Chih-Yuan (Edward) Chang

New York (Open to relocate) +1 (646) 995-8648 | cc5375@columbia.edu | linkedin.com/edwardchang716 | Portfolio

EDUCATION

Columbia University *New York, NY*
Master of Science in Data Science *Aug 2025 - Dec 2026 (expected)*
Courses: Machine Learning, Algorithm Design and Analysis, Exploratory Data Analysis

National Taiwan University (NTU) *Top1 university in Taiwan* *Taipei, Taiwan*
BS in Environmental Engineering, BS in Psychology**,** GPA 3.8/4.0
Courses: Python, C/C++, Data structures, Natural Language Processing, Artificial Neural Networks, Experiment Design

SKILLS

Programming: Python, C/C++, JavaScript, R, SQL, MATLAB, Git/Github
Machine Learning & AI: TensorFlow, PyTorch, Keras, scikit-learn, Hugging Face, OpenPose, LLMs, NLP
Analytics & Statistics: Pandas, NumPy, Matplotlib/Seaborn, NLTools, statsmodels, Pymer4/lme4, FactorAnalyzer

EXPERIENCE

CAYIN Technology *Worldwide top 11 digital signage software provider* *Taipei, Taiwan*
AI Engineering Intern *Jul 2024 - Aug 2024*

- Initiated an LLM-powered generative AI project using **OpenAI API**, enabling customers to generate advertising images and text via prompts, cutting manual design effort by ~70%
- Developed and deployed a web application (**Python**, HTML/CSS, JavaScript) to showcase AI-generated ad copies and templates for the sales team; later collaborated with R&D to integrate into CAYIN's flagship software Poster
- Identified limitations in the company's initial Canva API plan and independently proposed an alternative solution, which led to the successful implementation of the above generative AI project

Computational Human-sociocultural Experimental Neuroscience Lab, NTU *Taipei, Taiwan*
NLP Research Assistant *Jan 2023 - Jan 2025*

- Applied NLP (**NLTK**), LLMs (**OpenAI API, Transformers**) and ML (**TensorFlow, scikit-learn**) models to analyze verbal, nonverbal, and questionnaire data, achieving 80%+ accuracy on stance classification
- Initiated an A/B test experiment (**Pandas, NumPy, OpenAI API, Transformers**) revealing significant effects of text length on embedding similarity ($p < .001$), presented at the largest psychology conference in Taiwan
- Co-authored a study on mental representations of friendship and well-being, applying statistical modeling (**Pandas, NumPy**, NLTools, Pymer4/lme4); manuscript under submission as first author
- Led the NLP team and mentored undergraduates on LLMs, data analysis and Python coding practices, collaborated with lab members to publish and present findings at conferences

Water Resources and Hydroinformatics System Lab, NTU *Taipei, Taiwan*
Deep Learning Research Assistant *Sep 2022 - Aug 2023*

- Proposed an Encoder-Decoder LSTM framework (**PyTorch**) for time series water quality prediction, awarded the College Student Research Scholarship by Taiwan's National Science and Technology Council (~9,000 applicants)
- Scraped and preprocessed datasets (**Pandas, NumPy**, statsmodels) around 3 years of hourly data, performed Random Forest (**scikit-learn**) for feature selection, and applied MICE-RF imputation to address high missing-data rates
- Designed and optimized an LSTM-ED model, achieving $R^2 > 0.9$ and extending prediction horizons beyond prior short-term limits; validated predictive stability under extreme weather scenarios

PROJECTS

Financial Argument Mining with LLMs, NTU *Spring 2024*

- Implemented feature-based embeddings with Random Forest and fine-tuned GPT-3.5/BERT (**PyTorch, scikit-learn**) for argumentative relation classification, with BERT achieving weighted F1 of 0.72, surpassing strong baseline
- Applied in-context learning and fine-tuned GPT-3.5/LLaMA-3 for stance and relation classification, reaching F1 up to 0.77 and outperforming prompting methods