

CHIAO-WEI HSU

631-710-8463

cwhsu.career@gmail.com

chiao-wei-hsu

ml.chiaoweihsu.com

EDUCATION

Stanford University (Auditing Courses)

Digital Nomad at Startup House

Palo Alto, CA

Since Sep. 2024

- AI Agents and Simulations (Individual/Group/Population-Level Generative-Agent Based Model and Simulation for Emergence and Equilibria) with my final project: *RecruitSim*.
- 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

- Project: “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

Himax Imaging Corp.

Irvine, CA

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.
- Integrated Websocket, Shared Memory, and Named Pipe for Inter-Process Communication among the microcontroller, the host PC, and the application.
- Developed a library to communicate with the microcontroller FT4222 driver.
- Developed an LLM-powered Windows Desktop App that integrates with the edge device and operating system on AIPC by ElectronJS, ReactJS, and C++/C# with microcontroller library and Qt GUI framework.
- Developed a RESTful API for the AI PC by Django and Django REST framework with PostgreSQL.

Machine Learning Engineer (Research Assistant)

Academia Sinica

Taiwan

Nov. 2018 – Dec. 2022

- Developed a Person-Company Profile Generation System with Django, MySQL Database, and typical Natural Language Processing pipelines.
- nlds: Developed a python library of natural language data structure with Design Patterns—Composite, Factory, Template, Command, Facade, Adapter, Delegate, Builder, Strategy, Iterator, Dependency Injection, and Meta Class.
- 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.

Software Developer

Luye Gaotai

Taiwan

Oct. 2016 –Part-Time– Aug. 2017 –Full-Time– Jul. 2018

- Developed a 3D vision graphics (autostereogram) program by JavaScript.
- Developed a portal website using HTML5/CSS3, BootStrap, PHP, and JavaScript.

Alternative Military Service (Software Developer)

Taitung County Government

Taiwan

Sep. 2016 – Aug. 2017

- Developed a **Windows GUI Desktop Application** (CODE [🔗](#)) to automate newspaper clipping, publisher annotation, scanning, emailing, and titling with AutoHotKey (an interpreted scripting language with OOP support) and **OCR** (Optical Character Recognition).

Co-Founder & COO

MiDs/Tidymen [🔗](#)

Taiwan

Mar. 2016 – Oct. 2016

- Managed, organized and led team with CEO to make big decisions, and set up goals for our team.
- Left the team after pivoting from a tech startup to a non-tech service startup, which fundraised 2M dollars in 2023.

Co-Founder & Software Developer

Unite

Belgium

Aug. 2015 – Feb. 2016

- Developed two **mobile app** prototypes with **Java** and **Ionic** respectively for “tribe gatherer” (PITCH [🔗](#)) based on crowdsourcing network (i.e., crowd coordination) in timely connection to street security, scientific work, coordinative project, or community work. (DEMO [🔗](#))

PROJECTS

RecruitSim [🔗](#) CODE

Jan. 2019 – Jan. 2021

- ...

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**.

Decentralization and Blockhcains System [🔗](#) VIDEO

Mar. 2024 – Jun. 2024

- Developed a **P2P Network** for the blockchain system using **Rust**.
- Implemented the data structures (e.g., Transaction, Block, Blockchain, Mempool) and algorithms for **Gossiping Protocol** and the **Proof of Work** consensus for the mining process.

Interpreter by Racket/Lisp [🔗](#) CODE

Jan. 2024 – Mar. 2024

- Built a **parser**, a **compiler**, and a **garbage collector** using **Mark & Sweep** and **Two-Space Collection**.
- Implemented features of deferred substitution, higher order functions, state, error handling, and recursions.

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.

Robotic Arm with Arduino [🔗](#) DEMO

Sep. 2015 – Dec. 2016

- Implemented and complied the inverse kinematics algorithm, and burnt **bootloader** in **microcontroller** of **Arduino**.
- Designed the robot arm with **3D printing**, **laser cutter** and assembled the components with **C++**.

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.
- ❖ **3rd Place** in the Taitung **Mobile App Design Competition** 2015 for innovative app development.
- ❖ **Silver Award** at Global Startup Weekend 2015; **Silver Award** at Sanmen Startup Competition 2016; **2nd Place** at PitchCamp for building an upcycling C2C/B2C platform and consultancy as a **then-cofounder**.
- ❖ **Outstanding Social Youth** of 2017 awarded by Chiayi City for exceptional community service and leadership.