

# Chih-Yuan (Edward) Chang

New York (Open to relocate) +1 (646) 995-8648 | [cc5375@columbia.edu](mailto:cc5375@columbia.edu) | [linkedin.com/edwardchang716](https://www.linkedin.com/in/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 Analysis, Data structures, Statistics, Experiment Design, Deep Learning

## SKILLS

---

**Programming:** Python, R, SQL, MATLAB, C/C++, JavaScript, Git/Github

**Analytics & Statistics:** Pandas, NumPy, Matplotlib/Seaborn, NLTools, statsmodels, Pymer4/lme4, FactorAnalyzer, PCA

**Machine Learning & AI:** TensorFlow, PyTorch, Keras, scikit-learn, Hugging Face, LLMs, NLP, RAG, LangChain

## EXPERIENCE

---

### CAYIN Technology Worldwide top 11 digital signage software provider

Taipei, Taiwan

#### Data Scientist Intern

Jul 2024 - Aug 2024

- Proposed and implemented an LLM-based solution to replace an ineffective API plan, reducing manual advertising design effort by ~60%
- Developed a prototype (Python, JavaScript, OpenAI API) to generate ad copies, collaborated with the R&D team to integrate into the company's software and received 80%+ positive client feedback

### Computational Human-sociocultural Experimental Neuroscience Lab, NTU

Taipei, Taiwan

#### Data Science Research Assistant

Jan 2023 - Jan 2025

- Conducted large-scale analysis of conversational and questionnaire datasets using NLP (NLTK), ML (scikit-learn), and statistical models (Python, R, statsmodels, Pymer4/lme4), achieving 80%+ accuracy on stance classification
- Designed and executed an A/B test experiment (Pandas, NumPy, OpenAI API, Transformers), uncovering significant effects of text length on embedding similarity ( $p < .001$ ); results presented at Taiwan's largest psychology conference
- 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; built an automated data preprocessing pipeline (Pandas, NumPy) that reduced data cleaning time by ~80%

### Water Resources and Hydroinformatics System Lab, NTU

Taipei, Taiwan

#### Deep Learning Research Assistant

Sep 2022 - Aug 2023

- Conducted large-scale time series analysis (~20,000 hourly water quality records) using Python, R and MATLAB, providing insights for environmental monitoring and resource management; research awarded the College Student Research Scholarship (9,000+ applicants)
- Scraped and preprocessed datasets (Pandas, NumPy, statsmodels) ~ 20,000 hourly records, 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 (PyTorch), achieving  $R^2 > 0.9$  and extending prediction horizons beyond prior short-term limits; validated predictive stability under extreme weather scenarios

## PROJECTS

---

### Consumer Sentiment Analysis of HiMirror, NTU

Fall 2023

- Developed a pipeline to scrape, translate, and preprocess 3K+ Amazon product reviews for cross-lingual text consistency (Python, Pandas, BeautifulSoup)
- Conducted sentiment analysis and keyword extraction (Transformers, SpaCy, NLTK) to identify drivers of satisfaction and dissatisfaction, visualized patterns through word clouds, and delivered actionable insights for product improvement