WEI-CHIH HUANG

noctildon2@gmail.com | /in/wei-chih-huang | noctildon | Personal Site | Publications

EDUCATION

PhD in Physics, Texas A&M University, US — Aug. 2019 - Dec. 2025
BS in Physics, National Tsing Hua University, Taiwan — Aug. 2015 - Jun. 2019

EXPERIENCE

Data scientist internship - Capital One Auto Finance — Jun. 2025 - Aug. 2025

- Built PyTorch machine learning models (GBM, NN, LSTM) to optimize loss mitigation strategies for auto loans
- Designed customized time series model, training loop and loss function to better aling with business needs
- Developed predictive models for future payments and chargeoff probability with 99% accuracy
- Fetched 10 TB of data from SnowFlake and did statistical data analysis on AWS
- Collaborated with product managers to translate model outputs into action-based decisions

Research Assistant - Physics Department, Texas A&M University (researcher profile) — Aug. 2019 - Jun. 2025
PhD dissertation on dark matter search

- Designed machine learning models, eg random forest, neural network, to reduce 90% of time on particle simulation
- Built physics models and conducted the statistical analysis on 1000M rows of multi-dimensional data by Python
- Automized and visualized the analysis with NumPy, SciPy, Pandas, and Matplotlib
- Accelerated the analysis by 1000 times with dedicated algorithm, multiprocessing, caching, and C++
- Completed computational heavy calculations in MPI/OpenMP computer cluster (20 TB RAM, 3000 CPU cores)
- Published 7 papers in high impact journals and presented successful talks at international conferences

Quantitative Engineer - Aggie Quant Fund — Jan. 2024 - Dec. 2024
Application of cutting-edge technologies to financial market

- Managed $100,000 fund and developed models for stock forecasting and portfolio optimization
- Used cloud and local LLM and GitHub Actions to extract information from finance market news feed periodically
- Built an efficient, automatic and high-performance stock price database and dashboard using InfluxDB
- Built a high-performance backtest framework supporting automization and visualization and saved 70% of time
- Benchmarked, tested and validated trading strategies (outperform S&P500 by 200%)

Independent Data Science Researcher - Pro Cyclists Race Analysis (Github repo) — Apr. 2022 - present
Data analysis of professional sports and forecast using machine learning models

- Construct machine learning models with PyTorch and scikit-learn for 90% accuracy prediction (RNN model)
- Efficiently processed 12M rows of data with NumPy, Pandas, SciPy, scikit-learn, and PySpark
- Saved 80% costs compared to AWS, GCP, Azure by deploying data and model to Runpod (GPU cloud)
- Web scraped a website using BeautifulSoup and increased the performance by 500% with multi-threading

CERTIFICATIONS

- Fundamentals of Accelerated Computing with CUDA C/C++
- Machine Learning Foundations: Algorithmic Foundations
- Machine Learning Foundations: Mathematical Foundations
- Machine Learning Techniques
- Divide and Conquer, Sorting and Searching, and Randomized Algorithms
- A Crash Course in Causality: Inferring Causal Effects from Observational Data