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

- Reduced 10% loss for auto loans with machine learning models (GBM, NN, LSTM) deployed on AWS
- Improved AUC by 10% for future payments and default probability
- Designed customized PyTorch model, training loop and loss function to better align with business needs
- Built a scaleable data pipeline that fetched 10TB+ data from SnowFlake and processed analysis on AWS
- Collaborated with 4 product managers to translate model outputs into action-based decisions

Machine learning engineer - Firelight Innovations — Aug. 2025 - Dec. 2025

- Engineered a cross-platform SDK (Python/TypeScript) to abstract BigQuery interactions, standardizing telemetry collection and simplifying data ingestion for internal teams.
- Deployed a scalable data lake architecture on GCS and BigQuery to drive visualization metrics for system health and user behavior
- Implemented PEFT pipelines for fine-tuning SLMs using Ollama and PyTorch, optimizing model inference for deployment on edge devices

Quantitative Researcher/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
- Outperformed S&P500 by 200% by AI-driven strategies (sentiment analysis, LLMs, alpha research)
- Automated market insight extraction with LLM-powered pipelines, GitHub Actions, and Kubernetes
- Saved 70% query time by high-performance market database (InfluxDB and PostgreSQL)
- Collaborated in a 10-person team to optimize portfolio, mitigate risks and monitor trades

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

- Published 7 papers in top journals and presented at international conferences
- Processed 1B+ rows of multi-dimensional data with Python/PySpark and accelerated analysis 1000× using C++, multiprocessing, and caching
- Orchestrated large-scale distributed data processing and HPC optimization using MPI/OpenMP and PySpark
- Reduced particle simulation runtime by 90% with ML models (scikit-learn and PyTorch)
- Cut visualization runtime by 80% using optimized NumPy/Pandas/Matplotlib workflows
- Built an interactive platform agnostic data dashboard using Streamlit and Docker

SKILLS

ML/AI	PyTorch, scikit-learn, TensorFlow, NLP, Transformers
MLOps/Infra	AWS, GCP, Snowflake, Linux, Docker, Kubernetes, Git, GitHub Actions, CI/CD
Data/Database	NumPy, Pandas, SciPy, PySpark, InfluxDB, SQL, PostgreSQL
Languages	Python, Bash, C/C++, SQL
Visualization	Matplotlib, Seaborn, Streamlit