

Chen-Han Lin

✉ allenchenhan99@gmail.com ☎ (+886) 978-261-955 🌐 <https://allenchenhan99.github.io/website/>
in Allen Lin 🌐 allenchenhan99

Education

National Yang Ming Chiao Tung University, College of Computer Science Sept 2023 – Present
Master's in Data Science and Engineering

- **Coursework:** Machine Learning, Deep Learning, Data Science Project, Biostatistics

National Cheng Kung University, College of Science Sept 2018 – June 2022
Bachelor's degree in Mathematics


- **Coursework:** Probability and Statistics, Fourier Analysis, Data Structures and Algorithms

Research Topic

Model Enhancement and Implementation of Human Pose Estimation Based on the Mamba State Space Architecture

- Advisor: Chin-Tien Wu
- Completion time: Dec 2025
- Enhanced a Video Pose Transformer (VPT) by integrating Mamba-based sequence blocks for GPU-accelerated 2D-to-3D human pose estimation. Reduced inference latency and training cost through CUDA-level optimization, parallelized data flow, and modular system integration.

Projects

Low-Rank Adaptation and Combined Parameter-Efficient Tuning for Large Models [Poster Link](#) 

- Investigated the integration of parameter-efficient fine-tuning (PEFT) methods such as LoRA, BitFit, Adapter, and Prefix-tuning across both vision and NLP tasks.
- Demonstrated that **LoRA + Adapter** achieved a **+3.6% improvement in accuracy** on vision classification tasks (from 0.8775 to 0.9134), with only **1.6% trainable parameters**.
- Showed that combining LoRA and BitFit improved NLP accuracy on RoBERTa-base from 0.931 to **0.945** (SST-2), confirming the complementary nature of these PEFT methods.

Data Science Project: Pima Indian Diabetes Dataset [Report Link](#) 

- Performed comprehensive data preprocessing and feature engineering, including handling missing values, outlier detection, imputation with median values, scaling, and creation of interaction and domain-specific features (e.g., glucose-to-insulin ratio) to enhance model performance.
- Built and compared multiple classification algorithms (logistic regression, decision tree, random forest, XGBoost) with stratified k-fold cross-validation and grid-search hyperparameter tuning.
- Evaluated models with ROC-AUC, precision, recall, F1-score, and calibration curves, achieving **Accuracy = 83.6%**, **ROC-AUC = 91.2%**, **Precision = 84.5%**, **Recall = 88.2%**, and **F1 = 86.3%** on the PIMA Diabetes dataset.

Simulating Oceanic Environments with OpenGL

- Built a real-time ocean scene in C++/OpenGL with dynamic wave animation, per-fragment lighting, and texture mapping.
- Simulated refraction using DuDv distortion and implemented a cubemap skybox for rendering.
- Applied Fresnel reflectance to blend reflections and refractions on water surfaces.

Game Theory Applied to Darts Strategies

- Analyzed professional match data and framed checkout/target selection as a game-theoretic decision process.
- Investigated the trade-offs between risk and reward in critical gameplay scenarios to model optimal strategies.

Competition Experience

TSMC IT CareerHack

Jan 2024

A fourth-place ranking

- Utilized the Llava base model with fine-tuning techniques using LoRA to improve task-specific performance.
- Applied data augmentation techniques to enhance dataset diversity and improve model generalization.
- Implemented image filtering to remove noisy data that could negatively impact model accuracy.
- Researched alternative base models, such as Flamingo and Honeybee, to potentially replace Llava for improved performance.

TSMC Intelligent Manufacturing Workshop

March 2025

Achieved top-performing team in scheduling optimization

- Explored multiple approaches, including Greedy, Dynamic Programming, Reinforcement Learning, and Genetic Algorithm, to address complex scheduling tasks.
- Recorded and analyzed combinations of methods to evaluate the distribution of possible solutions and identify convergence patterns.
- Formulated the problem as a graph-based job-shop optimization, converting locations to nodes and distances to weighted edges, and applied Integer Linear Programming (ILP) to obtain the optimal schedule.

AI CUP 2024 Image Data Generation for UAVs

May 2024 - June 2024

Honorable Award

- Processed and augmented training data with image normalization, and mask-conditioned augmentation.
- Tuned key hyperparameters to improve generation quality and evaluated models using Fréchet Inception Distance (FID).

TBrain AI GO Housing Price Prediction

Sept 2023 - Nov 2023

Ranked in top 4%

- Cleaned and merged multi-source datasets to improve model reliability.
- Ensembled LightGBM and XGBoost to boost regression accuracy.
- Optimized prediction accuracy through hyperparameter tuning and model evaluation on validation datasets.

AI GO Skyrocketed Stocks Forecasting

Mar 2025 – Apr 2025

- Conducted data cleaning by handling missing values and removing non-informative features.
- Applied correlation-based feature selection to identify the most relevant predictors for forecasting.
- Trained and hyperparameter-tuned ensemble models (XGBoost and Random Forest).
- Implemented TabNet, a deep learning architecture for tabular data, to capture nonlinear feature interactions and enhance model interpretability.

Relevant Skills

Programming Languages:

- C++, C, Python, SQL, Bash, Java, JavaScript
- Familiar with Linux development, scripting, and system designing

Software Engineering and Tools:

- Git, Docker, VirtualBox, MATLAB, GCP
- Debugging, version control, and performance profiling

Machine Learning and System Optimization:

- PyTorch, TabNet, Transformer, Mamba-based model optimization
- Model deployment, runtime profiling, CUDA acceleration, and Linux-based data pipelines.