Andrew Liao

+1 (347) 271-1164 | yl8520@nyu.edu | https://github.com/ALiao18 | https://medium.com/@oail.werdna

Education

New York University — New York, NY

Bachelor of Arts in Computer Science & Data Science — *Expected Dec 2025*

- Relevant Coursework: Natural Language Processing, **Advanced ML & DL**, Algorithms, Responsible DS, Machine Learning, Data Management & Analysis, **Causal Inference**, Computer Systems, **Linear Algebra**, **Probability & Statistics**, Discrete Mathematics
- Study Abroad: Prague, Czech Republic (Fall 2022), Paris, France (Fall 2024)

Research Experience

NYU Langone CN^3 Lab — New York, NY

Research Volunteer - Computational Neuroscience — *Jan 2025 - Present*

- Designed and implemented **febrile seizure prediction pipeline** on Family Interview Dataset **(348 samples, 279 features)** achieving **35% AUC improvement**, advised by Prof. Zhe (Sage) Chen
- Co-authoring abstract submission for seizure prediction research, demonstrating interdisciplinary expertise
- Presented research findings to panel of **2 leading experts in Febrile Seizure** at NYU Langone, translating complex ML concepts for clinical audience

Independent Research with Prof. Parijat Dube — New York, NY

Research Volunteer - Reinforcement Learning — *Feb 2025 - Present*

- Analyzing internal world model representations in **reinforcement learning systems** using causal approaches
- Implementing **Probabilistic Graphical Models (PGMs)** including Causal Structure Cognitive Graphs (CSCGs) on novel card-game environment
- Investigating whether RL models learn true world models or efficient heuristics through rigorous experimentation

Fine Ace Asset Management Co. — Taipei, Taiwan

Quantitative Analyst Intern — *May 2022 - Sep 2022*

- Led financial analysis of Tianli Offshore Wind Technology, synthesizing insights from investor presentations, quarterly earnings, and annual reports to deliver comprehensive Q2 update to **8 senior directors**
- Streamlined data reporting processes, reducing weekly report generation time by **40%** while enhancing analyst presentations with actionable insights

Technical Projects

MovieLM: Counterfactual Storytelling & Visual Storyboarding — *Spring 2025*

- Developing novel multi-modal framework for predicting plot progression and generating screenplays with storyboards of alternative plotlines for romance motion pictures
- Implementing romance genre-specific counterfactual modeling and character state embeddings for consistent long-form storytelling
- Leveraging **NYU's HPC cluster (Greene)** for diffusion model training and fine-tuning stable diffusion with DreamBooth for style transfer

Advanced CartPole Stabilization via Deep Reinforcement Learning — *Dec 2024*

- Designed and implemented an **Actor-Critic framework** in PyTorch, combining policy gradient optimization with Monte Carlo estimation to achieve a perfect **500/500 score** on the CartPole problem
- Engineered a novel state representation by decoupling acceleration dependencies while preserving the Markov property, achieving **475/500 performance** with minimal training

Optimized Click-Through-Rate Prediction — *Nov 2024*

- Developed sophisticated deep learning model for click-through rate prediction using embedding layers in PyTorch complemented by XGBoost to enhance performance, achieving **0.767 AUC (25.5% improvement)** over baseline
- Created advanced feature engineering pipeline leveraging **empirical Bayes estimation** and temporal optimizations, processing datasets with **50+ high-dimensional features**

NLP Projects (Information Retrieval; HMM & POS Tagging) — *Spring 2025*

- Designed and implemented an Information Retrieval system utilizing **TF-IDF vectorization** and **cosine similarity** scoring with query expansion, achieving **85%+ precision** on test dataset
- Built Hidden Markov Model (HMM) Part-of-Speech tagger using the **Viterbi algorithm**, handling out-of-vocabulary words through strategic likelihood assignment with **95% accuracy**

Technical Skills

Languages: Python, R, SQL, Java, LaTeX | **ML/DL Frameworks**: PyTorch, Scikit-learn, NumPy, Pandas, Matplotlib | **Research Areas**: Reinforcement Learning, Computational Neuroscience, Causal Inference, PGMs