

CHIAO-WEI HSU

☎ 631-710-8463

✉ cwhsu.career@gmail.com

🌐 chiao-wei-hsu

🌐 ml.chiaoweihsu.com

🎓 EDUCATION

Stanford University (Auditing Courses)

Palo Alto, CA

Digital Nomad at Startup House 📄

Since Sep. 2024

- 📌 AI Agents and Simulations (Individual/Group/Population-Level Generative-Agent Based Model and Simulation and Equilibria), Parallel Computing (CUDA, CUTLAS, SIMD, RDD), Big Data Systems (nanoDB, transaction scheduling and synchronization).

Northwestern University

GPA: 3.88/4.0

Evanston, IL

Master of Science in Computer Science

Since Sep. 2023

- 📌 **Thesis:** “*Scaling Law for Deep Learning*” supervised by Han Liu 📄.
- 📌 **Seminar Project:** “*Financial LLMs and FinGPT*” supervised by Han Liu 📄.
- 📌 ML: Minimax, Bayesian Learning, Causal Inference, Decision Tree, Logistic Regression, EM Algorithm, Naive Bayes, Knowledge Representation and Reasoning, Deep Reinforcement Learning: A2C, PPO.
- 📌 DL: Quantization, LoRA, Retrieval-Augmented Generation, Diffusion Model, Linear RNN (Mamba), Transformer with RLHF, Mixture of Experts.
- 📌 CV/Graphics: Color Constancy, In-Pixel/On-Sensor Computation, Coded Aperture, Lensless Imaging.

👛 EXPERIENCE

Machine Learning Intern

Irvine, CA

Himax Imaging Corp.

Jun. 2024 – Sep. 2024

- Engineered a unified model for **Emotion, Attention, Body Detection, Face Expression** model based on **YOLO** with **TFLite** and **Vela**.
- Utilized **Model Pruning, Quantization, and Quantization-aware Training** to obtain a compression ratio of 46% on an **endpoint AI device** (milliwatts).
- Implemented Task Automation with Natural Language Commands by a fine-tuned **LLM** and **Text-to-Speech** and **Speech-to-Text** for AIPC.

Machine Learning Engineer (Research)

Taiwan

Academia Sinica

Nov. 2018 – Dec. 2022

- Engineered a **Knowledge-driven Multi-Modal QA System** integrating **vision, speech, and text** data, advancing capabilities in cross-modal information processing [3].
- Enhanced **state-of-the-art** performance by 4% on the **Named Entity Recognition (NER)** task through strategic **fine-tuning of Transformers**, applying **Multi-Task** and **Incremental Learning** techniques, and leveraging **Data Augmentation** [1][2].
- Designed a **Hierarchical Multi-Head Multi-Layer Attention** mechanism from the ground up, enhancing model interpretability and performance.
- Implemented **Adaptive Boosting** for Ensemble Learning.

📁 PROJECTS

Stock Trading using Deep Reinforcement Learning 📄 VIDEO

Mar. 2024 – Jun. 2024

- Trained a **Stock Trading Agent** by A2C and PPO to obtain an annual return of 29%.

Information Retrieval and Multi-Document Analysis 📄 CODE

Jan. 2019 – Jan. 2021

- Implemented **Entity Linking** and **Relation Extraction** for **Wikidata Knowledge Graph**.
- Incorporated **K-Mean** and **Hierarchical Clustering** algorithms to group similar documents.
- Improved the **Re-ranking** model of an Information Retrieval System by 0.24 f1-score.
- Implemented a BERT-based **Query Generation** model and threshold-tuning P-R curve.
- Extracted syntactic patterns based on **Dependency Parsing** and **Part-of-Speech Tagging** based on LSTM.

📖 PUBLICATIONS

- [1] Hsu et al., “A Redundancy-Enhanced Framework For Error Correction in Named Entity Recognition”, pre-print ’24.
- [2] Hsu et al., “GlobalNER: Incorporating Non-local Information into Named Entity Recognition”, pre-print ’23.
- [3] Liang et al., “A Flexible and Extensible Framework for Multiple Answer Modes Question Answering”, ROCLING ’21.

HONORS

- 🏆 **First Place** at the 2020 Formosa **Artificial Intelligence Grand Challenge**; developed a Multi-Round Dialogue System and Question Answering System.
- 🏆 Led the team of 4 students to **top 11%** in 2024 eBay Machine Learning Competition.

OTHER PROJECTS

AI-Driven Grocery Logging and Recipe Suggestions App 📱 DEMO

Sep. 2023 – Apr. 2024

- Adopted **ReactJS**, **Vite**, and **Firebase**; conducted tests using **Cypress**.
- Developed a **Serverless RESTful API** with **ExpressJS** as back-end deployed with **AWS ElasticBean**.
- Incorporated **API Gateway** with **AWS Lambda Functions**.
- **Engineered prompts for GPT-4 Vision**, deployed **OCR** to scan receipts, and implemented features to send expiry notifications and receipt suggestions.

AI-Powered Driver Companion Mobile App 📱 DEMO

Mar. 2024 – Apr. 2024

- Incorporated AI models, **YOLOv4**, **GPT-4**, **Text-To-Speech**, **Speech-To-Text**, to streamline from image recognition to natural language processing.
- Adopted **Django** as back-end, and **React Native**, **Expo** for mobile development.