipelines, and effectively communicating complex findings to non-technical audiences.

EDUCATION

University of Miami
PhD in Meteorology and Physical Oceanography, 2018

National Taiwan University
Bachelor of Science in Atmospheric Sciences, 2010

SOFT SKILLS

- Oral and written communication
- Creative problem-solving
- Research project management

TECHNICAL SKILLS

- Advanced Analytics
- Machine/Deep Learning: Scikit-Learn, PyTorch
- Languages: Python (numpy, pandas, matplotlib), SQL
- Workflow: Airflow, Docker, Kubernetes, dagster
- Cloud: Google Cloud Platform, BigQuery, AWS, SageMaker

CERTIFICATES

- Deep Learning Specialization
- AWS Certified Machine Learning – Specialty

05.18.2023

PROJECTS

Dashboard of Dark Ship Detection and Monitoring
Orchestrated a CNN-based illegal fishing vessels detection and labeling system, designing and implementing end-to-end data workflows for large-scale data ingestion, processing, tagging, and publishing.

Wildfires and climate change
explored the relationships between wildfires and climate variables, using Jupyterlab and matplotlib. Preliminary results showed that the wildfire season over the western U.S. has extended by nearly 70 days since 1980s. [blog post]

EXPERIENCE

Web3 Builders   2022 - Present
Founding Senior Data Scientist   Boston, MA
- Integrated disparate data sources to develop labeled datasets for ML models, predicting scam probabilities. Established ETL and ML pipelines using cloud infrastructures.

SFL Scientific (acquired by Deloitte)   2021 – 2022
Senior Data Scientist   Quincy, MA
- Implemented and deployed ML solutions tailored to Fortune 500 clients' objectives, resulting in improved business performance through meticulous exploratory analysis, metric definition, and continuous evaluation of data science life-cycle outcomes.

Tomorrow.io   2018 – 2021
Senior Atmospheric Data Scientist   Boston, MA
- Created novel weather products by integrating computer vision and ML techniques, resulting in a near real-time global precipitation pipeline using geostationary satellite images and cloud-based technologies using Google Cloud Platform, xarray, dask and satpy. [blog post]

RSMAS, University of Miami   2012 – 2018
Research Assistant   Coral Gables, FL
- Explored Agulhas Leakage variability through a state-of-the-art high-resolution climate model using Lagrangian particle analysis and open-source software including PyData stack (pandas, numpy, matplotlib, xarray, jupyter), CDO and NCL, resulting in 4 peer-reviewed articles.